Ocean Dynamics (2004) 54: 8–26
DOI 10.1007/s10236-003-0054-7

Kirsten Zickfeld · Thomas Slawig · Stefan Rahmstorf

A low-order model for the response of the Atlantic thermohaline circulation to climate change

Received: 20 January 2003 / Accepted: 30 June 2003

Abstract Concern has been expressed that anthropogenic climate change may lead to a slowdown or even collapse of the Atlantic thermohaline circulation (THC). Because of the possibly severe consequences that such an event could have on the northern North Atlantic and northwestern Europe, integrated assessment models (IAMs) are needed to explore the associated political and socioeconomic implications. State-of-the-art climate models representing the THC are, however, often too complex to be incorporated into an integrated assessment framework. In this paper we present a low-order model of the Atlantic THC which meets the main requirements of IAMs: it (1) is physically based, (2) is computationally highly efficient, (3) allows for comprehensive uncertainty analysis and (4) can be linked to globally aggregated climate models that are mostly used in IAMs. The model is an interhemispheric extension of the seminal Stommel model. Its parameters are determined by a least-squares fit to the output of a coupled climate model of intermediate complexity. Results of a number of transient global warming simulations indicate that the model is able to reproduce many features of the behaviour of coupled ocean–atmosphere circulation models such as the sensitivity of the THC to the amount, regional distribution and rate of climate change.

Keywords Thermohaline circulation · Box model · Climate change · Sensitivity analysis

Responsible Editor: Richard Greatbatch

K. Zickfeld (✉) · S. Rahmstorf
Potsdam Institute for Climate Impact Research (PIK),
PO Box 60 12 03, 14412 Potsdam, Germany
e-mail: zickfeld@pik-potsdam.de

T. Slawig
Technical University of Berlin, Institute for Mathematics,
Straße des 17. Juni 136, 10623 Berlin, Germany

1 Introduction

The Atlantic thermohaline circulation (THC) transports large amounts of heat northward (up to 10^{15} W), contributing a major part to the heat budget of the North Atlantic region (Ganachaud and Wunsch 2000). Paleoreconstructions (Dansgaard et al. 1993), theoretical considerations (Stommel 1961) and model simulations (Manabe and Stouffer 1993; Stocker and Schmittner 1997; Rahmstorf and Ganopolski 1999a) suggest that the THC behaves non-linearly, with a well-defined threshold beyond which the circulation shuts down. Manabe and Stouffer (1993), for example, have simulated a complete shutdown of the THC by quadrupling atmospheric CO_2 concentrations. The consequences of a collapse of the THC for the climate of the North Atlantic region are highly uncertain. The reason is that estimates of the magnitude and the exact location of the cooling anomaly associated with a reduction of the oceanic heat transport differ widely between models. Vellinga and Wood (2002), Ganopolski et al. (2001) and Manabe and Stouffer (1988), for example, simulated a maximum mean annual cooling of 6–12 °C. While in Ganopolski et al. (2001) and Manabe and Stouffer (1988) the maximum cooling is located over the western North Atlantic, in Vellinga and Wood (2002) it is centred over the eastern North Atlantic. Also, it is unclear to what extent the cooling over the North Atlantic could be compensated locally by greenhouse warming (Houghton et al. 2001). Model simulations, however, reveal that a greenhouse gas-induced collapse of the THC could lead to a sudden reversal of the warming trend and a cooling of 2–3 °C (relative to the peak warming) within a few decades, touching northwestern Europe and northeastern America (Rahmstorf and Ganopolski 1999a).

Because of its non-linearity and the possibly severe consequences that its collapse could cause, the THC has become a prominent example for a class of events called ''climate surprises'' (Schneider et al. 1998) or ''large-scale

discontinuities'' (McCarthy et al. 2001). These refer to abrupt and severe events of regional or global-scale triggered in the complex Earth system in response to smooth forcing such as, for example, the buildup of greenhouse gases in the atmosphere. The recognition that the evolution of the climate system under global warming could feature such non-linear behaviour has led several scientists to propose that climate policy analysis and the tools it relies upon – so-called integrated assessment models (IAMs) – should take into account climate surprises (Schellnhuber 1997; Schneider and Thompson 2000).

IAMs usually cover the entire causal chain from socioeconomic activities giving rise to greenhouse gas emissions, to concentrations, climate, and impacts – the reason why they often rely on simplified representations of the subsystems involved. Furthermore, a broad class of IAMs is run in an optimization (rather than prediction) mode, which adds further demands on simplicity and computational efficiency. For these reasons, state-of-the-art models representing the THC, i.e. three-dimensional ocean general circulation models (OGCMs), as well as models of intermediate complexity (e.g. zonally averaged ocean models) are often too complex to be incorporated into integrated assessment frameworks.

Efforts to include the non-linearity of the THC in IAMs have relied mainly on reduced-form approaches. Tóth et al. (1998), Keller et al. (2000) and Nordhaus (2000) use the stability diagram of the THC determined by Stocker and Schmittner (1997) as a static constraint on emission behaviour. Mastrandrea and Schneider (2001) use a dynamic box model, the ''Simple Climate Demonstrator'' (Schneider and Thompson 2000). This model represents the Northern Hemisphere only, which is a limitation, given that the THC is known to be an interhemispheric system (Rahmstorf 1996; Thorpe et al. 2001).

The scope of this paper is the development of a very simple Atlantic Ocean component which (1) is physically based, (2) reflects the basic structure of the THC, (3) reproduces key features of the behaviour of state-of-the-art climate models, (4) is computationally highly efficient, (5) can be linked to globally aggregated atmospheric models such as, for example, energy balance (Harvey 1988) or impulse response function models (Hooss et al. 2001) and (6) allows for comprehensive uncertainty analysis. Clearly a model as simple as the one presented here cannot do justice to the complexity of the real ocean. Nevertheless, we hope to demonstrate in this paper that it can still be a useful tool.

The paper is organized as follows: in Section 2 we describe the model setup in detail. Section 3 discusses the procedure by which the box model is fitted to a state-of-the-art climate model of intermediate complexity, the CLIMBER-2 model (Petoukhov et al. 2000; Ganopolski et al. 2001). The coupling of the box model to a globally aggregated climate model is described in Section 4. Section 5 characterizes the box model response to a specific global mean temperature scenario for a wide range of parameter settings and compares it to the behaviour of comprehensive coupled ocean–atmosphere models. Section 6 explores the sensitivity of the box model response on additional physical processes such as vertical and horizontal mixing. Finally, Section 7 draws some conclusions.

2 Model description

The model is a four-box extension of the classic Stommel model (Stommel 1961) – a conceptual model that has already been successfully applied for the investigation of bifurcations and the stability of the Atlantic thermohaline circulation (Rahmstorf 1996; Rahmstorf and Ganopolski 1999b; Scott et al. 1999; Titz et al. 2002). For the purpose of this paper, the model is extended to allow for the consideration of different box volumes and is used in the time-dependent (or ''transient'') mode.

The model configuration is shown in Fig. 1. It consists of four well-mixed boxes, representing the southern, tropical, northern and deep Atlantic, respectively. Neighbouring boxes are connected to allow for a continuous, closed-loop circulation. The surface boxes exchange heat and freshwater with the overlying atmosphere. These surface fluxes render the water in the northern box denser than that in the southern box, giving rise to a pressure-driven circulation with northward flow near the surface and southward flow at depth. This picture reflects the view of the Atlantic THC as a cross-hemispheric circulation, with sinking in the North Atlantic and upwelling in the Southern Ocean, driven by thermohaline forcing (Rahmstorf 1996; Rahmstorf and England 1997). It should be noted that this view is challenged by some authors arguing that it is the winds in the Southern Ocean which drive the Atlantic overturning circulation (Toggweiler and Samuels 1995, 1998). Their arguments, however, are controversial (cf. e.g. Rahmstorf and England 1997).

In this four-box model the meridional volume transport m (or ''overturning'') is proportional to the density difference $\rho_2 - \rho_1$ between boxes 1 and 2:

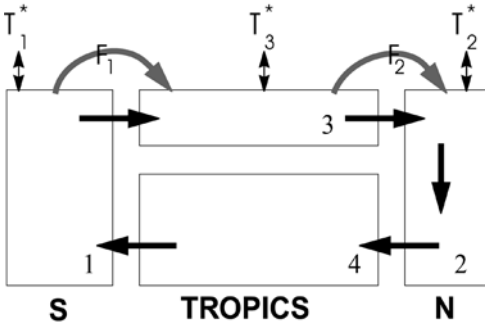

Fig. 1 Schematic of the four-box model of the Atlantic thermohaline circulation. The temperatures of boxes *1*, *2* and *3* are relaxed toward the values T_1^*, T_2^* and T_3^*, respectively. The salinities are forced by the freshwater fluxes F_1 and F_2. The meridional flow (*black arrows*) is proportional to the density gradient between boxes 1 and 2

$$m = \frac{k(\rho_2 - \rho_1)}{\rho_0} = k[\beta(S_2 - S_1) - \alpha(T_2 - T_1)] \ , \qquad (1)$$

where $S_2 - S_1$ and $T_2 - T_1$ are the north–south salinity and temperature gradients, respectively, k is a hydraulic constant linking volume transport m to the density difference, ρ_0 is the reference density, α and β are thermal and haline expansion coefficients. Note that the rightmost side of Eq. (1) is obtained by presuming a linear equation of state.

The temperatures and salinities of the four boxes adjust to the oceanic transport of heat and freshwater. Further, temperatures and salinities of the surface boxes are forced by the overlying atmosphere. The surface heat fluxes are computed from the condition $Q = \Gamma(T^* - T)$. Note that this formula, although having the same form as classical Haney restoring (Haney 1971), has a different interpretation: the thermal coupling constant Γ subsumes the radiative relaxation constant and a constant related to atmospheric heat diffusion, and T^* defines the oceanic temperature in the absence of heat transport, rather than the surface air temperature (Rahmstorf and Willebrand 1995). Though extremely simple, this approach allows us to capture the feedback of the oceanic heat transport on thermal forcing.

The freshwater fluxes at the ocean surface consist of prescribed freshwater transports between the upper boxes. Note that the latter represent not only atmospheric water vapour transport but also wind-driven oceanic transport. This is the reason for the asymmetry in the freshwater fluxes F_i: the equatorward transport of freshwater by the wind-driven circulation more than compensates for the loss of freshwater due to evaporation, and hence the flux F_1 is directed into the Atlantic catchment (Rahmstorf 1996).

This results in the following set of ordinary differential equations for temperatures T_i and salinities S_i of the four boxes:

$$\dot{T}_1 = \frac{m}{V_1}(T_4 - T_1) + \lambda_1(T_1^* - T_1) \qquad (2)$$

$$\dot{T}_2 = \frac{m}{V_2}(T_3 - T_2) + \lambda_2(T_2^* - T_2) \qquad (3)$$

$$\dot{T}_3 = \frac{m}{V_3}(T_1 - T_3) + \lambda_3(T_3^* - T_3) \qquad (4)$$

$$\dot{T}_4 = \frac{m}{V_4}(T_2 - T_4) \qquad (5)$$

$$\dot{S}_1 = \frac{m}{V_1}(S_4 - S_1) + \frac{S_0 F_1}{V_1} \qquad (6)$$

$$\dot{S}_2 = \frac{m}{V_2}(S_3 - S_2) - \frac{S_0 F_2}{V_2} \qquad (7)$$

$$\dot{S}_3 = \frac{m}{V_3}(S_1 - S_3) - \frac{S_0(F_1 - F_2)}{V_3} \qquad (8)$$

$$\dot{S}_4 = \frac{m}{V_4}(S_2 - S_4) \ . \qquad (9)$$

Here V_i are box volumes, λ_i thermal coupling constants, and T_i^* the temperatures towards which the southern, northern and tropical boxes are relaxed. F_1 and F_2 are the freshwater fluxes (multiplied by a reference salinity, S_0, for conversion to a salt flux) into the tropical and northern Atlantic, respectively. Since the overturning m depends on both temperature T_i and salinity S_i, the differential equations are non-linear. Note that in the above system of differential equations we have used upwind differencing assuming a positive (northward at the surface) meridional flow. For negative flow, the finite-difference form changes, although this is not of concern to this work since we are mainly interested in the evolution of today's THC.

