MIT Joint Program on the Science and Policy of Global Change

Thermohaline Circulation Stability: *A Box Model Study*

Part I: ***Uncoupled Model***. **Part II**: ***Coupled Atmosphere-Ocean Model***

Valerio Lucarini and Peter H. Stone

Report No. 99
June 2003

The MIT Joint Program on the Science and Policy of Global Change is an organization for research, independent policy analysis, and public education in global environmental change. It seeks to provide leadership in understanding scientific, economic, and ecological aspects of this difficult issue, and combining them into policy assessments that serve the needs of ongoing national and international discussions. To this end, the Program brings together an interdisciplinary group from two established research centers at MIT: the Center for Global Change Science (CGCS) and the Center for Energy and Environmental Policy Research (CEEPR). These two centers bridge many key areas of the needed intellectual work, and additional essential areas are covered by other MIT departments, by collaboration with the Ecosystems Center of the Marine Biology Laboratory (MBL) at Woods Hole, and by short- and long-term visitors to the Program. The Program involves sponsorship and active participation by industry, government, and non-profit organizations.

To inform processes of policy development and implementation, climate change research needs to focus on improving the prediction of those variables that are most relevant to economic, social, and environmental effects. In turn, the greenhouse gas and atmospheric aerosol assumptions underlying climate analysis need to be related to the economic, technological, and political forces that drive emissions, and to the results of international agreements and mitigation. Further, assessments of possible societal and ecosystem impacts, and analysis of mitigation strategies, need to be based on realistic evaluation of the uncertainties of climate science.

This report is one of a series intended to communicate research results and improve public understanding of climate issues, thereby contributing to informed debate about the climate issue, the uncertainties, and the economic and social implications of policy alternatives. Titles in the Report Series to date are listed on the inside back cover.

Henry D. Jacoby and Ronald G. Prinn,
Program Co-Directors

For more information, please contact the Joint Program Office

Postal Address: Joint Program on the Science and Policy of Global Change
MIT E40-428
77 Massachusetts Avenue
Cambridge MA 02139-4307 (USA)

Location: One Amherst Street, Cambridge
Building E40, Room 428
Massachusetts Institute of Technology

Access: Phone: (617) 253-7492
Fax: (617) 253-9845
E-mail: globalchange@mit.edu
Web site: http://mit.edu/globalchange/

Printed on recycled paper

Thermohaline circulation stability: a box model study - Part I: uncoupled model

Valerio Lucarini
lucarini@mit.edu
and
Peter H. Stone
phstone@mit.edu
Joint Program on the Science and Policy of Global Change, MIT
Cambridge, MA 02138 USA

June 19, 2003

Abstract

A thorough analysis of the stability of the uncoupled Rooth inter-hemispheric 3-box model of Thermohaline Circulation (THC) is presented. The model consists of a northern high latitudes box, a tropical box, and a southern high latitudes box, which respectively correspond to the northern, tropical and southern Atlantic ocean. We adopt restoring boundary conditions for the temperature variables and flux boundary conditions for the salinity variables. We study how the strength of THC changes when the system undergoes forcings that are analogous to those of global warming conditions by applying to the equilibrium state perturbations to the moisture and heat fluxes into the three boxes. In each class of experiments, we determine, using suitably defined metrics, the boundary dividing the set of forcing scenarios that lead the system to equilibria characterized by a THC pattern similar to the present one, from those that drive the system to equilibria where the THC is reversed. Fast increases in the moisture flux into the northern high-latitude box are more effective than slow increases in leading the THC to a breakdown, while the presence of moisture flux increases into the southern high-latitude box strongly inhibits the breakdown and can prevent it, in the case of slow increases in the Northern Hemisphere. Similarly, a high rates of heat flux increase in the North Hemisphere destabilize the system more effectively than low ones, and increases in the heat fluxes in the Southern Hemisphere tends to stabilize the system.

1 Introduction

One of the main issues in the study of climate change is the fate of the Thermohaline Circulation (THC) in the context of global warming [Weaver and Hughes 1992, Manabe and Stouffer 1993, Stocker and Schmittner 1997, Rahmstorf 1997, Rahmstorf 1999a, Rahmstorf 1999b, Rahmstorf 2000, Wang et al. 1999a, Wang et al. 1999b]. The THC plays a crucial role in determining the main features of the North Atlantic climate, since the advection of warm water prevents the formation of sea ice up to very high latitudes even during winter, and the heat exchange between relatively warm water and cold air warms high latitudes [Levitus 1982, Broecker 1994, Stocker 2000]. The THC plays a major role in the global circulation of the oceans as pictured by the conveyor belt scheme [Weaver and Hughes 1992, Stocker 2001]; therefore the effect on the climate of a change of its pattern may be global [Broecker 1997, Manabe and Stouffer 1999b, Cubasch et al. 2001]. The THC is sensitive to changes in the climate since the NADW formation is sensitive to variations in air temperature and in precipitation in the Atlantic basin [Rahmstorf and Willebrand 1995, Rahmstorf 1996]. There are several paleoclimatic data sets indicating that changes in the patterns or collapses of the THC coincide with large variations in climate, especially in the North Atlantic region [Broecker et al. 1985, Boyle and Keigwin 1987, Keigwin et al. 1994, Rahmstorf 2002].

Box models have historically played a major role in the understanding of the fundamental dynamics of the THC [Weaver and Hughes 1992]: the Stommel two-box oceanic model [Stommel 1961] built the first conceptual bridge between the THC strength and density gradients. The oceanic model proposed by Rooth [Rooth 1982] introduced the idea that the driver of the THC was density difference between the high latitude basins of the northern and southern hemispheres, thus implying that the THC is an inter-hemispheric phenomenon. Both the Stommel and Rooth models allow two stable equilibria, one -the warm mode- characterized by downwelling of water in the North Atlantic, which resembles the present oceanic circulation, and the other - the cold mode - characterized by upwelling of water in the North Atlantic. Perturbations to the driving parameters of the system, i.e. freshwater and heat fluxes in the oceanic boxes, can cause transitions between the two regimes.

More recently various GCM experiments have shown that multiple equilibria of the THC are possible [Bryan 1986, Manabe and Stouffer 1988, Stouffer et al. 1991, Stocker and Wright 1991, Manabe and Stouffer 1999a, Marotzke and Willebrand 1991, Hughes and Weaver 1994], and most GCMs have shown that the GHGs radiative forcing could cause a weakening or eventually a shutdown of the THC by the inhibition of the sinking of the water in the Northern Atlantic. Large increases of the moisture flux and/or of the sea surface temperature in the downwelling regions are the driving mechanisms of such a process. Nevertheless in a recent global warming simulation run with the *ECHAM-4* model the THC remained almost constant since the destabilizing mechanisms have been offset by the increase of the salinity

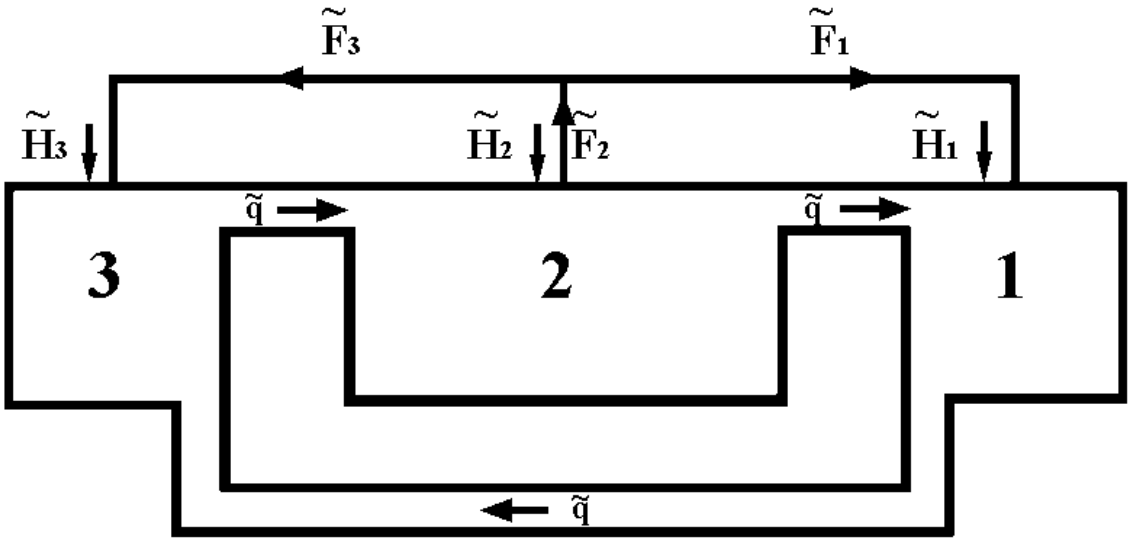

Figure 1: Schematic picture of the interhemispheric box model

advected into the downwelling region due to net freshwater export from the whole North Atlantic [Latif et al. 2000]. The weakening of the THC could greatly limit the regional warming in North Atlantic [Rahmstorf 1997, Rahmstorf 1999a, Rahmstorf 1999b, Rahmstorf 2000].