2.1 Steady-state solutions

We now consider equilibrium conditions: then according to Eqs. (5) and (9) the deep box plays no role ($T_4 = T_2, S_4 = S_2$), and the steady-state equations for temperature and salinity read:

$$\frac{m}{V_1}(T_2 - T_1) = \lambda_1(T_1 - T_1^*) \qquad (10)$$

$$\frac{m}{V_2}(T_3 - T_2) = \lambda_2(T_2 - T_2^*) \qquad (11)$$

$$\frac{m}{V_3}(T_1 - T_3) = \lambda_3(T_3 - T_3^*) \qquad (12)$$

$$m(S_2 - S_1) = -S_0 F_1 \qquad (13)$$

$$m(S_3 - S_2) = S_0 F_2 \qquad (14)$$

$$m(S_1 - S_3) = S_0(F_1 - F_2) \ . \qquad (15)$$

The combination of Eqs. (1) and (13) yields a quadratic equation in m:

$$m^2 + k\alpha(T_2 - T_1)m + k\beta S_0 F_1 = 0 \ . \qquad (16)$$

The north–south temperature gradient $T_2 - T_1$ can be expressed as a function of the parameters by combining Eqs. (10–12):

$$T_2 - T_1 = (T_2^* - T_1^*) + \frac{P_1^2(m)}{P_2^2(m)} \ , \qquad (17)$$

where the P_i^2 are 2nd-order polynomials in m:

$$P_1^2 = \left(\frac{\lambda_1\lambda_3}{V_2}(T_3^* - T_2^*) - \frac{\lambda_2\lambda_3}{V_1}(T_2^* - T_1^*)\right)m - \left(\frac{\lambda_1}{V_2V_3} + \frac{\lambda_2}{V_1V_3} + \frac{\lambda_3}{V_1V_2}\right)(T_2^* - T_1^*)m^2 \qquad (18)$$

$$P_2^2 = \left(\frac{m}{V_1} + \lambda_1\right)\left(\frac{m}{V_2} + \lambda_2\right)\left(\frac{m}{V_3} + \lambda_3\right) - \frac{m^3}{V_1V_2V_3} \ . \qquad (19)$$

Equation (16) in combination with Eq. (17) results in a 4th-order equation in m. The structure of the physical solutions is identical to that discussed in Rahmstorf (1996) and Titz et al. (2002) for prescribed box

temperatures T_i. For present-day climate conditions, i.e. $F_1 > 0$ and $T_1 > T_2$, two stable solution exists: an "on" mode with a vigorous overturning ($m > 0$) and a "reverse" mode with sinking of water masses in the southern and upwelling in the northern box ($m < 0$). If F_1 is increased beyond a critical threshold (the "bifurcation point"), the on mode becomes unstable and the only stable solution is the reverse mode.

Equation (16) indicates that the equilibrium overturning depends only on F_1, not on F_2. This is due to the fact that under steady-state conditions it makes no difference whether the freshwater enters the northern box via atmospheric transport or oceanic advection. F_2 merely determines the salinity gradient between the tropical and the northern boxes. In the transient case, however, it affects the response of the overturning, as on finite time scales atmospheric freshwater transport is more effective in freshening the northern box than oceanic advection. This implies that a large transient freshwater forcing F_2 may trigger a complete shutdown of the circulation. For an in-depth discussion of the role of the freshwater fluxes F_i and the stability properties of the four-box model cf. Rahmstorf (1996) and Titz et al. (2002).

As initial values for the transient experiments we use the steady-state salinities S_i^{eq} and temperatures T_i^{eq} of the single boxes. The latter can be derived analytically from Eqs. (10–12) and Eq. (17):

$$T_1^{eq} = \frac{m_{eq}^2 \lambda_3 T_3^*}{V_1 V_2 P_2^2(m_{eq})} + \frac{\left(\frac{m_{eq}}{V_3} + \lambda_3\right)\left(\frac{m_{eq}}{V_1}\lambda_2 T_2^* + \left(\frac{m_{eq}}{V_2} + \lambda_2\right)\lambda_1 T_1^*\right)}{P_2^2(m_{eq})}$$

$$T_2^{eq} = \frac{m_{eq}^2 \lambda_1 T_1^*}{V_2 V_3 P_2^2(m_{eq})} + \frac{\left(\frac{m_{eq}}{V_1} + \lambda_1\right)\left(\frac{m_{eq}}{V_2}\lambda_3 T_3^* + \left(\frac{m_{eq}}{V_3} + \lambda_3\right)\lambda_2 T_2^*\right)}{P_2^2(m_{eq})}$$

$$T_3^{eq} = \frac{m_{eq}^2 \lambda_2 T_2^*}{V_1 V_3 P_2^2(m_{eq})} + \frac{\left(\frac{m_{eq}}{V_2} + \lambda_2\right)\left(\frac{m_{eq}}{V_3}\lambda_1 T_1^* + \left(\frac{m_{eq}}{V_1} + \lambda_1\right)\lambda_3 T_3^*\right)}{P_2^2(m_{eq})},$$

where P_2^2 is defined as above. m_{eq} is the equilibrium overturning for the present-day value of F_1. The steady-state salinity gradients are derived from Eqs. (13–15) as:

$$S_2^{eq} - S_1^{eq} = -\frac{S_0 F_1}{m_{eq}}$$

$$S_3^{eq} - S_2^{eq} = \frac{S_0 F_2}{m_{eq}}$$

$$S_1^{eq} - S_3^{eq} = \frac{S_0 (F_1 - F_2)}{m_{eq}}.$$

3 Determination of the model parameters

As stated in the Introduction, the goal of our work is the development of a tool which is able to reproduce key features of the behaviour of more comprehensive models. As reference we have chosen CLIMBER-2, a coupled ocean–atmosphere model of intermediate complexity, which has been proved to successfully describe crucial elements of the climate system, including the THC (Petoukhov et al. 2000; Ganopolski et al. 2001). The advantage of this model lies in its computational efficiency, which allows performing a multitude of runs in a manageable time frame. Our intention is to tune the box model parameters such that not only the present-day state of the Atlantic ocean as simulated by CLIMBER-2 is reproduced, but also key features of its dynamic behaviour such as, for example, the location of the bifurcation point of the THC. We therefore determine part of the unknown parameters (i.e. the restoring temperatures T_1^*, T_2^* and T_3^*, the flow constant k and the thermal coupling constant Γ) by a least-squares fit of the equilibrium solution $m(F_1)$ of the box model to a hysteresis experiment performed with CLIMBER-2 (cf. Fig. 2). We impose additional constraints on the optimal parameters by requiring equilibrium temperatures T_i^{eq} and overturning m_{eq} to be equal to the present-day CLIMBER-2 values (cf. Table 2). Since the latter are averaged over the corresponding regions in the Atlantic, the extents of the southern, tropical and northern boxes (i.e. V_i and z_i, $i \in \{1, 2, 3\}$) have to be fixed beforehand and are not part of the optimization.

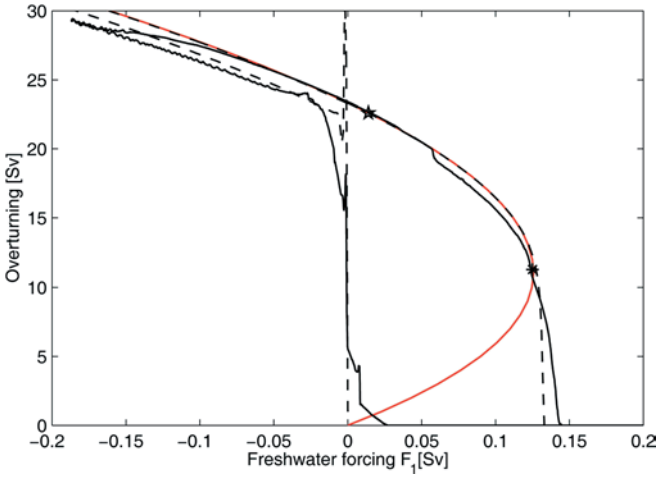

Fig. 2 Positive branch of the equilibrium solutions $m(F_1)$ and hysteresis of the box model (*red* and *dashed curves*, respectively) and hysteresis derived with CLIMBER-2 (*solid curve*). The *star* and the *asterisk* denote the present-day overturning m_{eq} and the bifurcation point (F_1^{crit}, m_{crit}) = (0.125 Sv, 11.3 Sv), respectively. The hysteresis curves are traced by slowly increasing and subsequently decreasing the freshwater flux into the Atlantic south of 50°N. The freshwater perturbation is applied at a rate of 0.05×10^{-3} Sv a^{-1} following the method described in Rahmstorf (1995). Note that in tracing the hysteresis of the box model the transition to the reverse circulation mode is inhibited by setting m to zero for $m < 0$

It should be noted that for a given model the bifurcation diagram (i.e. the curve depicting the asymptotic response of a system against some parameter) differs conceptually from the hysteresis curve. In fact, the hysteresis is traced by diagnosing the transient response of the system to a transiently varying parameter (cf. legend to Fig. 2). However, if the rate at which the parameter is varied is slow enough, the system can be assumed to be in quasi-equilibrium such that the hysteresis curve is close to the asymptotic response. Figure 2 demonstrates that in the case of the box model the hysteresis even coincides with the equilibrium curve for subcritical values of the freshwater forcing F_1. In CLIMBER-2, the equilibrium value of the overturning for some fixed values of the freshwater perturbation is found to lie slightly below the quasi-equilibrium response, the distance to the latter increasing in the proximity of the bifurcation point (T. Schneider, personal communication).

The extents of the surface boxes are chosen to roughly reflect the distribution of observed water masses in the Atlantic ocean. The tropical box extends from 30°S to 45°N, representing low-latitude water above the main thermocline in the Atlantic at ~1000 m depth. The northern box extends from 45–70°N, corresponding to the formation region of North Atlantic Deep Water (NADW). This water spreads southward at ~3000 m depth and joins the water masses of the Antarctic Circumpolar Current at approximately the latitudes of the Cape of Good Hope (30°S). This is where the analogy between the box model and the real Atlantic ends. The southern box is therefore somewhat arbitrarily chosen to extend from 60 to 30°S.

The remaining parameters, i.e. the freshwater transport F_2 and the volume of the deep box V_4, do not affect the equilibrium solutions of the box model (cf. Eqs. 16 and 17) and therefore cannot be constrained by the optimization procedure outlined below. F_2 is derived by requiring a salinity gradient between the tropical and the northern box close to the CLIMBER-2 value. V_4 is determined such that the critical freshwater forcing in the transient case corresponds to the CLIMBER-2 value (cf. Sect. 5). Note that this results in a relatively small value for V_4. This is not unrealistic, since only the portion of the deep Atlantic adjacent to the American continental slope is involved in the southward transport of NADW.

The individual coupling constants λ_i can be computed from the thermal coupling constant Γ and the respective box thickness z_i:

$$\lambda_i = \frac{\Gamma}{c\rho_0 z_i} \quad ,$$

where c is the specific heat capacity of seawater and ρ_0 its density.

The fitting procedure itself is performed by applying a mathematical optimization method to the weighted least-squares cost functional

$$J(\mathbf{m},\mathbf{x}) := \frac{1}{2}\sum_{i=1}^{N} \omega_i (m_i - \bar{m}_i)^2 \quad . \tag{20}$$

Here $\mathbf{x} = (T_1^*, T_2^*, T_3^*, k, \Gamma)$ is the vector of parameters to be adjusted, $(F_{1i}, \bar{m}_i)_{i=1,\ldots,N}$ are data from the CLIMBER-2 hysteresis curve, $\mathbf{m} \in \mathbb{R}^N$ is a vector of approximations m_i to the equilibrium solutions $m(F_{1i}, \mathbf{x})$ of the box model for given F_{1i} and $\mathbf{x}$, and the ω_i are non-negative weights used to enforce a closer fit in certain regions of the hysteresis curve, for example near the bifurcation point.

Each equilibrium solution m of the box model has to satisfy an equation

$$G(F_1, m, \mathbf{x}) = 0 \quad , \tag{21}$$

which is a 4th-order polynomial derived from Eqs. (16) and (17). To realize the coupling between F_{1i}, m_i, and $\mathbf{x}$ we thus have to impose Eq. (21) for all $i = 1, \ldots, N$ as additional constraints. Since Eq. (21) can only be solved approximately, we replace it by the least-squares inequality

$$|G(F_{1i}, m_i, \mathbf{x})|^2 \leq \delta_i \quad . \tag{22}$$

Here, the δ_i can be used to adjust the desired accuracy point-wise. This is crucial, since a very high prescribed accuracy for m_i may cause the optimization to fail because it stops at parameter values where some of the constraints are not satisfied.

The equality constraints for T_2^{eq} and T_3^{eq} and the inequality constraint for T_1^{eq} (cf. Table 2) are written as

$$c_j^{eq}(m_{eq}, \mathbf{x}) = 0, \quad j = 1, 2 \quad , \tag{23}$$

$$c_j(m_{eq}, \mathbf{x}) \leq 0, \quad j = 1, 2 \quad , \tag{24}$$

using non-linear functions c_j^{eq}, c_j.

We thus end up with the minimization problem

$$\min_{(\mathbf{m},\mathbf{x})} J(\mathbf{m},\mathbf{x}) \text{ subject to Eqs. } (22), (23), (24)$$

and additional bounds on m_{eq} and Γ.

This problem is solved by the routine `fmincon` from MATLAB's optimization toolbox (distributed by The Mathworks Inc., Natick, MA, USA). It incorporates the constraints by introducing Lagrange multipliers.

The method makes use of provided sensitivity information of cost functional and constraints with respect to variations in $\mathbf{m}$ and $\mathbf{x}$. The computation of the sensitivities of the cost functional J is straightforward; those of the more complex non-linear constraints are computed exactly, i.e without any approximation errors, by the software tool ADMAT (Verma 2002). This tool algorithmically computes the derivatives along with the function value itself by evaluating G, c_j^{eq}, and c_j for given F_{1i}, m_i and $\mathbf{x}$.

The optimal fit to the CLIMBER-2 hysteresis is shown in Fig. 2 and the parameters of the box model are summarized in Table 1. The equilibrium mass transport, salinities and temperatures resulting from these parameters are shown in Table 3. For comparison,

Table 1 Parameters of the box model. Note that the model is calibrated such that the equilibrium overturning $m_{eq} = m(F_1^{eq})$ is given for a freshwater flux F_1^{eq} equal to 0.014 Sv, which is the present-day value diagnosed in the CLIMBER-2 model

Constants		
c	4000 J kg^{-1} °C^{-1}	Specific heat capacity of seawater
ρ_0	1025 kg m^{-3}	Density of seawater
α	1.7×10^{-4} °C^{-1}	Thermal expansion coefficient
β	8×10^{-4} psu^{-1}	Haline expansion coefficient
S_0	35 psu	Reference salinity
Fixed parameters		
V_1	1.1×10^{17} m^3	Volume of the southern box
V_2	0.4×10^{17} m^3	Volume of the northern box
V_3	0.68×10^{17} m^3	Volume of the tropical box
V_4	0.05×10^{17} m^3	Volume of the deep box
z_1	3000 m	Depth of the southern box
z_2	3000 m	Depth of the northern box
z_3	1000 m	Depth of the tropical box
F_2	0.065 Sv	Freshwater transport from tropical to northern box
Tunable parameters		
T_1^*	6.6 °C	Relaxation temperature of the southern box
T_2^*	2.7 °C	Relaxation temperature of the northern box
T_3^*	11.7 °C	Relaxation temperature of the tropical box
Γ	7.3×10^8 J a^{-1} m^{-2} °C^{-1}	Thermal coupling constant
k	25.4×10^{17} m^3 a^{-1}	Empirical flow constant

Table 2 Parameters and constraints used in the optimization. Note that the constraint for T_1^{eq} is given as inequality since the CLIMBER-2 value is not known exactly (cf. Table 3)

Elements of the optimization	
Parameters: T_1^*, T_2^*, T_3^*, Γ, k	
Constraints:	
$\lvert G(F_{1i}, m_i, \vec{\mathbf{x}})\rvert^2 \leq \delta_i$	Match CLIMBER-2 hysteresis
$5 < T_1^{eq} < 7$ °C, $T_2^{eq} = 4.7$ °C, $T_3^{eq} = 11.4$ °C	Match CLIMBER-2 present-day oceanic temperatures
$m_{eq} = 22.6$ Sv	Match CLIMBER-2 present-day overturning
$10 < \Gamma < 75$ W m^{-2}	Keep Γ in realistic range

Table 3 Equilibrium values of the box model and present-day values of the CLIMBER-2 model (averaged over corresponding regions). Note that the values for T_1^{eq} and S_1^{eq} are determined only in approximate terms from CLIMBER-2 since it is unclear what the exact analogy to the southern box is

Equilibrium values			
	Box model	CLIMBER-2	
T_1^{eq}	6.5 °C	5 to 7 °C	Equil. temp. of the southern box
T_2^{eq}	4.7 °C	4.7 °C	Equil. temp. of the northern box
T_3^{eq}	11.4 °C	11.4 °C	Equil. temp. of the tropical box
$S_1^{eq} - S_2^{eq}$	−0.02 psu	−0.3 to 0 psu	Equil. salinity gradient
$S_3^{eq} - S_2^{eq}$	0.1 psu	0.1 psu	Equil. salinity gradient
$S_3^{eq} - S_1^{eq}$	0.08 psu	0.1 to 0.4 psu	Equil. salinity gradient
m_{eq}	22.6 Sv	22.6 Sv	Equil. overturning

the present-day values of the CLIMBER-2 model are also given.

4 Heat and freshwater forcing

Our purpose is the development of a low-order model of the Atlantic THC which can be driven by scenarios of global mean temperature (GMT) change. The reason is that the latter is the output of the globally aggregated climate models mostly used in integrated assessment frameworks. We therefore need an appropriate procedure for downscaling changes in GMT into basin-wide patterns of changes in restoring temperatures ΔT_i^* and net freshwater fluxes ΔF_i which force the box model. A common method for the efficient construction of regionally explicit climate-change projections is the so-called scaled scenario approach (Smith and Pitts 1997; Mitchell et al. 1999). It describes future climate change by scaling spatial patterns of climate anomalies by the respective global mean temperature change ΔT^{GL}. Following this approach, changes in restoring temperatures ΔT_i^* evolve in accordance with:

$$\Delta T_i^*(t) = p_i \Delta T^{GL}(t), \quad i \in \{1, 2, 3\} \ , \tag{25}$$

with constant values p_i. For determining the freshwater forcing, we take advantage of the fact that in CLIMBER-2, as well as in other models, changes in the meridional atmospheric water vapour transports are approximately proportional (Manabe and Stouffer 1994; Ganopolski et al. 2001) to the mean temperature change in the Northern and Southern Hemispheres, i.e. ΔT^{NH} and ΔT^{SH}, respectively:

$$\Delta F_1(t) = h_1 \Delta T^{SH}(t) = h_1 p_{SH} \Delta T^{GL}(t) \ , \quad (26)$$

$$\Delta F_2(t) = h_2 \Delta T^{NH}(t) = h_2 p_{NH} \Delta T^{GL}(t) \ . \quad (27)$$

The constant factors p_i, p_j and h_k are derived from greenhouse gas simulations with the CLIMBER-2 model. The p_i ($i \in \{1,2,3\}$) are calculated by means of a constant least-squares fit to the ratio $\Delta T_i^{SAT}(t)/\Delta T^{GL}(t)$, where the ΔT_i^{SAT} are changes in surface air temperature over the respective Atlantic box. It must be noted that here we implicitly assume that the ΔT_i^* correspond to changes in surface air temperatures ΔT_i^{SAT}, although the T_i^* have a somewhat different physical interpretation. In fact, the T_i^* are the temperatures the oceanic boxes would assume in the absence of overturning m and the associated heat transport, as discussed in Section 2. Since we deal with anomalies in the two quantities, however, the approximation $\Delta T_i^* \approx \Delta T_i^{SAT}$ is justified. The p_j ($j \in \{NH, SH\}$) and the h_k ($k \in \{1,2\}$) are obtained as above from the ratios $\Delta T^j(t)/\Delta T^{GL}(t)$ and $\Delta F_k(t)/\Delta T^j(t)$, respectively. The factor h_2, which considers the flux into the Atlantic north of 50°N, is in the following referred to as "North Atlantic hydrological sensitivity" and is diagnosed as 0.013 Sv °C^{-1} in CLIMBER-2. h_1 and the temperature constants are given in Table 4. Note that under greenhouse warming, changes in the poleward atmospheric water vapour transport are positive in both hemispheres in CLIMBER-2. h_1 is negative since the perturbation ΔF_1 is as F_1 directed towards the Equator. Note also that changes in the freshwater transport by the wind-driven ocean circulation are assumed to be negligible.

The above formulation of temperature and freshwater forcing allows for a large flexibility in the consideration of uncertainties about the amount and regional distribution of the warming and changes in the hydrological cycle, which are among the main uncertain factors in predicting the response of the THC to climate change. Indeed, quantities which are "diagnostic" in conventional climate models are here represented as parameters and can easily be varied. Not included is the feedback of THC changes on freshwater fluxes and GMT (recall that the feedback on thermal forcing is taken into account through the value of the thermal coupling constant Γ; cf. Sect. 2). The feedback on GMT, however, appears to be small (cf. Fig. 2b in Rahmstorf and Ganopolski 1999a), while the feedback of the circulation on meridional freshwater transport is in CLIMBER-2 and other models found to become relevant only after a complete shutdown of the THC (Hughes and Weaver 1996).

Table 4 Climate-change scenario parameters derived from simulations with the CLIMBER-2 model

Climate-change scenario parameters	
Regional temperature constants	
p_1	0.86
p_2	1.07
p_3	0.79
p_{SH}	0.93
p_{NH}	1.07
Hydrological sensitivities	
h_1	−0.005 Sv °C^{-1}
h_2	0.013 Sv °C^{-1}

5 Transient experiments

In this section we characterize the response of the THC as simulated by the box model and compare it to the behaviour of CLIMBER-2 and other comprehensive models. As forcing scenario we assume that GMT rises linearly by 4.5 °C over 150 years and stays constant thereafter. This scenario approximately mimics the response of the CLIMBER-2 model to a $4 \times CO_2$ stabilization scenario, with CO_2 concentrations increasing at 1% per year. As initial conditions for the transient runs we use the steady-state values given in Table 3. In Sect. 5.1 we discuss the box model response for standard parameter settings. In Sect. 5.2 we investigate the box model behaviour for a broad range of forcing parameters and propose extensions of the model for taking into account ulterior sources of uncertainty.

5.1 Response of the THC for standard parameters

In the first experiment, the model parameters are set to their standard values (cf. Tables 1 and 4). The time series of the forcing and the main box model variables are shown in Fig. 3. As the temperatures T_i^* and the poleward atmospheric freshwater transports increase along with global mean temperature, the overturning slows down to 16.5 Sv. It recovers as soon as the forcing is stabilized. In qualitative terms, this corresponds to the behaviour of CLIMBER-2 (cf. curve labelled 0.013 in Fig. 4) as well as of several other coupled ocean–atmosphere models (cf. Fig. 9.21 in Houghton et al. 2001). Quantitatively the overturning weakens less than in CLIMBER-2 and also recovers more rapidly. This is a consequence of our choice for the volume of the deep box V_4 which is taken to be rather small in order to reproduce the sensitivity of CLIMBER-2 to freshwater forcing ΔF_2 (see next section). Indeed, the smaller V_4, the shorter the time scale at which heat and freshwater perturbations are distributed among the boxes and the initial north–south density gradient is restored. This reduces the sensitivity of the system to perturbations and leads to a faster recovery of the circulation. One could have chosen the free parameter V_4 such that the amount of weakening and the recovery time scale of CLIMBER-2 were reproduced. We preferred the first option, as in view of specific integrated assessment applications (Keller et al. 2000; Mastrandrea and Schneider 2001; Zickfeld and Bruckner 2003) our main interest lies in correctly reproducing the THC threshold. Also, the amount of weakening and the recovery time scale of the overturning as simulated by CLIMBER-2 are

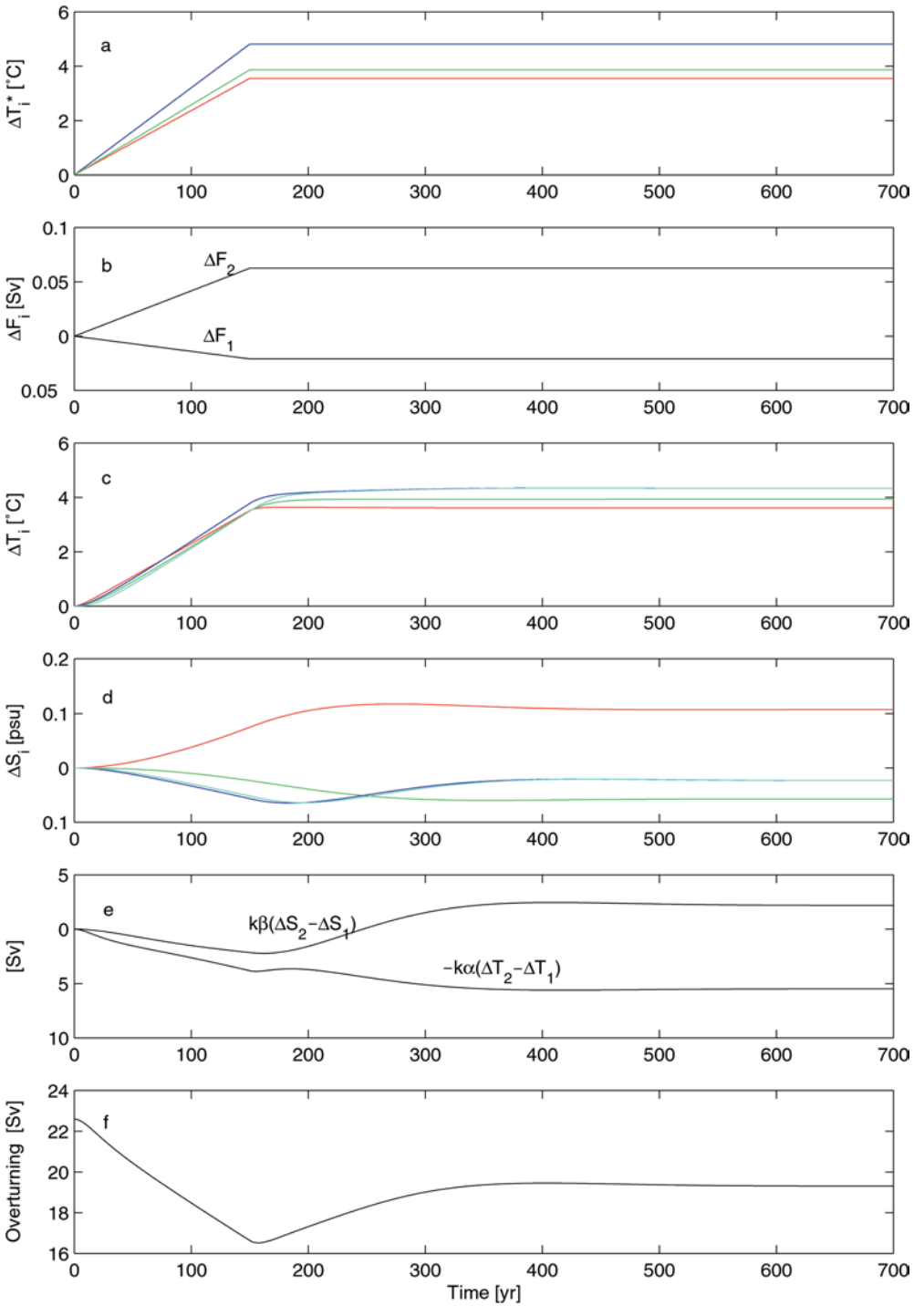

Fig. 3a–f Time series of key model variables for standard parameter values. **a** Regional temperature forcings. **b** Freshwater forcing. **c** Temperature anomaly of each box. **d** Salinity anomaly of each box. **e** Contributions of temperature and salinity gradient anomalies to the overturning (cf. Eq. 1). **f** Overturning. The *deep blue curves* denote changes related to the northern box, the *green curves* changes related to the southern box, the *red curves* changes related the tropical box and the *light blue curves* changes related to the deep ocean

considerably lower than in other models (cf. Fig. 9.24 in Houghton et al. 2001).