Analysis of GCM data have also indicated that the THC is an inter-hemispheric phenomenon, and in some cases it has been found that the THC strength is approximatively proportional to the density difference between the Northern and Southern Atlantic [Hughes and Weaver 1994, Rahmstorf 1996, Klinger and Marotzke 1999, Wang et al. 1999a], thus supporting Rooth's approach. While hemispheric box models like Stommel's have been extensively studied [Weaver and Hughes 1992, Nakamura et al. 1994, Marotzke 1996, Krasovskiy and Stone 1998, Tziperman and Gildor 2002], and the stability of their THC has been assessed in the context of various levels of complexity in the representation of the coupling between the ocean and the atmosphere, the literature on inter-hemispheric box models is far less abundant [Rahmstorf 1996, Scott et al. 1999, Stone and Krasovskiy 1999, Titz et al. 2002a, Titz et al. 2002b], and relatively few box-model studies have included the effect of coupling an inter-hemispheric model of the ocean to the atmosphere [Scott et al. 1999, Stone and Krasovskiy 1999, Tziperman and Gildor 2002].

In this study we perform an analysis of the stability of the present pattern of the THC against changes to the heat and freshwater atmospheric fluxes using an uncoupled version of the Rooth oceanic model. The model here considered has no real explicit coupling between the ocean and the atmosphere: the atmospheric freshwater fluxes are prescribed and the atmospheric heat fluxes relax the oceanic temperatures towards prescribed values. In Part II of this paper we present a 3-box model having explicit coupling between the ocean and the atmosphere: the atmospheric freshwater and heat fluxes are expressed as functions of

the oceanic temperatures.

We explicitly analyze, to the extent that the simplicity of the model allows, what is the role of the spatial patterns *and* of the rates of increase of the forcings to the driving parameters and characterize the response of the system. We consider a suitably defined metric and determine which are the thresholds beyond which we have destabilization of the warm mode of the THC. We underline that our treatment goes beyond quasi-static analysis since the effect of the rate of forcing is explicitly addressed, so that we analyze how the effects on the system of slow [Wang et al. 1999a] and rapid [Wiebe and Weaver 1999] changes join on; in particular in the limit of very slow changes, our results coincide with those that could be obtained with the analysis of the bifurcations of the system [Stone and Krasovskiy 1999,Scott et al. 1999,Titz et al. 2002a,Titz et al. 2002b]. We underline that since we consider only poleward freshwater fluxes, our study differs from Rahmstorf's [Rahmstorf 1996]. Only relatively few studies have explicitly addressed how the spatial [Rahmstorf 1995, Rahmstorf 1996, Rahmstorf 1997] or temporal [Stocker and Schmittner 1997] patterns of forcing determine the response of the system in the context of more complex models. Obvious limitations due to computing time did not allow an extensive exploration of the parameter space of the forcings applied, nor a parametric study of the influence of both spatial *and* temporal patterns.

Our paper is organized as follows. In the second section we provide a description and the general mathematical formulation of the dynamics of the three-box model used in this study. In the third section we describe the parametrization of the physical processes. In the forth section we analyze the feedbacks of the system. In the fifth section we describe the stability properties of the system when it is forced with perturbations in the freshwater flux. In the sixth section we describe the stability properties of the system when it is forced with perturbations in the heat flux. In the seventh section we present our conclusions.

2 Model description

The three-box model consists of a northern high latitude box (box 1), a tropical box (box 2), and a southern high latitude box (box 3). The volume of the two high latitudes boxes is the same, and is $1/V$ times the volume of the tropical box. We choose $V = 2$, so that box 1, box 2 and box 3 respectively can be thought as describing the portions of an ocean like the Atlantic north of $30°N$, between $30°N$ and $30°S$, and south of $30°S$. At the eastern and western boundaries of the oceanic boxes there is land; our oceanic system spans $60°$ in longitude, so that it covers $\epsilon = 1/6$ of the total surface of the planet. The boxes are 5000 m deep, so that the total mass $M_{i=1,3} = M$ of the water contained in each of the high-latitude boxes is $\approx 1.1 \cdot 10^{20}\ Kg$, while $M_2 = V \cdot M_{i=1,3} = 2 \cdot M$. The tropical box is connected to both high

latitude boxes; moreover the two high latitude boxes are connected by a deep current passage (which bypasses the tropical box) containing a negligible mass. The three boxes are assumed to be well mixed, so that the *polar halocline disaster* [Bryan 1986, Zhang et al. 1993] is automatically excluded from the range of phenomena that can be described by this model. The physical state of the box i is described by its temperature T_i and its salinity S_i; the box i receives from the atmosphere the net flux of heat $\tilde{H}_i$ and the net flux of freshwater $\tilde{F}_i$; the freshwater fluxes $\tilde{F}_i$ globally sum up to 0, since the freshwater is a globally conserved quantity. The box i is subjected to the oceanic advection of heat and salt from the upstream box through the THC, whose strength is $\tilde{q}$.

In Figure 1 we present a scheme of our system in the northern sinking pattern: note that the arrows representative of the freshwater fluxes are arranged in such a way that the conservation law is automatically included in the graph. The dynamics of the system is described by the tendency equations for the heat and the salinity of each box. We divide the three heat tendency equations by $c_p \cdot M_i$, where c_p is the constant pressure specific heat of water per unit of mass, and in the salinity tendency equations we neglect the contribution of the freshwater fluxes in the mass balance, so that virtual salinity fluxes and freshwater fluxes are equivalent [Marotzke 1996, Rahmstorf 1996, Scott et al. 1999]. We obtain the following final form for the temperature and salinity tendency equations for the three boxes [Scott et al. 1999]:

$$\dot{T}_1 = \begin{cases} q(T_2 - T_1) + H_1, & q \geq 0 \\ q(T_1 - T_3) + H_1, & q < 0 \end{cases} \tag{1}$$

$$\dot{T}_2 = \begin{cases} \frac{q}{V}(T_3 - T_2) + H_2, & q \geq 0 \\ \frac{q}{V}(T_2 - T_1) + H_2, & q < 0 \end{cases} \tag{2}$$

$$\dot{T}_3 = \begin{cases} q(T_1 - T_3) + H_3, & q \geq 0 \\ q(T_3 - T_2) + H_3, & q < 0 \end{cases} \tag{3}$$

$$\dot{S}_1 = \begin{cases} q(S_2 - S_1) - F_1, & q \geq 0 \\ q(S_1 - S_3) - F_1, & q < 0 \end{cases} \tag{4}$$

$$\dot{S}_2 = \begin{cases} \frac{q}{V}(S_3 - S_2) - F_2, & q \geq 0 \\ \frac{q}{V}(S_2 - S_1) - F_2, & q < 0 \end{cases} \tag{5}$$

$$\dot{S}_3 = \begin{cases} q(S_1 - S_3) - F_3, & q \geq 0 \\ q(S_3 - S_2) - F_3, & q < 0 \end{cases} \tag{6}$$

where $q = \rho_0 \cdot \tilde{q}/M$, $H_i = \tilde{H}_i/(c_p \cdot M)$, and $F_i = \rho_0 \cdot S_0 \cdot \tilde{F}_i/M$.

We impose that the average salinity is a conserved quantity; this means that $\dot{S}_1 + V\dot{S}_2 + \dot{S}_3$=0, which implies that:

$$F_2 = -\frac{1}{V}(F_1 + F_3). \tag{7}$$

This conservation law holds at every time, and so rules out the possibility of including in our study the effects of the melting of continental ice sheets. On the other hand we do not impose any strict conservation law for the transient behavior of the global heat budget of the ocean, since we essentially want to include the effects of radiative imbalances. We note that the system can (asymptotically) reach an equilibrium only if its feedbacks can drive it towards a state in which the following equation holds:

$$H_2 = -\frac{1}{V}(H_1 + H_3). \tag{8}$$

The strength of the specific THC, as in the Stommel model, is parameterized as proportional to the difference between the density of the water contained in the box 1 and the density of the water contained in the box 3. Given that the water density can approximately be expressed as:

$$\rho(T, S) \approx \rho_0(1 - \alpha T + \beta S); \tag{9}$$

where α and β are respectively the thermal and haline expansion coefficients, set respectively to $1.5 \cdot 10^{-4}\ ^\circ C^{-1}$ and $8 \cdot 10^{-4}\ psu^{-1}$, and ρ_0 is a standard reference density, we obtain for the normalized THC strength q the following relation:

$$q = \frac{k}{\rho_0}(\rho_1 - \rho_3) = k(\alpha(T_3 - T_1) + \beta(S_1 - S_3)), \tag{10}$$

where k is the hydraulic constant,chosen to obtain a reasonable value of the THC strength. The northern (southern) sinking state of the circulation is therefore characterized by $q >$ ($<$) 0.

Considering the case $q > 0$, we obtain a diagnostic relation for steady state q by setting to 0 equations 3 and 6, and using equation 9:

$$q_{eq} = (k(\alpha((H_3)_{eq} + \beta(F_3)_{eq})))^{1/2}; \tag{11}$$

in the case $q < 0$, a diagnostic relation for the steady state q can be obtained using the same

procedure as above but setting to zero equations 1 and 4:

$$q_{eq} = -(k((\alpha(H_1)_{eq} + \beta(F_1)_{eq}))^{1/2}. \tag{12}$$

We conclude that the *equilibrium* value of the THC strength is determined by the equilibrium values of the heat and freshwater fluxes of the box where we have upwelling of water. These results generalize the expressions given by Rahmstorf [Rahmstorf 1996]. The transient evolution of q is determined by its tendency equation:

$$\begin{aligned} \dot{q} &= -q^2 + kq(\alpha(T_1 - T_2) - \beta(S_1 - S_2)) + \\ &\quad + k(\alpha(H_3 - H_1) + \beta(F_3 - F_1)) \qquad , \quad q \geq 0, \end{aligned} \tag{13}$$

$$\begin{aligned} \dot{q} &= q^2 + kq(\alpha(T_3 - T_2) - \beta(S_3 - S_2)) + \\ &\quad + k(\alpha(H_1 - H_3) + \beta(F_1 - F_3)) \qquad , \quad q < 0, \end{aligned} \tag{14}$$

We observe that the difference between the forcings applied to the freshwater and heat fluxes into the two boxes 1 and 3 directly effect the evolution of q; the presence of terms involving the gradient of temperature and salinity between box 2 and box 1 (box 3) if $q > 0$ ($q < 0$) breaks the symmetry of the role played by the two high latitude boxes. In our experiments we perturb an initial equilibrium state -which is the same for all the experiments in all of the three versions of the model analyzed- by changing at various rates the parameters controlling the fluxes H_i and/or F_i and observe under which conditions we obtain a reversal of the THC. The reversal of the THC causes a sudden cooling and freshening in the box 1 and a sudden warming and increase of the salinity in the box 3, because the former loses and the latter receives the advection of the tropical warm and salty water.