In the box model the initial weakening of the overturning is mainly due to thermal forcing, as can be inferred from Fig. 3e. Indeed, the northern box warms faster than the southern box (cf. Fig. 3c), leading to a reduction of the north–south temperature gradient which drives the flow. This effect is amplified by a negative feedback on salinity, since a weaker circulation implies reduced salt advection towards the northern latitudes. The recovery of the THC after stabilization of the forcing is primarily caused by the northward advection of the positive salinity anomaly that has formed in the tropics (cf. Fig. 3d and e). The overturning does not fully regain its original strength because of the asymmetric warming of the southern and northern boxes, which leads to a permanent weakening of the driving temperature gradient. A more detailed discussion of the role of the regional distribution of the warming in determining the response of the THC is given in the next section. The stabilization

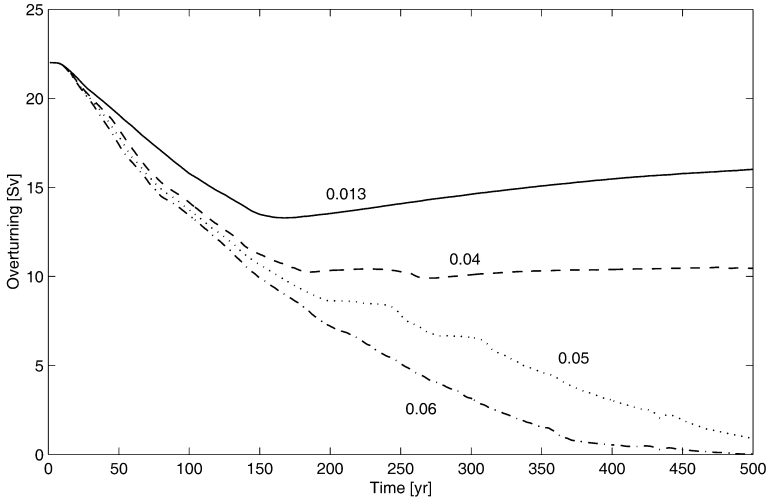

Fig. 4 Response of the Atlantic overturning as simulated by CLIMBER-2 for standard and enhanced values of the hydrological sensitivity h_2 (Sv °C^{-1}). The underlying forcing scenario is a $4 \times CO_2$ stabilization scenario with CO_2 concentrations increasing at 1% a^{-1}

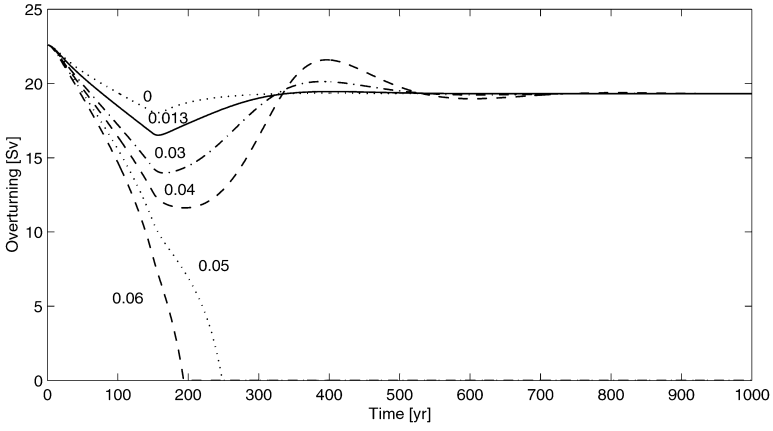

Fig. 5 Response of the Atlantic overturning as simulated by the box model for different values of the hydrological sensitivity h_2 (Sv °C^{-1})

of the THC at a level which lies below the present-day value was also found by Schmittner and Stocker (1999) with a model of intermediate complexity. Note that in this case GCM experiments can hardly be used as reference since the runs are too short to allow the models to reach a new equilibrium.

In order to further test the hypothesis that the initial weakening of the circulation is caused by thermal forcing, we isolate the latter from the freshwater forcing by setting h_2 to 0. We found that temperature effects alone, together with the negative feedback of the circulation on salinity, lead to a considerable weakening of the overturning (cf. curves labelled 0 and 0.013 in Fig. 5). This behaviour is similar to that exhibited by CLIMBER-2 as well as by many other models where thermal forcing was found to be responsible for the weakening of the THC (Rahmstorf and Ganopolski 1999a; Schmittner and Stocker 1999; Mikolajewicz and Voss 2000; Thorpe et al. 2001). There are models, however, where the freshwater fluxes play the key role in slowing down the THC (Dixon et al. 1999).

5.2 Sensitivity experiments

In a first set of experiments we investigated the response of the THC to the same scenario as above but for different values of the North Atlantic hydrological sensitivity h_2. This parameter is among the main uncertain quantities in predicting the fate of the THC. The reason is that estimates of evaporation, precipitation, river runoff and meltwater changes in the North Atlantic catchment differ largely. In the CLIMBER-2 model, for example, the value for h_2 is 0.013 Sv °C^{-1}, which is at the lower end of the range spanned by climate models (Rahmstorf and Ganopolski 1999a). The GFDL model shows a much stronger response of the hydrological cycle than CLIMBER-2: a 7.5 °C hemispheric warming leads to an additional freshwater flux to the North Atlantic of approximately 0.4 Sv, which results in a hydrological sensitivity of 0.053 Sv °C^{-1} (Manabe and Stouffer 1994). Recent observational studies were able to constrain at least part of the projected North Atlantic freshwater budget: Peterson et al. (2002) found that the Eurasian river discharge to the Arctic Ocean had increased by 7% since 1936. Regressing discharge against GMT, Peterson et al. (2002) provide an empirical estimate of the contribution of the Eurasian river discharge to the Atlantic hydrological sensitivity of 0.007 Sv °C^{-1}. Van der Veen (2002) estimated the ablation of the Greenland ice sheet at 10% per degree local warming, which also results in a sensitivity of 0.007 Sv °C^{-1}. Considering that Eurasian river discharge accounts for

only approximately 40% of the total discharge to the Arctic Ocean and that precipitation minus evaporation over the Atlantic is a major component of the freshwater budget, these empirical estimates suggest that a value of 0.05 Sv $^{\circ}C^{-1}$ for the North Atlantic hydrological sensitivity is not unrealistic.

We consider scenarios with enhanced hydrological response by increasing the hydrological sensitivity h_2 to 0.03, 0.04, 0.05 and 0.06 Sv $^{\circ}C^{-1}$. For the above temperature scenario this leads to an additional freshwater transport ΔF_2 into the northern box of 0.14, 0.19, 0.24 and 0.29 Sv, respectively. The response of the overturning to these scenarios is shown in Fig. 5. The responses diverge, falling into two categories: those in which the overturning recovers and those in which it collapses. This is in line with CLIMBER-2 experiments where a critical value for the hydrological sensitivity h_2 was determined (see Fig. 4 and Rahmstorf and Ganopolski 1999a). The critical value h_2^{crit} in the box model is 0.046 Sv $^{\circ}C^{-1}$, yielding a critical freshwater forcing ΔF_2 of 0.22 Sv. This value corresponds to that of CLIMBER-2 since the free parameter V_4 was tuned to reproduce this quantity (cf. Sect. 3).

A notable feature in Fig. 5 is the overshooting and subsequent oscillation of the overturning in the recovery phase for values of h_2 close to h_2^{crit}. This is due to the fact that in the latter case the overturning is slowed down considerably such that a strong positive salinity anomaly develops in the tropical box which is advected northward as soon as the circulation regains strength. This might suggest that the overshooting is a model artefact due to the neglect of mixing processes. Indeed, the latter would allow for diffusive transport of salt out of the tropics even in the absence of a vigorous overturning. The same mechanism (i.e. the massive advection of salt as the circulation regains strength) also causes the spike in the hysteresis curve traced with the box model (cf. Fig. 2).

Note that for a hydrological sensitivity h_2 of 0.05 Sv $^{\circ}C^{-1}$, which approximately corresponds to the value diagnosed in the GFDL model, the overturning collapses. This implies that for a suitable choice for h_2 the box model results are similar to those of the GFDL model, where a collapse of the THC was found for a quadrupling of atmospheric CO_2 (Manabe and Stouffer 1993, 1999).

Another source of uncertainty concerns the freshwater forcing ΔF_1. In a set of experiments we explored the sensitivity of the model response to the perturbation ΔF_1 by scanning the hydrological sensitivity h_1 in the range 0 to -0.02 Sv $^{\circ}C^{-1}$ while keeping all other parameters at their standard values. In all cases we found that qualitatively the model response changes little compared to that shown in Fig. 3: the overturning circulation is weakened and subsequently recovers. It reaches a minimum of 15.7 Sv for $h_1 = 0$ and 18.9 Sv for $h_1 = -0.02$ (compared to 16.5 Sv for the standard case $h_1 = -0.005$). We then increased h_2 as in the experiment above and computed the critical value for h_2: h_2^{crit} is lowered to 0.040 Sv $^{\circ}C^{-1}$ for $h_1 = 0$ and raised to 0.063 Sv $^{\circ}C^{-1}$ for $h_1 = -0.02$ (compared to 0.046 Sv $^{\circ}C^{-1}$ for $h_1 = -0.005$). This sensitivity can be explained as follows: the more negative h_1, the stronger the net freshwater export from the tropical Atlantic. The latter tends to counterbalance the freshening of the high latitudes through enhanced salt advection to the north. A positive h_1, on the other hand, would lead to a fresher tropical Atlantic and reduced salt advection, acting in the same direction as the freshening of the North Atlantic to destabilize the circulation.

So far we subsumed changes in the atmospheric freshwater transport F_2 and increased meltwater runoff from ice sheets and glaciers under the same forcing parameter, the hydrological sensitivity h_2. This is not fully correct since in the atmospheric vapour transport case, the freshwater gain in the North Atlantic is compensated for by freshwater loss in the tropics, whereas in the meltwater case it is not. For subsequent sensitivity experiments we modified the box model such that meltwater runoff from the Greenland ice sheet and meltwater from Arctic sea ice could be taken into account. This was achieved by introducing an additional meltwater flux F_2^{MW} into the northern box which is not compensated within the box model. We assume that changes in this flux can also be approximately described by a linear equation:

$$\Delta F_2^{MW} = h_2^{MW} \Delta T^{NH} \quad . \tag{28}$$

The total freshwater forcing $\Delta F_2^{\mathrm{tot}}$ of the northern box then is:

$$\begin{aligned} \Delta F_2^{\mathrm{tot}} &= \Delta F_2 + \Delta F_2^{MW} \\ &= (h_2 + h_2^{MW}) \Delta T^{NH} \equiv h_2^{\mathrm{tot}} \Delta T^{NH} \quad . \end{aligned} \tag{29}$$

We assume that 1/3 of the total additional freshwater $\Delta F_2^{\mathrm{tot}}$ entering the northern box is provided by meltwater (i.e. $h_2^{MW} = 1/3 h_2^{\mathrm{tot}}$). This implies that we expect the contributions from meltwater runoff and atmospheric transport to be of the same sign (i.e. positive) and order of magnitude. For Greenland, the latest report of the IPCC indicates a sensitivity of 0.001 to 0.005 Sv for a 1 °C local warming (cf. Table 11.7 in Houghton et al. 2001). Given that in CLIMBER-2 and in other models the temperature change over Greenland is approximately twice as large as that of the Northern Hemisphere as a whole (i.e. $p^{GRL}/p^{NH} \approx 2$, where p^{GRL} is the temperature constant for Greenland), this range corresponds to $h_2^{MW} = 0.002$–0.01 Sv $^{\circ}C^{-1}$ in our notation. Note that this range encompasses the empirical value of 0.007 Sv $^{\circ}C^{-1}$ discussed previously. The contribution of sea ice to h_2^{MW} can be estimated as follows: in the past two decades or so, about 14% of perennial Arctic sea ice have melted away (Johannessen et al. 1999). Since perennial sea ice is about 3–4 m thick and occupies an area of 6–7×10^6 km^2, the contribution to the freshwater budget can be estimated to about 0.01 Sv. During the same time, annual mean air temperature in the Northern Hemisphere has increased by about 0.8 °C (Houghton et al.

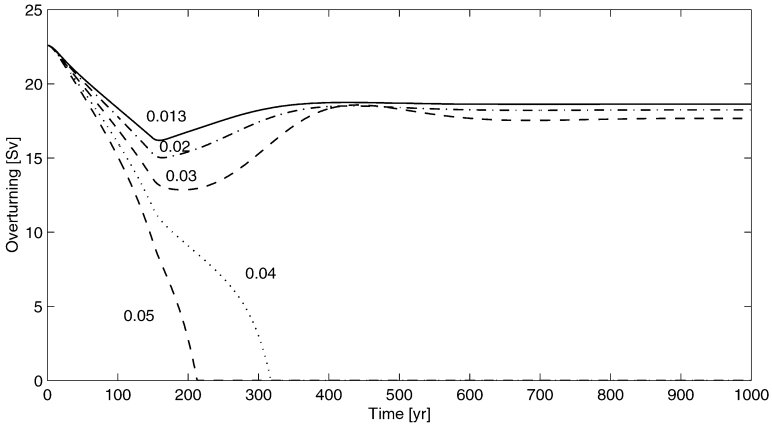

Fig. 6 Response of the Atlantic overturning for different values of the hydrological sensitivity h_2^{tot} (Sv $^{\circ}\mathrm{C}^{-1}$). 1/3 of the additional freshwater entering the northern box is provided by meltwater from ice sheets and glaciers in the North Atlantic catchment

2001). This would yield an estimate of 0.0125 Sv $^{\circ}\mathrm{C}^{-1}$ for the contribution of sea ice to the hydrological sensitivity h_2^{MW}. Note, however, that this sensitivity declines as the Arctic sea-ice cover fades away.

In a set of experiments we increase h_2^{tot} up to 0.05 Sv $^{\circ}\mathrm{C}^{-1}$ and compute the corresponding circulation response. In this case the overturning is more sensitive than under increased atmospheric freshwater transport ΔF_2 alone (cf. Fig. 6): the critical value for h_2^{tot} is 0.039 Sv $^{\circ}\mathrm{C}^{-1}$, corresponding to a total freshwater forcing of 0.19 Sv (compared to 0.22 Sv above). This greater sensitivity can be explained as follows: in the case of enhanced meridional freshwater transport F_2 alone, the tropical box becomes more saline. This salt anomaly is subsequently advected poleward, increasing the salinity of the northern box and therefore stabilizing the circulation. When part of the additional freshwater is supplied by meltwater, the freshening of the northern box is only partly compensated by salinity advection from the tropics. This also explains why in comparison to the experiments with enhanced meridional freshwater transport alone the overturning takes longer to reintensify and no overshooting is observed. The above effect was already noted by Rahmstorf and Ganopolski (1999a): in global warming experiments they found the critical freshwater flux to differ by approximately 25% between the meltwater and the meridional transport cases.

In a recent paper Latif et al. (2000) indicated another possible source of uncertainty in simulating the response of the THC. They showed that greenhouse-gas warming could lead to an additional freshwater export from the tropical Atlantic to the tropical Pacific with a stabilizing effect on the THC. The physical mechanism responsible for this freshwater transport was found to be an El-Niño-like warming pattern in the eastern equatorial Pacific associated with changes in the atmospheric circulation in the tropics. In order to allow for the consideration of interbasin freshwater transport in the box model, we introduce a (negative) freshwater flux F_4 into the tropical box which is compensated for in the Pacific (i.e. outside the box model). We again approximate changes in this freshwater flux by a linear relationship:

$$\Delta F_4 = h_4 \Delta T_3^* \quad . \tag{30}$$

Drawing upon Fig. 4 in Latif et al. (2000), we set h_4 to -0.1 Sv $^{\circ}\mathrm{C}^{-1}$ (i.e. a change in interbasin freshwater transport of -0.2 Sv for $\sim 2\,^{\circ}\mathrm{C}$ warming in the tropics), while all other parameters are held at their standard values. Figure 7 displays the response of the key box model variables under consideration of the freshwater export from the Atlantic to the Pacific. The overturning m is weakened slightly as long as the warming is sustained and then recovers, even exceeding the initial strength. This is due to the enhanced salinity in the tropics, which is advected northward and leads to a permanent increase in the density gradient between the northern and the southern box. Because of this stabilizing mechanism the freshwater perturbation ΔF_2 needed to shut down the circulation is almost twice as large as in experiments without interbasin freshwater transport F_4 (0.38 Sv compared to 0.22 Sv). Note that the sustained increase in the salinities of the four boxes (cf. Fig. 7d) is due to an idealization of the box model (i.e. the isolation of the Atlantic from the other oceanic basins), which prevents the freshwater exported to the Pacific from circulating back into the Atlantic.

So far we investigated mainly the role of the uncertain freshwater fluxes in determining the response of the THC. Another factor contributing to the response is the unequal thermal forcing of the three boxes, as expressed by the parameters p_i (cf. Table 4). To study its effect, we isolate the thermal from the freshwater forcing by assuming the p_j, $j \in \{NH, SH\}$, to be independent of the p_i, $i \in \{1, 2, 3\}$, and perform a sensitivity analysis with respect to the p_i. In the first experiment we assume that the northern and southern temperature factors p_1 and p_2 are the same. We find that the overturning is reduced much less and recovers more rapidly as compared to the standard case, and even exceeds its initial strength after the forcing is stabilized. This is because of the symmetric warming, which does not act to offset the north–south temperature gradient as in the case $p_1 < p_2$.

We then assume equality between the tropical and the northern temperature constants p_3 and p_2. The overturning is found to be more sensitive to the warming scenario. Indeed, as the thermal forcing of the tropical box is increased, the northward advection of warmer

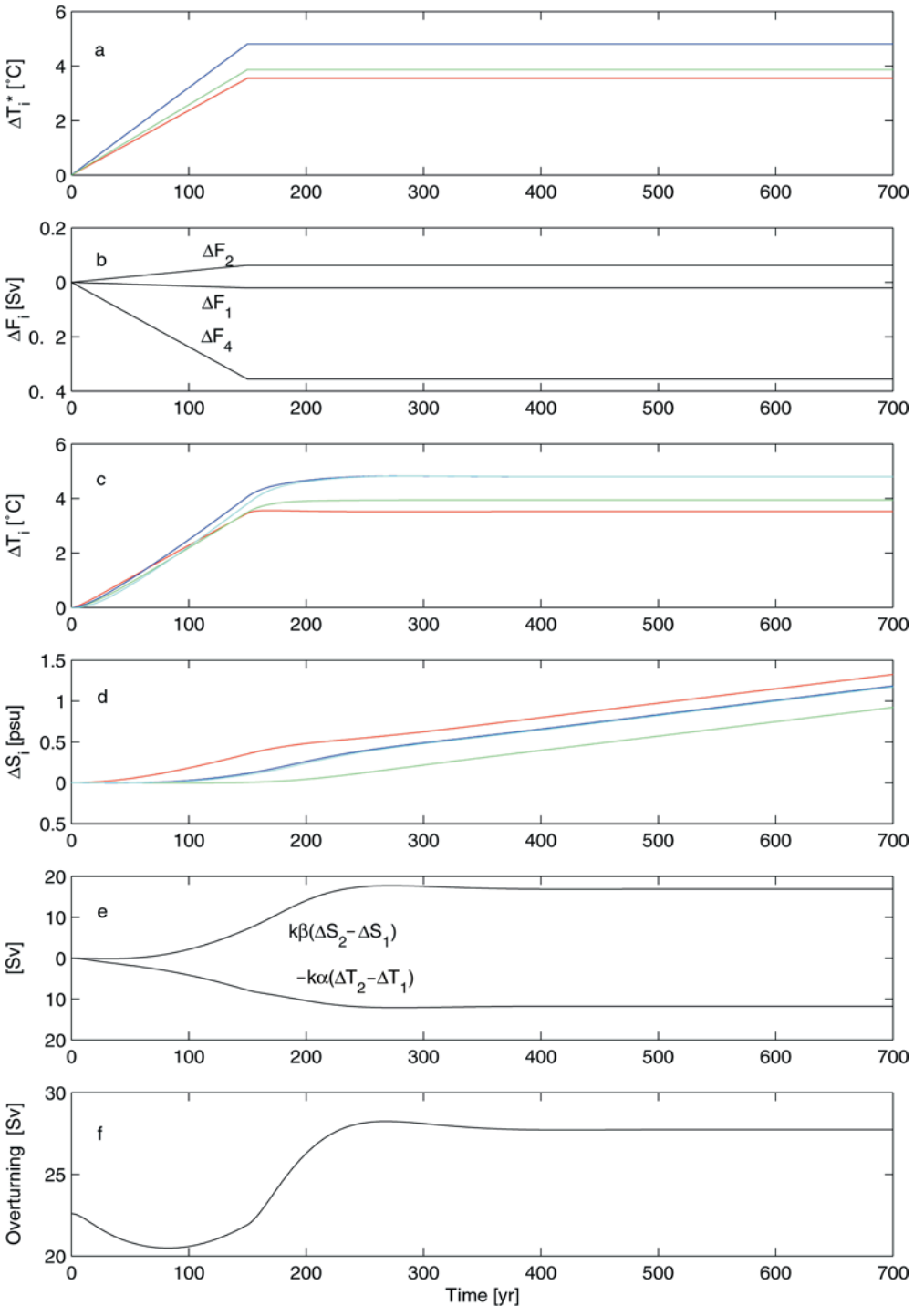

Fig. 7a–f Time series of key model variables under enhanced freshwater export ΔF_4 from the tropical Atlantic. h_4 is -0.1 Sv °C^{-1}, all other parameters are at their standard values. **a** Prescribed regional temperature forcings. **b** Prescribed freshwater forcings. **c** Temperature anomaly of each box. **d** Salinity anomaly of each box. **e** Contributions of temperature and salinity gradient anomalies to the overturning (cf. Eq. 1). **f** Overturning. The *deep blue curves* denote changes in the northern box, the *green curves* changes in the southern box, the *red curves* changes in the tropical box and the *light blue curves* changes in the deep ocean

waters acts to further reduce the north–south density gradient. In a last experiment we set $p_3 = p_1 = p_2$: the circulation weakens less and recovers faster than for the standard p_i and exceeds its initial strength. This indicates that in this case the stabilizing effect associated with the symmetric warming of both hemispheres outweighs the destabilizing effect due to advection of warmer tropical waters.

Furthermore, we found that the THC as simulated in the box model is sensitive to the rate of climate change: the greater the rate of temperature increase, the stronger the weakening of the circulation (cf. Fig. 8). In a set of experiments with h_2 at the critical value (0.046 Sv °C^{-1}) and the other sensitivities at their standard values ($h_1 = -0.005$ Sv °C^{-1}, $h_2^{MW} = h_4 = 0$ Sv °C^{-1}), we found that for slower rates of temperature change the circulation weakens and then recovers, while for faster rates of change it collapses irreversibly. The critical rate (i.e., 3 °C century^{-1}) is determined by the time scale at which excess heat and freshwater are redistributed among the

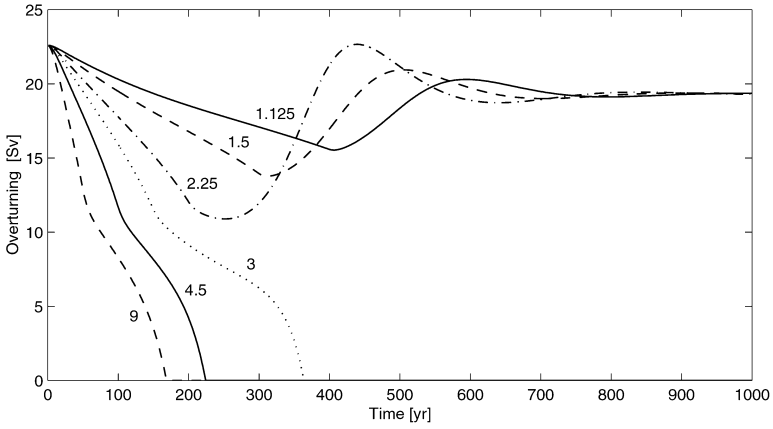

Fig. 8 Response of the Atlantic overturning for different rates of temperature increase (°C century^{-1}). The hydrological sensitivities are the same in all experiments ($h_2 = 0.046$ Sv °C^{-1}, $h_1 = -0.005$ Sv °C^{-1}, $h_2^{MW} = h_4 = 0$ Sv °C^{-1})

boxes. This mechanism is similar to the one identified by Stocker and Schmittner (1997), who first demonstrated that the response of the THC is dependent upon the rate of climate change.