3 Parametrization of the physical processes

We perform our study considering the original formulation of the Rooth Model [Rooth 1982]. For a more detailed description of the model see also the paper by Scott et al. [Scott et al. 1999]. In this model the atmosphere has a negligible heat capacity and water content compared to the ocean, and it only transports heat and moisture; the land has also negligible heat capacity, so that all the heat fluxes end up in the oceanic fraction of the total planetary surface. The heat flux into the box i is described by a newtonian relaxation law of the form $H_i = \lambda_i(\tau_i - T_i)$ where the parameter τ_i is the target temperature [Bretherton 1982], which represents a climatologically reasonable value towards which the box temperature is relaxed, and the parameter λ_i quantifies the efficacy of the relaxation. We then make a

choice of the values of these parameters and integrate from reasonable initial conditions in order to define a northern sinking equilibrium state that is the baseline for all the rest of the study. Along the lines of the paper of Scott et al. [Scott et al. 1999] and following the parametrization introduced by Marotzke [Marotzke 1996], we choose $\tau_1 = \tau_3 = 0°C$ and $\tau_2 = 30°\ C$, and select $\lambda_1 = \lambda_2 = \lambda_3 = \lambda = 1.29 \cdot 10^{-9}\ s^{-1}$, which mimics the combined effect of radiative heat transfer and of a meridional heat transport parameterized as linear with the meridional temperature gradient [Marotzke and Stone 1995, Marotzke 1996, Scott et al. 1999]. In physical terms, λ corresponds to a restoring coefficient of $\approx 4.3\ Wm^{-2}\ °C^{-1}$ affecting the whole planetary surface, which translates into an effective $\tilde{\lambda} \approx 1/\epsilon \cdot 4.3\ Wm^{-2}\ °C^{-1} = 25.8\ Wm^{-2}\ °C^{-1}$ relative to the oceanic surface fraction only.

Using equations 1, 2, and 3, we can obtain the following tendency equation for the globally averaged temperature $T_M = (T_1 + VT_2 + T_3)/(2 + V)$:

$$\dot{T}_M = \lambda(\tau_M - T_M) \tag{15}$$

where τ_M is the average target temperature; the average temperature relaxes towards its target value in about 25 y, which is very close to the value considered in [Titz et al. 2002a]; we observe that in the case of an ocean represented with vertical resolution, our heat flux parametrization corresponds to a restoring time of ≈ 3 months for a 50 m deep surface level [Marotzke and Stone 1995].

The virtual salinity fluxes F_i are not functions of any of the state variables of the system, and are given parameters. We choose $F_1 = 13.5 \cdot 10^{-11}\ psu\ s^{-1}$ and $F_3 = 9 \cdot 10^{-11}\ psu\ s^{-1}$, which respectively correspond to net atmospheric freshwater fluxes of $\approx 0.40\ Sv$ into the box 1 and of $\approx 0.27\ Sv$ into the box 3. The rationale for the chosen asymmetry in the freshwater fluxes lies in the analysis of the hydrology of the two hemispheres done by Baumgartner and Reiche [Baumgartner and Reichel 1975]. Thus we have a mixed boundary conditions (BC) dynamical system, since the salinities tendency equations contain flux terms, while the temperatures tendency equation have restoring conditions.

Choosing $k = 1.5 \cdot 10^{-6}\ s^{-1}$, at equilibrium we have $q \approx 1.47 \cdot 10^{-10}\ s^{-1}$, [Scott et al. 1999] which corresponds to an oceanic flushing time scale of $\approx 300\ y$; in physical terms the previous value for q describes a THC in the northern sinking pattern with a strength of $\approx 15.5\ Sv$: this value agrees reasonably well with estimates coming from observations [Roemmich and Wunsch 1985, Macdonald and Wunsch 1996]. In this equilibrium state boxes 1 and 3 receive a net poleward oceanic heat advection of $\approx 1.58\ PW$, and $\approx 0.17\ PW$ respectively: while the former agrees reasonably well with estimates, the latter does not, because the South Atlantic has a weak equatorward flux [Macdonald and Wunsch 1996]. This mistaken sign implies that our model will not be able to capture any dephasing between North and South hemisphere high latitudes temperature signals due to changes in the THC strength

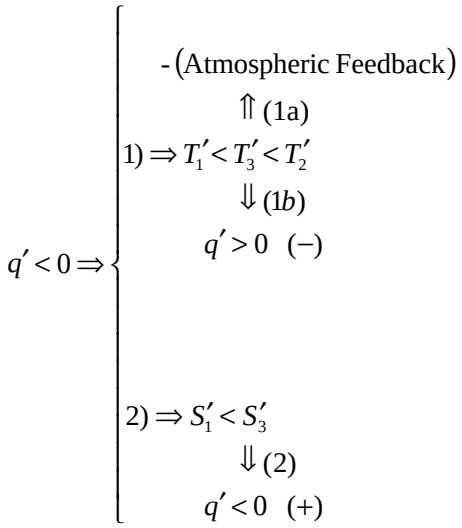

Figure 2: Feedbacks scheme of the model

as those occurred in the past [Vidal et al. 1999, Stocker 2000, Stocker 2002]. The equilibrium temperature of box 1 is larger than the equilibrium temperature of box 3, and the same inequality holds for the salinities, the main reason being that box 1 receives directly the warm and salty tropical water. Taking into account 9 and considering that the actual equilibrium q is positive, we conclude that the THC is haline-dominated: the ratio between the absolute value of the thermal and of the haline contribution on the right side of 9 is ≈ 0.8. We report in table 1 the values of the main model constants, while in table 2 we present the fundamental parameters characterizing this equilibrium state.

4 Feedbacks of the system

The newtonian relaxation law for the temperature implies that the atmosphere has a negative thermic feedback:

- T_1 increases more than $T_3 \Rightarrow H_1$ decreases more than $H_3 \Rightarrow T_1$ decreases more than T_3

Apart from the mean flow feedback, the reaction of the ocean to a decrease in the strength of the THC is depicted in figure 2 and can be described as follows:

- q decreases $\Rightarrow$

1. T_1 decreases more than $T_3 \Rightarrow$
   (a) q increases

(b) the change is limited by the Atmospheric feedback;

2. S_1 decreases more than $S_3 \Rightarrow q$ decreases;

Generally the feedback (1) is negative and stronger than the positive feedback (2), even if it triggers the previously described atmospheric mechanism. We observe that the strength of the second order feedback mechanism $1b$ is relevant in establishing the overall stability of the THC [Tziperman et al. 1994]; a very strong temperature-restoring atmospheric feedback like that in our model tends to decrease the stability of the system [Nakamura et al. 1994, Rahmstorf 2000]. The feedbacks (1) + (2) are governed by q, so that their time scale is around the flushing time of the oceanic boxes; a fast perturbation avoids those feedbacks.

We perform two sets of experiment in order to simulate global warming conditions. In the first, we increase the freshwater fluxes into the two high-latitude boxes. The rationale for this forcing is in the fact that an increase in the global mean temperature produces an increase in the moisture capacity of the atmosphere, and this is likely to cause the enhancement of the hydrological cycle, and especially in the transport of moisture from the tropics to the high latitudes. In the second, we represent the purely thermic effects of global warming by setting the increase in the target temperatures of the two high-latitude boxes larger than that of the tropical boxes, since virtually all GCM simulations forecast a larger increase in temperatures in the high latitudes (the so-called *polar amplification*).

5 Freshwater flux forcing

We force the previously defined equilibrium state with a net freshwater flux into the two high latitudes boxes which increases linearly with time; this is obtained by prescribing:

$$F_i(t) \quad = \quad \begin{cases} F_i(0) + F_i^t \cdot t, & 0 \leq t \leq t_0 \quad i = 1,3 \\ F_i(0) + F_i^t \cdot t_0 & t > t_0 \quad i = 1,3 \end{cases} \tag{16}$$

We impose the conservation of the globally averaged salinity of the oceanic system by requiring that equation 7 hold at all times. When the perturbation is over, we reach a final state that can be either qualitatively similar to the initial northern sinking equilibrium or radically different, i.e. presenting a reversed THC. Each perturbation can be uniquely specified by a set of three parameters, such as the triplet 1 $[t_0, \Delta F_3/\Delta F_1, \Delta F_1]$, or the triplet 2 $[t_0, \Delta F_3/\Delta F_1, F_1^t]$, where $\Delta F_i \equiv F_i^t \cdot t_0$ is the total change of the freshwater flux into box i. In order to describe a global warming scenario, we explore the effect of positive ΔF_1 and ΔF_3. Freshwater forcing into box 1 tends to destabilize the THC, since it induces a freshening and so a decrease of the density of the water there contained; the converse holds for increases in the net freshwater flux into box 3 (see 11). If the freshwater forcing is larger

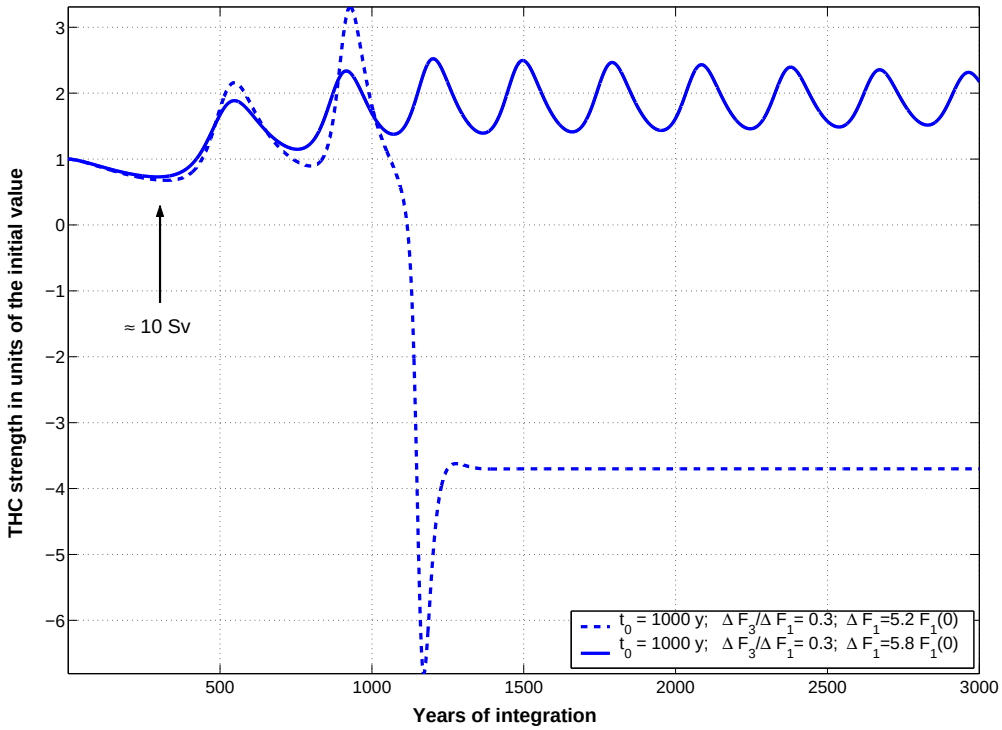

Figure 3: Evolution of the THC strength under a super- and sub-critical freshwater flux forcing

in the box 3, we cannot obtain a reversal of the THC, so that we limit ourselves to the case $\Delta F_3/\Delta F_1 < 1$. These perturbations cause an initial decrease in the THC strength q and so trigger the oceanic feedbacks described in Figure 2:

- F_1's increase larger than F_3's $\Rightarrow$ S_1 decreases more than S_3 $\Rightarrow$ q decreases.

We present in Figure 3 two trajectories of q describing the evolution of the system from the initial equilibrium state when two different forcings, one subcritical and one supercritical, are applied. The presence of slowly decaying large oscillations in the q trajectory of the subcritical perturbation shows the existence of strong feedbacks acting on time scales of the order of the flushing time of the oceanic boxes ($\approx$ 300 years); in the case of the supercritical perturbation the oscillations of the value of q grow in size until the system collapses to a southern sinking equilibrium. The two trajectories separate $\approx$ 400 years after the beginning of the perturbation, when the negative feedback due to the decrease of advection of warm, salty water in box 1 becomes of critical relevance. The threshold value for q is of the order of 10 Sv: more complex models have shown that weak THC are not stable [Rahmstorf 1995, Tziperman 2000a]. When the THC does not change sign, it recovers to its initial unperturbed value (if $\Delta F_3 = 0$) or overshoots it ($\Delta F_3 > 0$), as can be deduced from 11; this reminds us of results obtained with much more complex models [Wiebe and Weaver 1999].

5.1 Critical Forcings

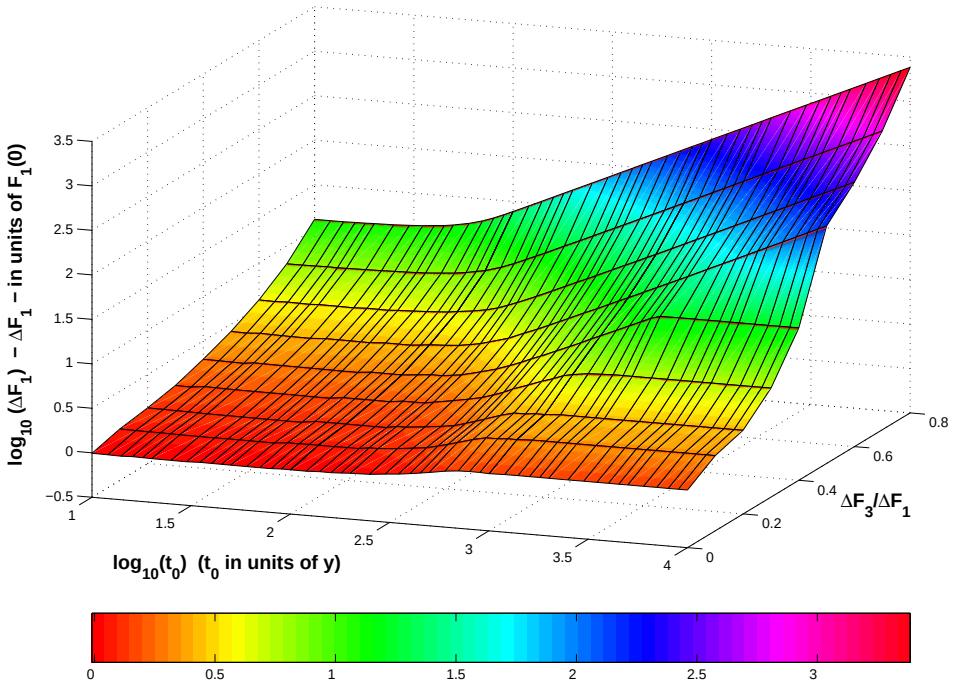

Figure 4: Critical values of the total increase of the freshwater flux

Figures 4 and 5 present, using respectively the coordinates described by the triplets 1 and 2, the manifold dividing the subcritical from the supercritical forcings. In Figure 4 we observe that, as expected, for a given ΔF_1, the lower is the value of the ratio $\Delta F_3/\Delta F_1$, the lower is the total change ΔF_1 needed to obtain the reversal of the THC. For a given value of the ratio $\Delta F_3/\Delta F_1$, more rapidly increasing perturbations (larger F_1^t) are more effective in disrupting the circulation. For values of $\Delta F_3/\Delta F_1 \leq 0.4$ we see a changeover between a *slow* and a *fast* regime, geometrically described by a relatively steep portion dividing two portions of the manifold having little x-dependence. Figure 5 shows more clearly that for large values of $\Delta F_3/\Delta F_1$ (≥ 0.5) the collapse of the THC occurs only for fast increases, because of the presence of a threshold in the rate of increase depicted by the little t_0-dependence of the manifold. The changeover corresponds to $t_0 \approx 250$ y, which essentially matches the flushing time of the oceanic boxes [Scott et al. 1999]. For low values of $\Delta F_3/\Delta F_1$ the transition takes place even in the case of a very low F_1^t because, as shown in the bifurcation analysis performed by [Scott et al. 1999], there is a critical ratio $\Delta F_3/\Delta F_1$ below which the northern sinking equilibrium becomes unstable: such a ratio can be reached after a long enough t_0 if $\Delta F_3/\Delta F_1$ is low enough.