In order to allow for a quantitative comparison between our results and those obtained by Stocker and Schmittner (1997) (henceforth referred to as S&S), we calculated the stability diagram of the THC in the temperature, rate of temperature increase (i.e. $(T, \dot{T})$) phase space assuming high, medium and low values of the hydrological sensitivity h_2 (i.e. 0.01, 0.03, 0.05 Sv °C^{-1}, respectively). The result is displayed in Fig. 9. Each marked point on the stability curves is obtained by prescribing the rate of temperature change and by increasing the temperature up to a value for which the circulation collapses. Figure 9 also displays the stability curve calculated by S&S with a zonally averaged ocean model. Note that originally this curve was computed in the CO_2 concentrations and rate of concentrations increase phase-space. For the purpose of comparability, we use the S&S curve translated to $(T, \dot{T})$ phase-space (cf. Tóth et al. 1998). Figure 9 shows that the box model and the S&S curves are similar in shape, with a stronger sensitivity of the THC threshold on temperature increase for smaller rates of change. Our results also indicate that for the emulation of the stability properties of the model used by S&S, the hydrological sensitivity h_2 would need to be set to 0.04–0.05 Sv °C^{-1}.

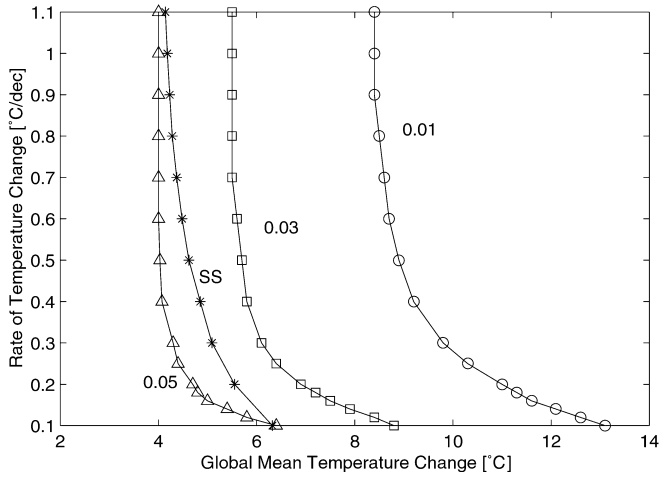

Fig. 9 Stability diagram of the THC in (T, $\dot{T}$) phase space for different assumptions about the hydrological sensitivity h_2 (Sv °C^{-1}). The stable (unstable) domains are located to the *left* (*right*) of the respective curves. For comparison we show the stability curve derived from Stocker and Schmittner (1997) (*SS*)

Lastly, we investigated the sensitivity of the circulation with respect to the initial overturning intensity m_{eq}. The initial overturning can be varied in two ways: by varying the model parameters, i.e. the shape of the equilibrium solution curve $m(F_1)$ (cf. Fig. 2), or by varying the freshwater flux F_1 (i.e. the position of m_{eq} on the solution curve). We here consider the model parameters to be fixed, and restrict ourselves to the second case. Figure 2 suggests that the position of m_{eq} on the curve (i.e. the strength of the m_{eq}) determines the distance of the system to the bifurcation point and hence the freshwater perturbation which is needed to shut down the circulation (see also Rahmstorf 2000). To test the validity of this hypothesis for the transient case, we adjusted the freshwater flux F_1 so as to create weaker ($m_{eq} = 10$–20 Sv) and stronger ($m_{eq} = 25$–30 Sv) overturning rates compared to our standard setting ($m_{eq} = 22.6$ Sv). Note that these values for m_{eq} match the range of 10–30 Sv for the initial (i.e. control) overturning spanned by state-of-the-art climate models (Houghton et al. 2001). Observation-based studies estimate the present-day overturning at approximately 15 Sv (Ganachaud and Wunsch 2000; Stammer et al. 2002). For each case we then performed a sensitivity analysis with respect to the parameter h_2. The results are shown in Fig. 10. As conjectured, the stronger the initial overturning, the higher the critical value for the hydrological sensitivity. For $m_{eq} = 20$ Sv h_2^{crit} is 0.027 Sv °C^{-1}, for $m_{eq} = 25$ Sv h_2^{crit} is 0.067 Sv °C^{-1} and for $m_{eq} = 30$ Sv h_2^{crit} is 0.122 Sv °C^{-1}, corresponding to critical freshwater fluxes of 0.13 Sv, 0.32 Sv and 0.59 Sv, respectively. In the cases $m_{eq} = 10$–15 Sv (not shown), the circulation collapses even without any additional freshwater input to the northern box, meaning that

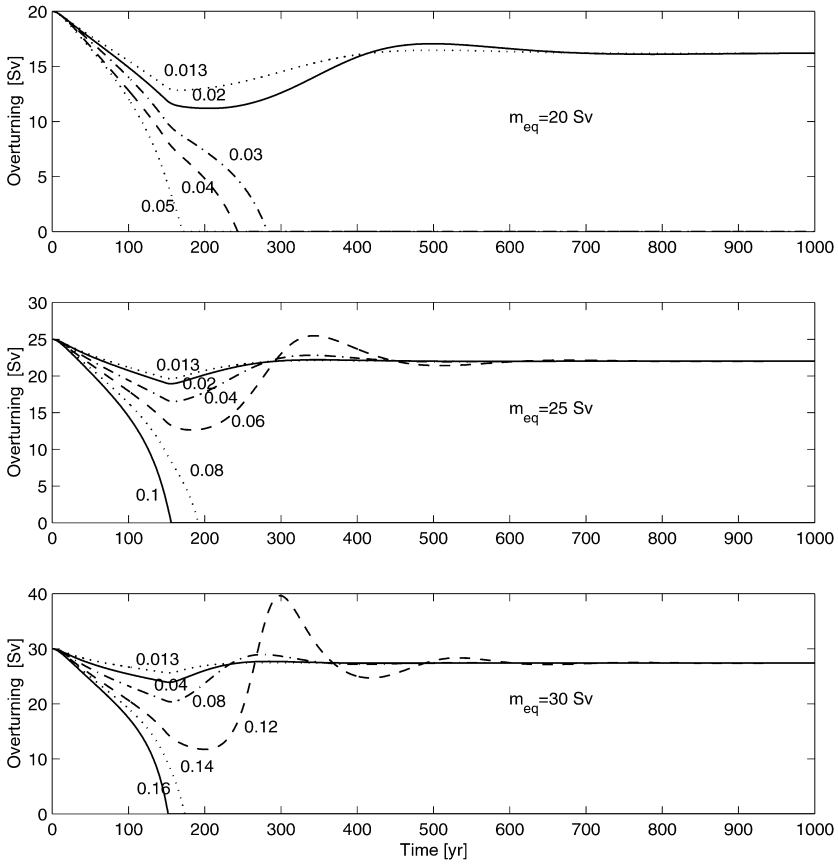

Fig. 10 Response of the Atlantic overturning for different values of the initial overturning m_{eq} (Sv) and hydrological sensitivity h_2 (Sv °C^{-1}). In the cases $m_{eq} = 10$–15 Sv (not shown) the circulation collapses already for $h_2 = 0$

thermal forcing alone is sufficient to reverse the driving north–south density gradient.

So far one single attempt has been made to perform similar sensitivity experiments with a coupled ocean–atmosphere GCM (Tziperman 1997). The results, however, cannot be taken as benchmark for the behaviour of the box model, since the initial states prepared by Tziperman (1997) are physically not fully consistent (weaker THC states than today's with present-day observed sea-surface temperature fields; see note by (Rahmstorf and Ganopolski 1999b).

6 The effects of vertical and horizontal mixing

A number of studies have used three or four ocean boxes in one hemisphere, with a vertically stacked high-latitude box (Huang et al. 1992; Tziperman et al. 1994; Griffies and Tziperman 1995; Gargett and Ferron 1996). In two hemispheres, such a model setup would imply six or eight boxes, depending on whether the low-latitude box is split into a northern and a southern part or not (cf. e.g. Gildor and Tziperman 2001). The advantage of such model configurations is that they provide the opportunity to introduce additional physical processes, such as vertical diffusion and convection in the southern and northern boxes.

Concerning the southern box in our model, the subdivision into an upper and a lower part and the implementation of vertical mixing would not deliver new physical insight. The reason is that in the conception of the THC underlying the four-box model, box 1 represents the entire ocean south of the latitudes of Drake Passage, rather than the polar Southern Ocean. Given that the heat forcing at the surface is chosen accordingly, and that the deep southern box would be filled with NADW, the vertical stratification of the water column would not become unstable, consistently with the absence of open-ocean deep convection in the real Southern Ocean. Also, vertical diffusion of the order of 10^{-4} m^2 s^{-1} would be small compared to the advective flux connecting lower and upper box (and mimicking upwelling). The split-up of the southern box would therefore merely amount to a redefinition of the box volumes, with NADW extending further south and the waters of southern origin being confined to the surface layer.

Also for the northern box, we do not expect the effects of the explicit representation of stratification and vertical mixing to be large. To test this conjecture, we subdivided the northern box in an upper and a lower part, and connected them by an advective (downward) and a diffusive (bidirectional) flux (cf. Appendix). In this model setup, convection is represented through the choice of the vertical diffusion coefficient, which is taken to be very high for conditions of static instability of the water column and low for static stability, similarly to Gildor and Tziperman (2001) and Shaffer and Olsen

(2001). This formulation is equivalent to the complete mixing of the water properties (i.e. heat and salt) of the two overlying boxes in the case of unstable stratification (Huang et al. 1992; Gargett and Ferron 1996). It should be noted, however, that the representation of convection in a coarse box model without implementation of the seasonal cycle is problematic. The reason is that in reality the convective events are spatially and temporally highly localized: they occur exclusively in winter, lasting but a few days, on spatial scales of 50–100 km.

In the Appendix, we give the equations for the model version representing vertical and horizontal mixing explicitly. In the case of vertical mixing in the northern box alone (i.e. for K_v^3 and K_h equal to zero), it is easily inferred from Eqs. (35) and (40) that in equilibrium T_2 and S_2 equal T_5 and S_5, such that the vertical diffusion coefficient for the northern box, K_v^2, does not influence the strength of the overturning. Also the transient THC behaviour remains largely unaffected by northern vertical mixing: the amount and time scale of the weakening and recovery of the overturning in response to the climate change scenario introduced in Section 5 are not at all sensitive to the vertical mixing coefficient. The same applies to the value of the freshwater perturbation $\Delta F_2^{\mathrm{crit}}$ that the overturning can sustain without collapsing. The reason for this insensitivity is that the stratification of the water column remains stable throughout the transient run, such that convection does not occur, and turbulent mixing of the order of 10^{-4} m^2 s^{-1} is small compared to the advective fluxes. The delay (by some decades) and reduction (by about 2 Sv) of the weakening of the overturning as compared to our standard model setup (cf. the dotted and red lines in Fig. 11) is solely due to the split-up of the northern box in two parts.

A process neglected so far in our model is diapycnal mixing across the main thermocline. In order to estimate the error made by omitting this process, we introduced a vertical flux between boxes 3 and 4 (while retaining the four-box formulation). We found that the low-latitude diffusion coefficient (K_v^3; cf. Appendix) affects the equilibrium overturning only slightly: increasing K_v^3 from 0 to 10^{-4} m^2 s^{-1} leads to a strengthening of the overturning by 0.5 Sv. The reason for this slight increase is downward diffusion of heat and subsequent southward advection, which strengthens the meridional density gradient. Also, the response of the overturning to our standard climate change scenario changes only little if vertical diffusion is introduced: the amount and time scale of the weakening are almost the same, just the stabilization level decreases by about 1 Sv for values of the vertical diffusion coefficient of 10^{-4} m^2 s^{-1} (cf. the red and solid lines in Fig. 11). As far as the sensitivity of the overturning to freshwater perturbation is concerned, it turns out that the critical freshwater forcing $\Delta F_2^{\mathrm{crit}}$ decreases for non-zero values of the diffusion coefficient: for K_v^3 equal to 10^{-4} m^2 s^{-1} $\Delta F_2^{\mathrm{crit}}$ is reduced by 30% compared to the case $K_v^3 = 0$. Note, however, that a diffusivity of 10^{-4} m^2 s^{-1} is at the high end of estimates of diffusion in the real ocean: measurements yield a value of about 10^{-5} m^2 s^{-1} at the thermocline (Ledwell et al. 1998). The reason for this stronger sensitivity is that in the presence of vertical diffusion the freshwater anomaly which is advected from the northern box towards the south mixes with high-salinity water from the tropics. As a consequence, the redistribution of the freshwater anomaly among the boxes, which acts to stabilize the overturning (cf. Sect. 5.1), takes longer for higher values of K_v^3. Note that this result is in contrast to the findings of studies using two- and three-dimensional ocean models (Schmittner and Weaver 2001; Prange et al. 2003), where the critical freshwater perturbation is found to increase significantly with the vertical diffusivity. The reason for this discrepancy is not clear.

Lastly, we augmented the four-box model with horizontal diffusion, which is meant to represent the effects of the wind-driven gyre circulation. This is achieved by introducing horizontal fluxes between adjacent boxes (cf. the appendix). As for vertical diffusion, the effects on the overturning are small. The equilibrium state decreases only slightly with increasing horizontal diffusivity (by about 0.4 Sv for K_h equal to 10^{-1} m^2 s^{-1}), and the characteristics of the response curve remain essen-

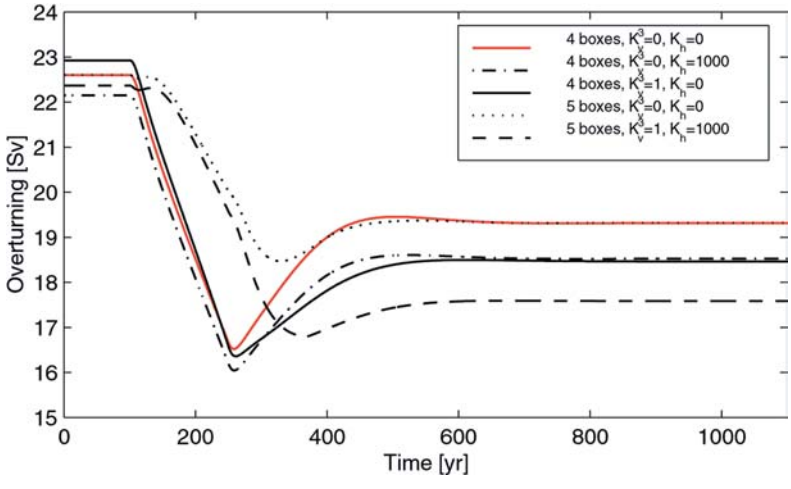

Fig. 11 Response of the Atlantic overturning to the climate change scenario introduced in Section 5 for different box configurations and different values of the vertical and horizontal mixing coefficients (10^{-4} m^2 s^{-1}). The *red curve* depicts the response for standard model settings. Note that the climate change scenario is applied from year 100 on

tially the same (cf. the dashed-dotted and red lines in Fig. 11). The freshwater perturbation needed to shut down the circulation is unaffected by horizontal diffusion. A notable result is that the oscillations evident during the recovery phase of the overturning for values of h_2 close to critical (cf. e.g. Figs. 5 and 10) are more strongly damped for larger values of the horizontal diffusivity, as was conjectured in Section 5.2.

What do the results discussed in this section imply in terms of the optimal fit to the CLIMBER-2 experiments? Given that the unperturbed equilibrium state is not affected at all by vertical mixing in the north and only slightly by vertical and horizontal diffusion across the low-latitude boxes, inclusion of these processes would change the parameters determined via a least-squares fit to the CLIMBER-2 hysteresis only slightly (cf. Fig. 12). The parameter V_4, which was adjusted to match the critical freshwater perturbation ΔF_2^{crit} determined with CLIMBER-2, would change only in the case of vertical diffusion across the thermocline (recall that ΔF_2^{crit} remains largely unaffected by northern vertical mixing and horizontal diffusion): since vertical diffusion was found to increase the sensitivity of the circulation to freshwater perturbations, V_4 would have to be decreased further with increasing vertical diffusivity. These results suggest that the freedom gained by introducing the additional tunable parameters K_v^2, K_v^3 and K_h would not help to improve the fit to the transient CLIMBER-2 behaviour.

7 Conclusions

In this paper we presented a low-order model of the Atlantic THC. The model is an interhemispheric extension of the classic Stommel (1961) model and is calibrated against results obtained with the state-of-the-art climate model of intermediate complexity CLIMBER-2 (Petoukhov et al. 2000; Ganopolski et al. 2001). It is designed to be forced by scenarios of global mean temperature change which are translated into transient fluxes of heat and freshwater through an appropriate down-scaling procedure.

The specific design and computational efficiency of the model allowed us to perform a large number of sensitivity experiments with respect to uncertain model and forcing parameters. We devoted particular attention to the hydrological forcing, which is among the major uncertain factors determining the response of the THC to global warming. Further we analyzed the response of the overturning to modified regional patterns of warming, different rates of climate change and initial model states.

Although highly simplified compared to the comprehensive, coupled ocean–atmosphere circulation models, the box model reproduces key features of their behaviour. In response to moderate climate-change scenarios, for example, the circulation weakens and, as soon as the forcing is stabilized, recovers. This is in line with the behaviour of the majority of comprehensive climate models as summarized in Houghton et al. (2001). For higher forcing scenarios the overturning collapses due to the existence of a threshold value in the freshwater perturbation beyond which the circulation cannot be sustained. The latter is similar to the behaviour discussed by Manabe and Stouffer (1993), Stocker and Schmittner (1997) and Rahmstorf and Ganopolski (1999a).

Furthermore, the box model reproduces convincingly the results obtained by Stocker and Schmittner (1997), who demonstrated that the stability of the THC is dependent upon the rate of climate change. This rate-sensitive response implies that the final amount of temperature increase that can be reached without inducing a collapse of the THC is higher for slower rates of temperature change.

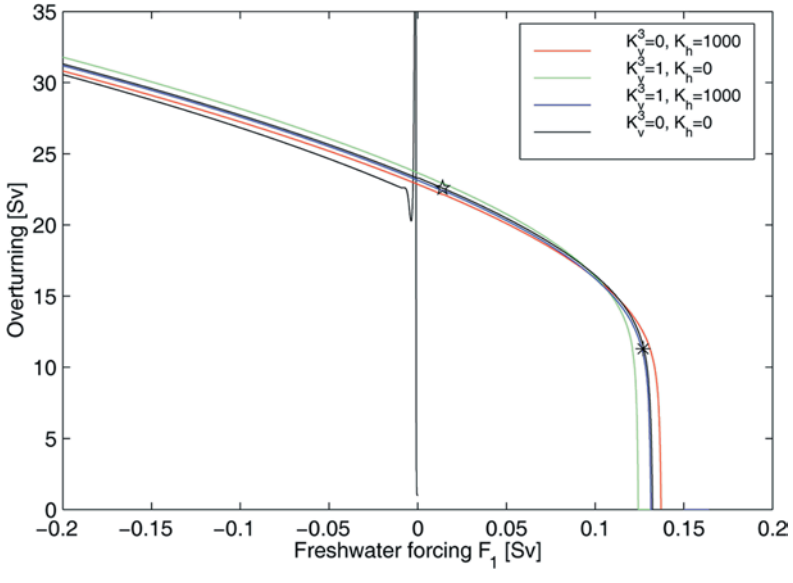

Fig. 12 Upper branch of the hysteresis for different values of the vertical and horizontal diffusion coefficients (10^{-4} m^2 s^{-1}). For reference, we display the full hysteresis curve calculated with our standard model setup. As in Fig. 2, the *star* and the *asterisk* denote the present-day overturning and the bifurcation point, respectively. Note that the hysteresis for $K_v^3 = 10^{-4}$ m^2 s^{-1}, $K_h = 10^{-1}$m^2 s^{-1} almost coincides with the hysteresis for the standard case ($K_v^3 = 0$, $K_h = 0$)

Lastly, we investigated the sensitivity of the THC to the initial model state. Given that the difference between the initial states is determined by the freshwater budget, our results suggest that the weaker the initial overturning, the lower the freshwater forcing needed to shut down the circulation.

Some recent studies have raised awareness that the evolution of the THC may become inherently unpredictable as the system approaches the critical threshold (Knutti and Stocker 2002; Schaeffer et al. 2002). The reason is that in the proximity of the bifurcation point the sensitivity of the THC to stochastic processes (e.g. weather) is increased. Noise in the heat and freshwater forcings could therefore induce a collapse of the THC before the threshold itself is reached. In the case of strong forcing scenarios, this evidence would suggest a probabilistic treatment of the THC, whereby its evolution is characterized by probability density functions (PDFs). If desired, such a probabilistic view can easily be attained with our box model through the parametrization of the fluctuations in the heat and freshwater forcing by a stochastic process (cf. e.g. Griffies and Tziperman 1995 Kleinen et al. 2003) and the performance of ensemble simulations.

Since our model does not explicitly include the effects of vertical and horizontal mixing, such as convection, small-scale turbulent mixing and the wind-driven gyre circulation, we attempted to estimate the error made by neglecting these processes. We found their influence on the unperturbed equilibrium model state to be almost negligible, implying that their inclusion would not compromise the optimality of the model parameters. In the transient case, the effects are not significant either: vertical mixing does not affect the weakening and recovery behaviour of the overturning at all, whereas vertical diffusion across the thermocline and horizontal diffusion lead to a slight decrease of the stabilization level. Most notably, horizontal mixing somewhat dampens the oscillations exhibited by the overturning in response to forcing scenarios which bring it close to the critical threshold.

Our results also suggest that the inclusion of the above physical processes does not necessarily improve the performance of the box model: the implementation of vertical mixing across the thermocline, for example, leads to a stronger sensitivity of the overturning to freshwater perturbation, contrarily to what is observed in two and three-dimensional ocean models (Schmittner and Weaver 2001; Prange et al. 2003). This raises the question about whether a model as simple as ours necessarily gains credibility through the inclusion of additional physical mechanisms. We maintain the view that such a low-order model is a valid predictive tool only if conceived as emulator of comprehensive ocean models. As such, it is desirable to keep the number of its tunable parameters as low as possible. Also, the inclusion of vertical and horizontal mixing would be at the expense of the mathematical simplicity and overall transparency of the model, which we strongly exploited for the design of the parameter-fitting procedure and the characterization of the transient model behaviour.

We conclude that our approach to use a dynamic four-box model of the Atlantic to mimic the response of the THC as simulated by comprehensive climate models was successful. We found that our model represents a useful tool for exploring the basic physical mechanisms underlying the response of the THC to climate change. Further, the model has proven to be fast, transparent and flexible, meeting the main requirements of interdisciplinary applications such as integrated assessment modelling.

Acknowledgements We wish to thank Andrey Ganopolski for performing the experiments with CLIMBER-2 and for giving helpful advice. We are grateful to Till Kuhlbrodt for suggestions for the revision of the manuscript, and to Alexa Griesel, Marisa Montoya and Miguel Morales Maqueda for fruitful discussions. Thanks are also due to two anonymous reviewers for their helpful comments. Financial support for the first author was provided by the German Federal Ministry for Education, Science, Research and Technology (under grant no. 01LD0016).

Appendix

In the following we give the equations for the version of the four-box model including vertical and horizontal mixing. For the representation of convection and vertical diffusion we subdivide the northern box into an upper and a lower part (the lower part being termed box 5) and introduce vertical diffusive fluxes between these boxes, and between boxes 3 and 4. Horizontal mixing is represented through the implementation of diffusive fluxes between adjacent boxes (i.e. between the surface boxes 1 and 3, and 3 and 2, and the deep boxes 1 and 4, and 4 and 5). Note that the diffusive fluxes are bidirectional, in order to allow for the exchange of heat and salt, but not of mass. This results in the following set of differential equations for the temperatures and salinities of the boxes:

$$\dot{T}_1 = \frac{m + D_h}{V_1}(T_4 - T_1) + \frac{D_h}{V_1}(T_3 - T_1) + \lambda_1\left(T_1^* - T_1\right) \tag{31}$$

$$\dot{T}_2 = \frac{m + D_h}{V_2}(T_3 - T_2) + \frac{D_v^2}{V_2}(T_5 - T_2) + \lambda_2\left(T_2^* - T_2\right) \tag{32}$$

$$\dot{T}_3 = \frac{m + D_h}{V_3}(T_1 - T_3) + \frac{D_h}{V_3}(T_2 - T_3) + \frac{D_v^3}{V_3}(T_4 - T_3) + \lambda_3\left(T_3^* - T_3\right) \tag{33}$$

$$\dot{T}_4 = \frac{m + D_h}{V_4}(T_5 - T_4) + \frac{D_h}{V_4}(T_1 - T_4) + \frac{D_v^3}{V_4}(T_3 - T_4) \tag{34}$$

$$\dot{T}_5 = \frac{m + D_v^2}{V_5}(T_2 - T_5) + \frac{D_h}{V_5}(T_4 - T_5) \tag{35}$$

$$\dot{S}_1 = \frac{m + D_h}{V_1}(S_4 - S_1) + \frac{D_h}{V_1}(S_3 - S_1) + \frac{S_0 F_1}{V_1} \tag{36}$$

$$\dot{S}_2 = \frac{m + D_h}{V_2}(S_3 - S_2) + \frac{D_v^2}{V_2}(S_5 - S_2) - \frac{S_0 F_2}{V_2} \tag{37}$$

$$\dot{S}_3 = \frac{m + D_h}{V_3}(S_1 - S_3) + \frac{D_h}{V_3}(S_2 - S_3) + \frac{D_v^3}{V_3}(S_4 - S_3) - \frac{S_0(F_1 - F_2)}{V_3} \tag{38}$$

$$\dot{S}_4 = \frac{m + D_h}{V_4}(S_5 - S_4) + \frac{D_h}{V_4}(S_1 - S_4) + \frac{D_v^3}{V_4}(S_3 - S_4) \tag{39}$$

$$\dot{S}_5 = \frac{m + D_v^2}{V_5}(S_2 - S_5) + \frac{D_h}{V_5}(S_4 - S_5) \ . \tag{40}$$

Here, D_v^2, D_v^3 denote the vertical diffusive fluxes between the northern and the low-latitude boxes, respectively, and D_h^s the horizontal diffusive flux.