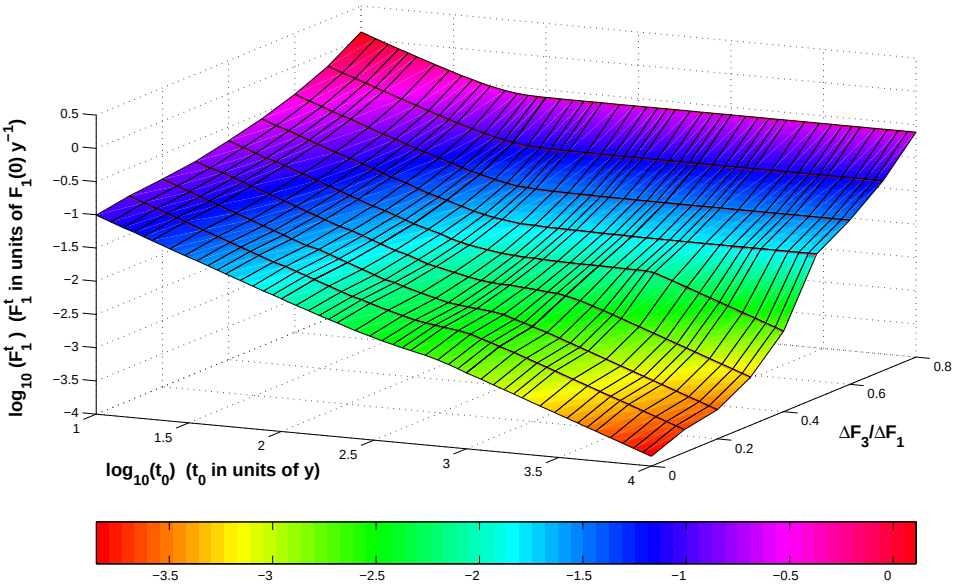

Figure 5: Critical values of the rate of increase of the freshwater flux

5.2 Bifurcation

In figure 6 we present the bifurcation diagrams describing the stability of this model to a slow freshwater flux forcing applied in both boxes 1 and 3; in this case the perturbations applied are quasi-static, i.e. t_0 is very large. We have plotted 5 curves, corresponding to $\Delta F_3/\Delta F_1 = [0, 0.1, 0.2, 0.3, 0.4]$. For a given curve, each point (x, y) plotted represents a stable state of the system having $q = y$ and $F_1 = x$. We scale all the graphs so that the common initial equilibrium state is the point $(1, 1)$. For a range of F_1, which depends on the value of $\Delta F_3/\Delta F_1$, the system is bistable, i.e. it possesses two distinct stable states, one having $q > 0$, the other one having $q < 0$. The history of the system determines which of the two stable states is realized. The boundaries of the domain in F_1 where the system is bistable correspond to subcritical Hopf bifurcations, which drive the system from the northern (southern) sinking equilibrium to the southern (northern) sinking equilibrium if F_1 crosses the right (left) boundary of the domain of bistability. As previously explained if $\Delta F_3/\Delta F_1$ is large enough (≥ 0.5) slow increases of the freshwater fluxes do not destabilize the system, so that no bifurcations are present. The presence of an hysteresis for slow freshwater flux perturbations has been presented both for simple [Rahmstorf 1995, Rahmstorf 1996, Scott et al. 1999, Titz et al. 2002a] (albeit with differences in the model formulation in some cases) as well as more complex interhemispheric models [Stocker and Wright 1991, Mikolajewicz and Maier-Reimer 1994, Rahmstorf 1995, Rahmstorf 1996]; the inclusion of changes in the freshwater flux into both high latitude boxes presented in this work had not been previously explicitly considered.

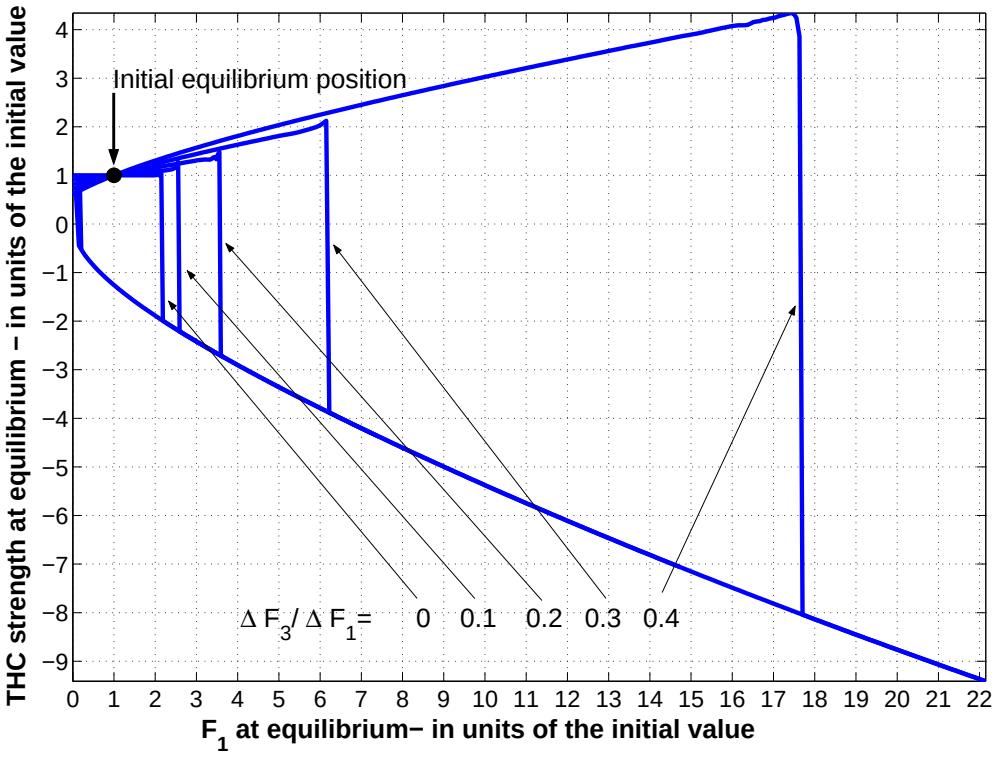

Figure 6: Bifurcation diagram of freshwater flux forcings

5.3 Relevance of the mixed boundary conditions

In order to explore the relevance of the mixed BC for the stability of the THC pattern against freshwater flux perturbations [Tziperman et al. 1994, Nakamura et al. 1994, Rahmstorf 2000], we have considered also a second version of the model where we replace the expression of the heat flux in the box i defined as $\lambda(\tau_i - T_i)$ with a free parameter H_i, thus realizing a flux BC model. Choosing the value of the initial H_i equal to the value of $\lambda(\tau_i - T_i)$ of the previous version of the model at equilibrium, we obtain the same equilibrium solution. We then perturb this equilibrium solution with freshwater flux forcing as in the previous case, and with explicit heat flux forcing. Since the heat flux H_i does not depend on T_i, the positive feedback ($1b$) described at the beginning of the section is off, therefore we expect this version of the system to be more stable than the previous one. Repeating the analysis of stability to freshwater flux perturbations for the flux BC model, it is indeed confirmed that temperature restoring conditions decrease the stability of the model, in agreement with the results of [Tziperman et al. 1994, Nakamura et al. 1994, Rahmstorf 2000]. Figure 7 shows the ratio between the values of the critical values of ΔF_1 presented in figure 4 and the corresponding values obtained with the flux BC model. the average value of this ratio over all the domain is ≈ 0.3. Observing figure 7 for the $\Delta F_3/\Delta F_1$ values $[0.3, 0.4]$, we see that the ratio goes to 0 with increasing t_0. This occurs because for these values the threshold in the rate of increase

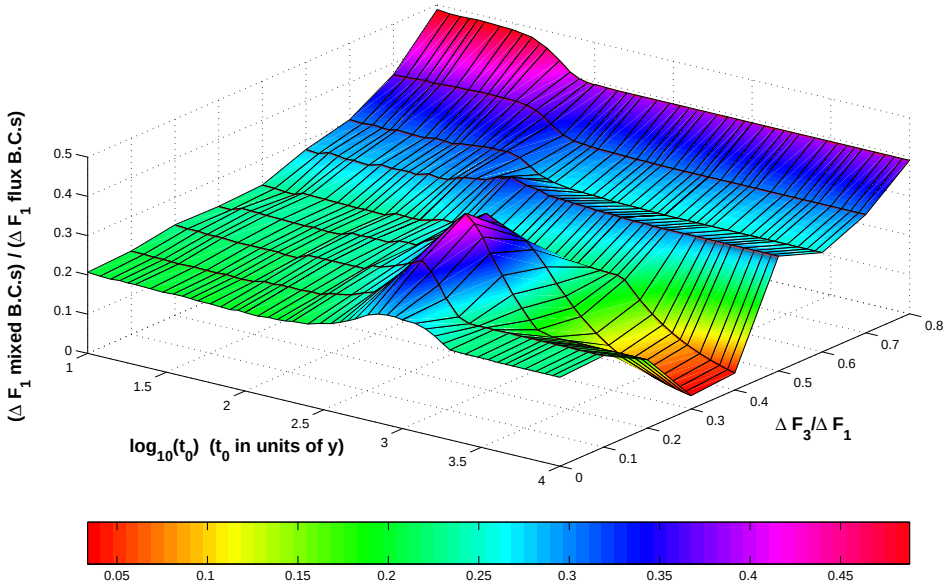

Figure 7: Comparison between the stability of the mixed and flux BCs uncoupled model to increases of the freshwater lfuxes

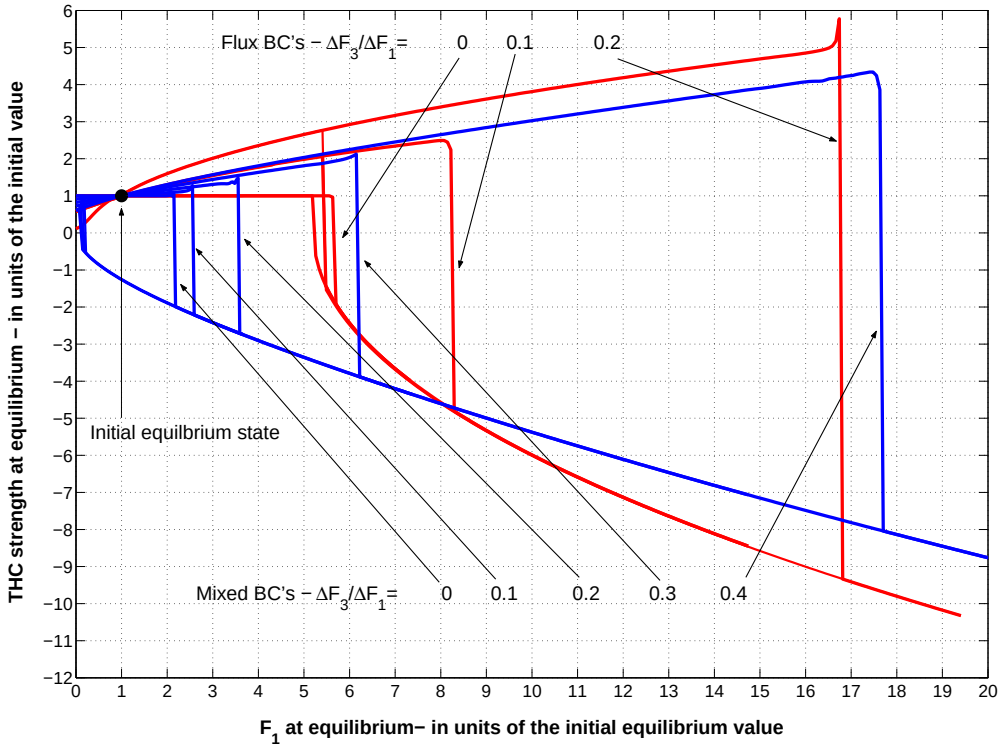

Figure 8: Bifurcation diagram of freshwater flux forcings - Mixed (blue lines) and Flux (red lines) BC's uncoupled models

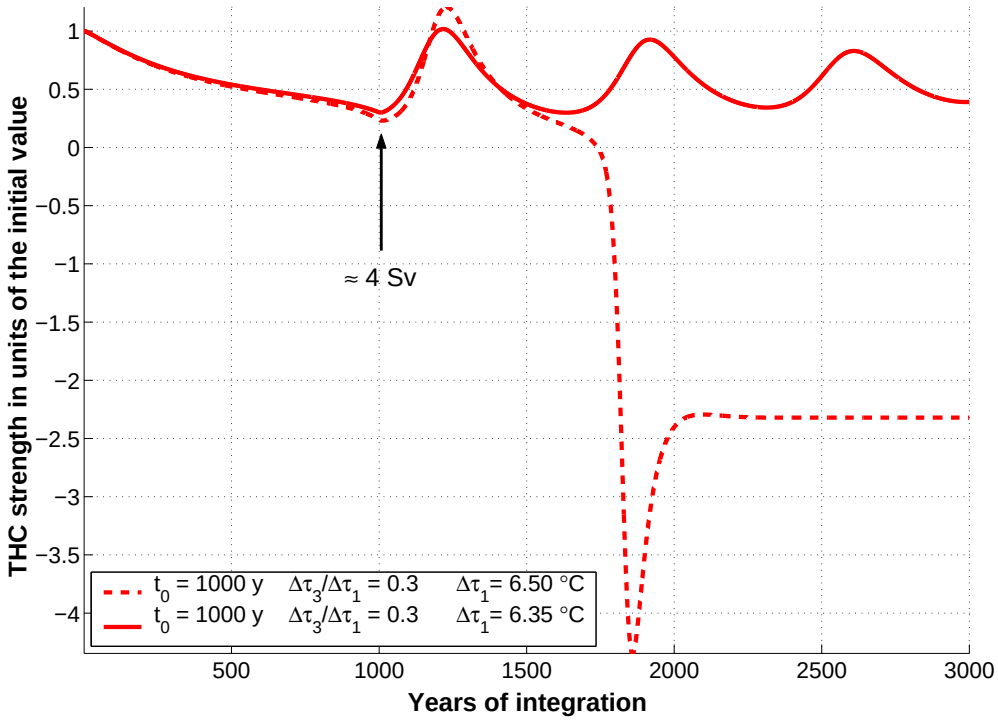

Figure 9: Evolution of the THC strength under a super- and sub-critical forcing to target temperatures

is present only in the flux BC model, therefore in this model for large t_0 the quantity ΔF_1 diverges, while in the mixed BC model it is finite.

These results agree with what can be concluded from the comparison of the bifurcation diagrams presented in figure 8. It is apparent that the mixed BC model is less stable with respect to freshwater flux perturbations. Excluding the temperature feedback mechanism, no bifurcations are present for $\Delta F_3/\Delta F_1 = [0.3, 0.4]$, while for each of the cases $\Delta F_3/\Delta F_1 = [0, 0.1, 0.2]$ the bifurcation point is much farther than in the mixed BC model from the point $(1, 1)$, which describes the unperturbed system.

6 Heat flux forcing

We explore the stability of the northern sinking equilibrium against perturbations to the initial heat fluxes expressed as changes in the target temperatures in the three boxes. We do not impose a constraint fixing the average target temperature τ_M, so that in this case not two but three parameters τ_1, τ_2, τ_3 can be changed. It is possible to recast the system equations 1−6 [Scott et al. 1999] so that the heat forcing is described completely by the changes in $\tau_N = \tau_1 - \tau_2$ and in $\tau_S = \tau_3 - \tau_2$. Therefore we set $\tau_2 = \tau_2(0)$ at all times, so

that $\tau_1^t = \tau_N^t$ and $\tau_3^t = \tau_S^t$. We apply the following forcings:

$$\tau_{i=1,3}(t) \quad = \quad \begin{cases} \tau_{i=1,3}(0) + \tau_{i=1,3}^t \cdot t, & 0 \leq t \leq t_0 \\ \tau_{i=1,3}(0) + \tau_{i=1,3}^t \cdot t_0 & t > t_0 \end{cases} \tag{17}$$

We characterize each forcing experiment, which leads to a final state that is characterized by a northern or a southern sinking equilibrium, using the parameter set 1 defined as $[t_0, \Delta\tau_3/\Delta\tau_1, \Delta\tau_1]$ and the parameter set 2 defined as $[t_0, \Delta\tau_3/\Delta\tau_1, \tau_1^t]$; consistently with the previous case we define $\Delta\tau_i \equiv \tau_i^t \cdot t_0$, that is the total change of the target temperature of the box i. In order to imitate global warming scenarios where polar amplification occurs, we consider $\Delta\tau_i \geq 0$, $i = 1, 3$. We observe that if $\Delta\tau_3 > \Delta\tau_1$, no destabilization of the THC can occur, because such a forcing reinforces the THC. The specified forcings trigger the previously described feedbacks of the system in the following way:

- τ_1's increase larger than τ_3's $\Rightarrow$ H_1's increase larger than H_3's $\Rightarrow$ T_1 increases more than T_3 $\Rightarrow$ q decreases $\Rightarrow$ feedbacks in Figure (2).

We present in Figure 9 two trajectories of q starting from the initial equilibrium state, characterized by slightly different choices of the forcing applied to τ_1 and τ_3. One of the forcing is supercritical, the other one is subcritical. The two trajectories are barely distinguishable

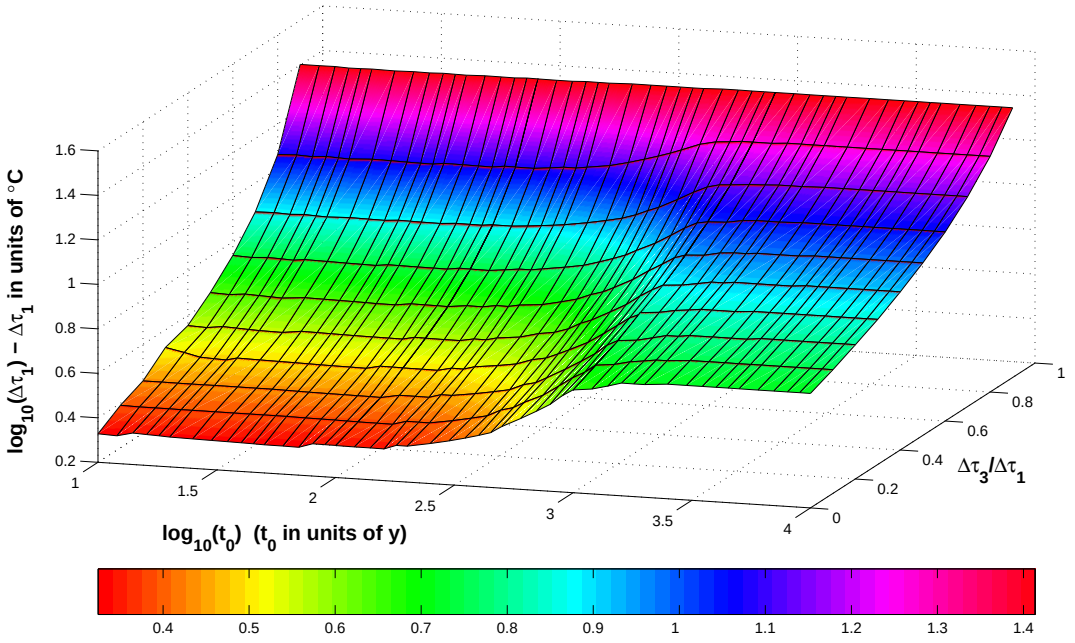

Figure 10: Critical values of the total increase of the target temperatures

up to the end of the perturbation, because the forcing dominates the internal feedbacks of

the system thanks to the large value of λ (although we notice that after around 400 years the decrease of q slows down); the mean flow feedback dampens the increase in T_1 and keeps the system in the northern sinking state, although the decreasing q lowers the amount of warm, salty tropical water injected into box 1; when q gets too small (in this case about 4 Sv) the mean flow feedback becomes negligible and the system eventually collapses. The period of the oscillations is comparable with the time scale of the flushing of the oceanic boxes.