Since the northern box is split into two parts, the pressure at its bottom results from the volume-averaged densities of boxes 2 and 5. Equation (1) is therefore rewritten as:

$$\rho_0 m = k\left(\frac{\rho_2 V_2 + \rho_5 V_5}{V_2 + V_5} - \rho_1\right) \ . \tag{41}$$

Note that in the limit of very large vertical mixing in the northern box (i.e. $\rho_5 \approx \rho_2$), Eqs. (1) and (41) are identical.

For the derivation of the fluxes D_v^i and D_h from the diffusion coefficients in the units they are usually expressed in (i.e. $m^2\ s^{-1}$), we need to specify a length. For D_v^i, $i \in \{2, 3\}$, this is taken to be the relation between box area A_i and average thickness $\bar{z}$:

$$D_v^2 = \frac{A_2}{\bar{z}} K_v^2 \tag{42}$$

$$D_v^3 = \frac{A_3}{\bar{z}} K_v^3 \ , \tag{43}$$

where K_v^2 and K_v^3 are the vertical diffusion coefficients for the northern and low-latitude boxes, respectively. For K_v^2 we use a value of $10^{-4}\ m^2\ s^{-1}$ or $10^{-2}\ m^2\ s^{-1}$ for the convecting and non-convecting case, and for K_v^3 a value of $10^{-4}\ m^2\ s^{-1}$.

For the calculation of the horizontal diffusive flux D_h, we use the relation between the area $B = \bar{x}\bar{z}$, where $\bar{x}$ is the average zonal box extent, and the average meridional box extent $\bar{y}$:

$$D_h = \frac{B}{\bar{y}} K_h \ , \tag{44}$$

where K_h is the horizontal diffusion coefficient. For K_h, $\bar{x}$, $\bar{y}$, $\bar{z}$ we use values of $10^{-1}\ m^2\ s^{-1}$, 6×10^6 m, 5×10^6 m and 1.5×10^6 m, respectively.

Note that for the investigation of the effects of the individual physical processes, as attempted in Section 6, we applied the following model settings:

- Northern vertical mixing: $K_v^3 = K_h = 0$.
- Mid-latitude vertical mixing: four boxes (i.e. $V_5 \to 0$, $V_4 \to V_4 + V_5$), $K_h = 0$.
- Horizontal mixing: four boxes, $K_v^3 = 0$.

References

Dansgaard W, Johnsen S, Clausen H, Dahl-Jensen N, Gundestrup NS, Hammer C, Hvidberg C, Steffensen J, Sveinbjoernsdottir A, Jouzel J, Bond G (1993) Evidence for general instability of past climate from a 250-kyr ice-core record. Nature 364: 218–220

Dixon K, Delworth T, Spelman M, Stouffer R (1999) The influence of transient surface fluxes on North Atlantic overturning in a coupled GCM climate-change experiment. Geophys Res Let (2)17: 2749–2752

Ganachaud A, Wunsch C (2000) Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. Nature 408: 453–457

Ganopolski A, Petoukhov V, Rahmstorf S, Brovkin V, Claussen M, Eliseev A, Kubatzki C (2001) CLIMBER-2: a climate system model of intermediate complexity, part II. Model sensitivity. Climate Dynamics 17: 735–751

Gargett A, Ferron B (1996) The effect of differential vertical diffusion of T and S in a box model of thermohaline circulation J Mar Res 54: 827–866

Gildor H, Tziperman E (2001) A sea-ice climate switch mechanism for the 100-kyr glacial cycles. J Geophys Res 105(C5): 9117–9133

Griffies M, Tziperman E (1995) A linear thermohaline oscillator driven by stochastic atmospheric forcing. J Climate 8: 2440–2453

Haney R (1971) Surface thermal boundary condition for ocean circulation models. J Phys Oceanogr 1: 241–248

Harvey L (1988) A semi-analytic energy balance climate model with explicit sea-ice and snow physics. J Climate 1: 1065–1085

Hooss G, Voss R, Hasselmann K, Maier-Reimer E, Joos F (2001) A nonlinear impulse response model of the coupled carbon cycle – climate system (NICCS). Climate Dynamics 18: 189–202

Houghton J, Ding Y, Griggs D, Noguer M, van der Linden P, Dai X, Maskell K, Johnson C (eds) (2001) Climate change 2001: the scientific basis – Contribution of working group I to the third assessment report of the IPCC. Cambridge University, Cambridge

Huang R, Luyten J, Stommel H (1992) Multiple equilibrium states in combined thermal and saline circulation. J Phys Oceanogr 22: 231–246

Hughes T, Weaver A (1996) Sea-surface temperature-evaporation feedback and the ocean's thermohaline circulation. J Phys Oceanogr 26: 644–654

Johannessen OM, Shalina EV, Miles MW (1999) Satellite evidence for an Arctic sea-ice cover in transformation. Science 286: 1937–1939

Keller K, Tan K, Morel F, Bradford D (2000) Preserving the ocean circulation: implications for climate policy. Climatic Change 47: 17–43

Kleinen T, Held H, Petschel-Held G (2003) The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. Ocean Dynamics (In press)

Knutti R, Stocker T (2002) Limited predictability of the future thermohaline circulation close to an instability threshold. J Climate 15(2): 179–186

Latif M, Röckner E, Mikolajewicz U, Voss R (2000) Tropical Stabilization of the Thermohaline circulation in a greenhouse Warming Simulation. J Climate 13: 1809–1813

Ledwell J, Watson A, Law C (1998) Mixing of a tracer in the pycnocline. J Geophys Res 103(C10): 21499–21529

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

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

Manabe S, Stouffer R (1994) Multiple-century response of a coupled ocean–atmosphere model to an increase of atmospheric carbon dioxide. J Climate 7: 5–23

Manabe S, Stouffer R (1999) The role of thermohaline circulation in climate. Tellus 51(A–B): 91–109

Mastrandrea M, Schneider S (2001) Integrated assessment of abrupt climatic changes. Climate Policy 1: 433–449

McCarthy J, Canziani O, Leary N, Dokken D, White K. (2001) Climate change 2001: impacts adaptation and vulnerability. Contribution of working group II to the third assessment report of the IPCC. Cambridge University, Cambridge

Mikolajewicz U, Voss R (2000) The role of the individual air-sea flux components in CO_2-induced changes of the ocean's circulation and climate. Climate Dynamics 16: 627–642

Mitchell J, Johns T, Eagles M, Ingram W, Davis R (1999) Towards the construction of climate change scenarios. Climatic Change 41: 547–581

Nordhaus W (2000) The economic impacts of abrupt climate change. Paper presented at the Yale-National Bureau for Economic Research Workshop on the Societal Impacts of Abrupt Climate Change, Snowmass, CO, July 24–25, 2000

Peterson B, Holmes R, McClelland J, Vörösmarty C, Lammers R, Shiklomanov A, Shiklomanov I, Rahmstorf S (2002) Increasing river discharge to the Arctic Ocean. Science 298: 2171–2172

Petoukhov V, Ganopolski A, Brovkin V, Claussen M, Eliseev A, Kubatzki C, Rahmstorf S (2000) CLIMBER-2: a climate system model of intermediate complexity, part I. Model description and performance for present climate. Climate Dynamics 16: 1–17

Prange M, Lohmann G, Paul A (2003) Influence of vertical mixing on the thermohaline hysteresis: analysis of an OGCM. J Phys Oceanogr (In Press)

Rahmstorf S (1995) Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. Nature 378: 145–149

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

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

Rahmstorf S, England M (1997) Influence of Southern Hemisphere winds on North Atlantic deep water flow. J Phys Oceanogr 27: 2040–2054

Rahmstorf S, Ganopolski A (1999a) Long-term global warming scenarios computed with an efficient coupled climate model. Climatic Change 43: 353–367

Rahmstorf S, Ganopolski A (1999b) Simple theoretical model may explain apparent climate instability. J Climate 12: 1349–1352

Rahmstorf S, Willebrand J (1995) The role of temperature feedback in stabilizing the thermohaline circulation. J Phys Oceanogr 25: 787–805

Schaeffer M, Selten F, Opsteegh J, Goosse H (2002) Intrinsic limits to predictability of abrupt regional climate change in IPCC SRES scenarios. Geophys Res Lett 29(16): 10.1029/2002GL0 1524

Schellnhuber H-J (1997) Integrated assessment of climate change: regularity and singularity. Paper presented at the Symposium Climate impact research: why, how and when? Berlin-Brandenburg Academy of Sciences and German Academy Leopoldina, Berlin, October 28–29, 1997

Schmittner A, Stocker T (1999) The stability of the thermohaline circulation in global warming experiments. J Climate 12: 1117–1133

Schmittner A, Weaver A (2001) Dependence of multiple climate states on ocean mixing parameters. Geophys Res Lett 28(6): 1027–1030

Schneider S, Thompson S (2000) Simple climate model used in economic studies of global change. In: DeCanio S, Howarth R, Sanstad A, Schneider S, Thompson S (eds) New directions in the economics and integrated assessment of global climate change. Pew Center on Global Climate Change, pp 59–80

Schneider S, Turner B II, Morehouse Garriga H (1998) Imaginable surprise in global change science. J Risk Res 1(2): 165–185

Scott J, Marotzke J, Stone P (1999) Interhemispheric thermohaline circulation in a coupled box model. J Phys Oceanogr 29: 351–365

Shaffer G, Olsen S (2001) Sensitivity of the thermohaline circulation and climate to ocean exchanges in a simple coupled model. Climate Dynamics 17: 433–444

Slawig T, Zickfeld K (2003) Parameter optimization using algorithmic differentiation in a reduced-form model of the Atlantic thermohaline circulation. Submitted.

Smith J, Pitts G (1997) Regional climate change scenarios for vulnerability and adaptation assessments. Climatic Change 36: 3–21

Stammer D, Wunsch C, Giering R, Eckert C, Heimbach P, Marotzke J, Adcroft A, Hill C, Marshall J (2002) Global ocean circulation during 1992–1997 estimated from ocean observations and a general circulation model. J Geophys Res 107(C9): 3118

Stocker T, Schmittner A (1997) Influence of CO_2 emission rates on the stability of the thermohaline circulation. Nature 388: 862–865

Stommel H (1961) Thermohaline convection with two stable regimes of flow. Tellus 13: 224–241

Thorpe RB, Gregory JM, Johns TC, Wood RA, Mitchell JFB (2001) Mechanisms determining the Atlantic thermohaline circulation response to greenhouse gas forcing in a non-flux-adjusted coupled climate model. J Climate 14: 3102–3116

Titz S, Kuhlbrodt T, Feudel U, Rahmstorf S (2002) On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulation. Tellus 54(1): 89–98

Toggweiler J, Samuels B (1995) Effect of drake passage on the global thermohaline circulation. Deep-Sea Res I 42(4): 477–500

Toggweiler J, Samuels B (1998) On the ocean's large-scale circulation near the limit of no vertical mixing. J Phys Oceanogr 28(9): 1832–1852

Tóth F, Petschel-Held G, Bruckner T (1998) Kyoto and the long-term climate stabilization. In: Proceedings of the OECD Workshop on Economic Modeling of Climate Change, Paris, September 17–18, 1998

Tziperman E (1997) Inherently unstable climate behaviour due to weak thermohaline ocean circulation. Nature 386: 592–594

Tziperman E, Toggweiler J, Feliks Y, Bryan K (1994) Instability of the thermohaline circulation with respect to mixed boundary conditions: is it really a problem for realistic models? J Phys Oceanogr 24: 217–232

Van der Veen CJ (2002) Polar ice sheets and global sea level: how well can we predict the future? Global Planetary Changes 12: 165–194

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

Verma A (2002) ADMAT: automatic differentiation in MATLAB using object-oriented methods. http://www.tc.cornell.edu/averma/AD/admatoo.ps

Zickfeld K, Bruckner T (2003) Reducing the risk of abrupt climate change: emissions corridors preserving the Atlantic thermohaline circulation. Integrated Assessment (In Press)