6.1 Critical Forcings

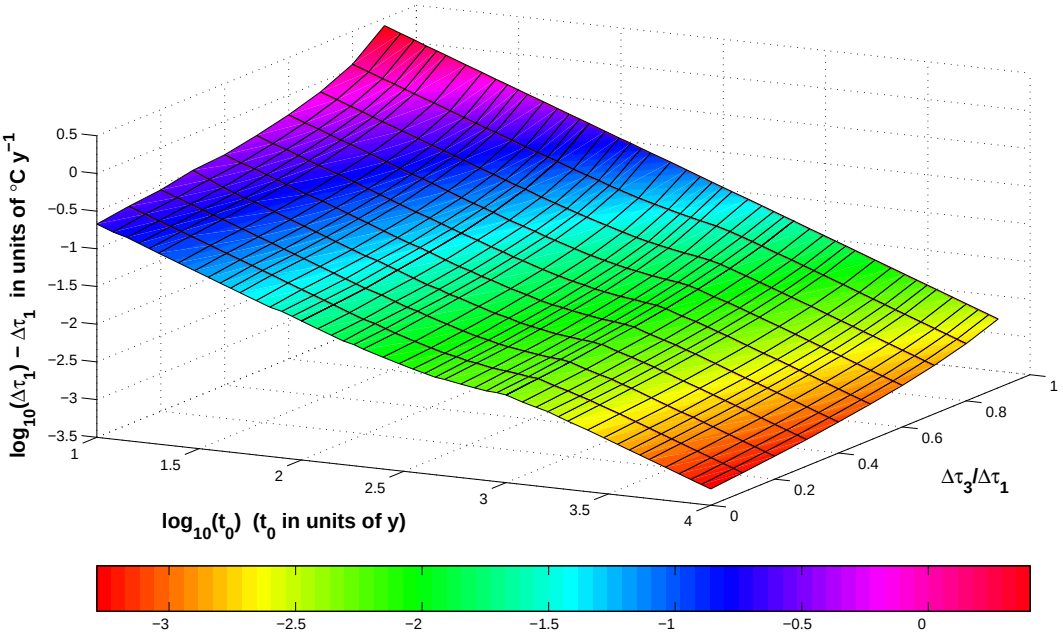

Figure 11: Critical values of the rate of increase of the target temperatures

Figures 10 and 11 present, using respectively the parameter sets previously designated with 1 and 2, the manifold of the critical forcings. In figure 10 we can observe that for all the values of $\Delta\tau_3/\Delta\tau_1$ there is a steep portion of the manifold (corresponding to a time scale of $\approx 250\ y$) dividing two flat regions describing the *fast* and *slow* regimes. In the fast region, a smaller total increase $\Delta\tau_1$ is required to achieve the reversal of the THC. For $\Delta\tau_3/\Delta\tau_1 = 0.9$, which represents the case of highly symmetrical forcings, the critical $\Delta\tau_1$ essentially does not depend on t_0. The absence of thresholds in τ_1^t in the case of more symmetric forcing (large $\Delta\tau_3/\Delta\tau_1$) is clear in figure 11: no flat regions, like those encountered in figure 5, can be found. The main reason for this property of the system's response to perturbations to the target temperatures, is that because of the large value of λ, which is about one order of magnitude larger than the initial q, the forcing is stronger than any negative feedback.

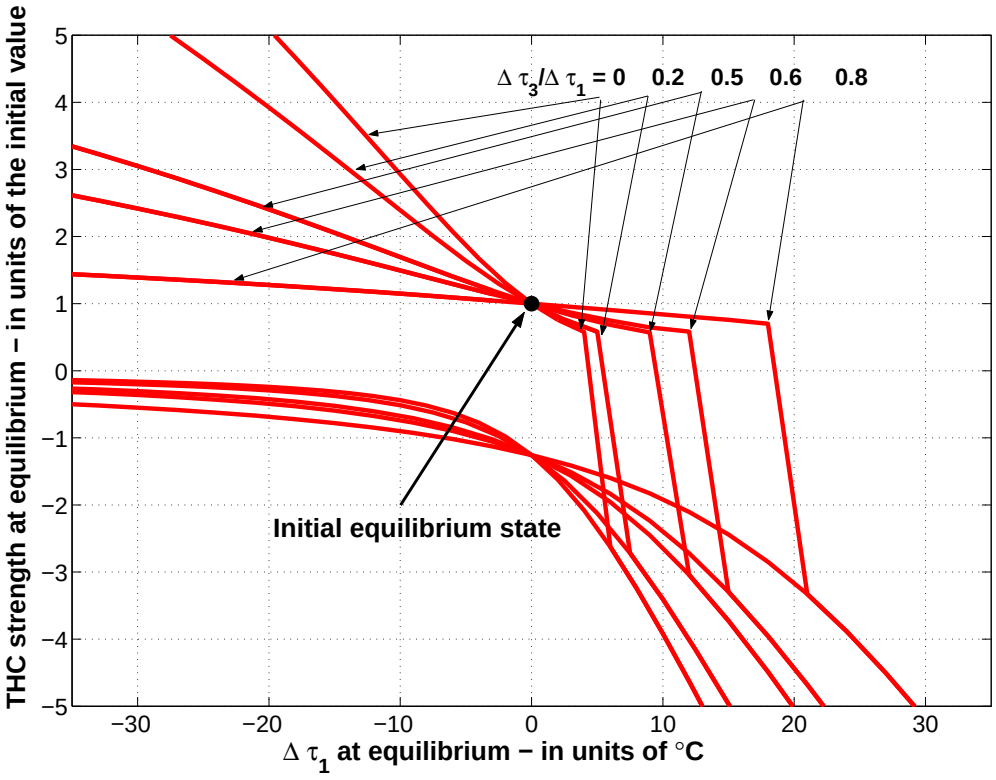

Figure 12: Bifurcation diagram of the target temperature forcings

If the increase in τ_1 is slow enough, the negative feedbacks make the destabilization more difficult, but do not prevent it.

6.2 Bifurcations

For this class of forcings, the achievement of the collapse of the THC is not restricted by any threshold on the rate of increase, therefore we expect the presence of bifurcations for every ratio of $\Delta\tau_3/\Delta\tau_1 < 1$. We present in Figure 12 only the graphs relative to $\Delta\tau_3/\Delta\tau_1 = [0, 0.2, 0.5, 0.6, 0.8]$ for sake of simplicity; the other curves that can be obtained for other values of $0 \leq \Delta\tau_3/\Delta\tau_1 < 1$ are not qualitatively different. The initial state is the point $(0, 1)$. The subcritical Hopf bifurcation points are respectively $(\approx 5, \approx 0.7)$, $(\approx 6, \approx 0.7)(\approx 9, \approx 0.7)(\approx 13, \approx 0.7)$ and $(\approx 17, \approx 0.5)$, which means that the bifurcation occurs when the THC has already declined down to $\approx 10\ Sv$. The bistable region is extremely large in all cases, because the x of the bifurcation points in the lower part of the circuits are below $-50°C$; they become farther and farther form the initial equilibrium value as $\Delta\tau_3/\Delta\tau_1$ increases. This means that once the circulation has reversed, the newly established pattern is extremely stable and can hardly be changed again. To our knowledge, this is the first time such an analysis of bifurcation for this class of forcing is presented in the literature.

7 Conclusions

In this paper we have accomplished a thorough analysis of the stability of the THC as described by the Rooth model.

We simulate the radiative forcing by applying to the equilibrium state perturbations to the moisture and heat fluxes into the three boxes. High rates of increase in the moisture flux into the northern high-latitude box lead to a THC breakdown at smaller final increases than low rates, while the presence of moisture flux increases into the southern high-latitude box strongly inhibit the breakdown. Similarly, fast heat flux increases in the North Hemisphere destabilize the system more effectively than slow ones, and again the enhancement of the heat fluxes in the Southern Hemisphere tend to drive the system towards stability. In particular the relevance of the role of changes in the freshwater flux into the Southern Ocean is determined, along the lines of [Scott et al. 1999] for a similar box model and [Wang et al. 1999a] for an OGCM. In all cases analyzed slow forcings, if asymmetric enough, lead to the reversal of the THC.

We have presented bifurcation diagrams for freshwater fluxes perturbations that reproduce and extend the results previously given in the literature, because the effect of freshwater flux increases in the Southern ocean are included, while for the first time bifurcation diagrams for heat flux forcings, obtained by perturbation in the target temperature, are presented.

We have shown that the restoring boundary conditions for temperature decrease the stability of the system, because they introduce an additional positive feedback that tends to destabilize the system. Flux boundary conditions provide the system with much greater stability with respect to perturbations in the freshwater fluxes at all time scales.

The study of the THC response to perturbations descriptive of global warming scenarios with uncoupled models, of any level of complexity, should take into much greater account the spatial pattern of the freshwater and/or heat flux forcing [Rahmstorf 1995, Rahmstorf 1996]. The THC is a highly nonlinear, nonsymmetric, system, and we suggest that the effect of changing the rate of increase of the forcings should be explored in great detail, to provide a bridge between the instantaneous and quasi-static changes. The THC dynamics encompasses very different time-scales, which can be explored only if the timing of the forcings is varied.

We conclude by pointing out some possible improvements to the present model and possible extensions of this work. In this study we have put emphasis on how the oceanic advection can reverse the symmetry of the pattern of the circulation in the context of a rigorously symmetric geometry. The system is likely to be very sensitive to changes in the volumes of the boxes and especially to the introduction of asymmetries between the two high-latitude boxes, which would also make the system more realistic since the the southern mid-high latitude portion of the Atlantic is considerably larger than the northern mid-high latitude portion. Asymmetries in the oceanic fractional areas would induce asymmetries in

the values of λ_i, thus causing the presence of different restoring times for the various boxes.

Adding a noise component in the tendency equations, along the lines of [Titz et al. 2002a, Titz et al. 2002b], i.e. making the evolution of the system not deterministic, would increase the model's realism and make the model conceptually more satisfying; recent studies have undertaken this strategy and shown that close to instability threshold the evolution of the THC has a very limited predictability [Wang et al. 1999a, Knutti and Stocker 2001], and that stochastic resonance could be responsible for glacial/interglacial climate shift [Ganopolski and Rahmstorf 2002]. It would be interesting to analyze how the intensity and the color of the noise would influence the results obtained and the conclusions drawn in this work. Finally, the model could be improved so that it could be able to represent the main features of the whole conveyor belt, by adding other boxes, descriptive of other oceanic basins, and carefully setting the connections between the boxes, along the lines of some examples presented in [Weaver and Hughes 1992].

In Part II of this paper we perform a similarly structured stability analysis of the THC under global warming conditions using a more complex version of the Rooth model, where an explicit coupling between the oceanic and the atmospheric part is introduced and the main atmospheric physical processes responsible for freshwater and heat fluxes are formulated separately.

Acknowledgments

The authors are grateful to J. Scott for interesting discussions and useful suggestions. One author (V.L.) wishes to thank R. Stouffer for having proposed improvements to an earlier version of the manuscript, and T. Stocker and E. Tziperman for having suggested a number of relevant references. This research was funded in part by the US Department of Energy's (DOE) Climate Change and Prediction Program, in part by the Alliance for Global Sustainability (AGS), and in part by the MIT Joint Program on the Science and Policy of Global Change (JPSPGC). Financial support does not constitute an endorsement by DOE, AGS, or JPSPGC of the views expressed in this article.

References

[Baumgartner and Reichel 1975] Baumgartner A., and E. Reichel, The World Water Balance (Elsevier, New York, 1975)

[Boyle and Keigwin 1987] Boyle E. A., and L. Keigwin, North Atlantic thermohaline circulation during the past 20000 years linked to high-latitude surface temperature, Nature 330, 35-40 (1987)

[Bretherton 1982] Bretherton F. P., Ocean climate modeling, Progr. Oceanogr., 11, 93-129 (1982)

[Broecker et al. 1985] Broecker W. S., D. M. Peteet, and D. Rind, Does the ocean-atmosphere system have more than one stable mode of operation?, Nature, 315, 21-26 (1985)

[Broecker 1994] Broecker, W. S., Massive iceberg discharges as triggers for global climate change, Nature 372, 421-424 (1994)

[Broecker 1997] Broecker W. S., Thermohaline circulation, the Achilles heel of our climate system: Will man-made CO2 upset the current balance? Science, 278, 1582-1588 (1997)

[Bryan 1986] Bryan, F., High-latitude salinity effects and interhemispheric thermohaline circulations, Nature, 323, 301-304 (1986)

[Charles et al. 1994] Charles, C. D., D. Rind, J. Jouzel, R. D. Koster, and R. G. Fairbanks, Glacial/interglacial changes in moisture sources for Greenland: Influences on the ice core record of climate, Science, 263, 508-518 (1994)

[Cubasch et al. 2001] Cubasch, U., G. A. Meehl, G. J. Boer, R. J. Stouffer, M. Dix, A. Noda, C. A. Senior, S. Raper, K. S. Yap, A. Abe-Ouchi, S. Brinkop, M. Claussen, M. Collins, J. Evans, I. Fischer-Bruns, G. Flato, J. C. Fyfe, A. Ganopolski, J. M. Gregory, Z.-Z. Hu, F. Joos, T. Knutson, R. Knutti, C. Landsea, L. Mearns, C. Milly, J. F. B. Mitchell, T. Nozawa, H. Paeth, J. Risnen, R. Sausen, S. Smith, T. Stocker, A.Timmermann, U. Ulbrich, A. Weaver, J. Wegner, P. Whetton, T. Wigley, M. Winton, and F. Zwiers, 9. Projections of future climate change, in Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge University Press, Cambridge, UK) pp. 526-582. (2001)

[Gill 1982] Gill A.E., Atmosphere-Ocean Dynamics (Academic Press, London, 1982)

[Held 1978] Held, I.M., The vertical scale of an unstable baroclinic wave and its importance for eddy heat flux parameterizations, J. Atmos. Sci. 35, 572-576 (1978)

[Ganopolski and Rahmstorf 2001] Ganopolski, A. and S. Rahmstorf, Stability and variability of the thermohaline circulation in the past and future: a study with a coupled model of intermediate complexity, in The Oceans and Rapid Climate Change: Past, Present and Future, edited by D. Seidov, M. Maslin and B. J. Haupt, AGU, Washington, pp. 261-275 (2001)

[Ganopolski and Rahmstorf 2002] Ganopolski, A. and S. Rahmstorf, Abrupt Glacial Climate Changes due to Stohastic Resonance, Phys. Rev. Lett. 88, DOI 038051 (2002)

[Ganopolski et al. 2001] Ganopolski, A., V. Petoukhov, S. Rahmstorf, V. Brovkin, M. Claussen and C. Kubatzki, 2001: CLIMBER-2: A climate system model of intermediate complexity. Part II: Model sensitivity. Clim. Dyn. 17, 735-751 (2001)

[Hughes and Weaver 1994] Hughes T. C. M., and Weaver A. J., Multiple equilibrium of an asymmetric two-basin model, J. Phys. Oceanogr. 24, 619-637 (1994)

[Keigwin et al. 1994] Keigwin L. D., Curry W. B., Lehman S. J., and Johnsen S., The role of the deep ocean in North Atlantic climate change between 70 and 130 ky ago, Nature 371, 323-327 (1994)

[Kitoh et al. 2001] Kitoh A., S. Murakami, and H. Koide, A simulation of the Last Glacial Maximum with a coupled atmosphere/ocean GCM, Geophys. Res. Lett., 28, 2221-2224 (2001)

[Klinger and Marotzke 1999] Klinger B. A., and J. Marotzke, Behavior of double hemisphere thermohaline flow in a single basin, J. Phys. Oceanogr. 29, 382-399 (1999)

[Knutti and Stocker 2001] Knutti, R., and T. F. Stocker, Limited predictability of the future thermohaline circulation close to an instability threshold, J. Climate 15 176-186 (2001)

[Krasovskiy and Stone 1998] Krasovskiy Y. P. and P. H. Stone, Destabilization of the thermohaline circulation by atmospheric transports: An analytic solution, J. Climate 11, 1803-1811 (1998)

[Krinner and Genthon 1998] Krinner, G. and C. Genthon, GCM simulations of the Last Glacial Maximum surface climate of Greenland and Antarctica, Climate Dyn. 14, 741-758 (1998)

[Latif et al. 2000] Latif M., E. Roeckner, U. Mikolajewicz, and R. Voss, Tropical Stabilization of the Thermohaline Circulation in a Greenhouse Warming Simulation, J. Climate 13, 1809-1813 (2000)

[Levitus 1982] Levitus, S., Climatological atlas of the world ocean, NOAA Professional Paper, vol. 13, U.S. Department of Commerce, NOAA, Washington DC (1982)

[Macdonald and Wunsch 1996] Macdonald A.M., and C. Wunsch, An estimate of global ocean circulation and heat fluxes, Nature 382, 436-439 (1996)

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

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

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

[Manabe and Stouffer 1999a] Manabe S. and R. J. Stouffer, Are two modes of thermohalince circulation stable?, Tellus, 51A, 400-411 (1999)

[Manabe and Stouffer 1999b] Manabe S. and R. J. Stouffer, The role of thermohaline circulation in climate, Tellus, 51A-B(1), 91-109 (1999)

[Manabe and Stouffer 2000] Manabe, S. and R. J. Stouffer, Study of abrupt climate change by a coupled ocean-atmosphere model, Quaternary Science Reviews, 19 285-299 (2000)

[Marotzke and Willebrand 1991] Marotzke J. and J. Willebrand, Multiple equilibria of the global thermohaline circulation, J. Phys. Oceanogr. 21, 1372 (1991)

[Marotzke and Stone 1995] J. Marotzke and P. H. Stone, Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. J. Phys. Ocean., 25, 1350-136 (1995) .

[Marotzke 1996] Marotzke J., Analysis of thermohaline feedbacks, Decadal Climate Variability: Dynamics and predicatibility, Anderson D.L.T. and Willebrand J. Eds., Springer-Verlag, Berlin, pp. 333-378 (1996)

[Mikolajewicz and Maier-Reimer 1994] Mikolajewicz, U. and E. Maier-Reimer, Mixed boundary conditions in ocean general circulation models and their influence on the stability of the model's conveyor belt. J. Geophys. Res. 99, 22633-22644 (1994)

[Mikolajewicz and Voss 2000] Mikolajewicz U. and R. Voss, The role of the individual air-sea flux components in CO2-induced changes of the ocean's circulation and climate, Climate Dyn. 16, 627-642 (2000)

[Nakamura et al. 1994] Nakamura M., P. H. Stone, and J. Marotzke, Destabilization of the thermohaline circulation by atmospheric eddy transports, J. Climate 7, 1870-1882 (1994)

[Peixoto and Oort 1992] Peixoto A. and Oort B., Physics of Climate (American Institute of Physics, 1992)

[Petoukhov et al. 2000] Petoukhov, V., A. Ganopolski, V. Brovkin, M. Claussen, A. Eliseev, C. Kubatzki, and S. Rahmstorf, A climate system model of intermediate complexity. Part I: model description and performance for present climate. Climate Dyn., 16, 1-17 (2000)

[Rahmstorf 1995] Rahmstorf S., Bifurcations of the Atlantic Thermohaline circulation in response to changes in the hydrological cycle, Nature 378, 145-149 (1995)

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

[Rahmstorf 1996] Rahmstorf S., On the freshwater forcing and transport of the Atlantic thermohaline circulation, Climate Dyn. 12, 799-811 (1996)

[Rahmstorf 1997] Rahmstorf S., Risk of sea-change in the Atlantic, Nature, 388, 825-826 (1997)

[Rahmstorf 1999a] Rahmstorf S., Shifting seas in the greenhouse?, Nature, 399, 523-524 (1999)

[Rahmstorf 1999b] Rahmstorf S., Rapid Oscillation of the thermohaline ocean circulation,in Reconstructing ocean history: A window into the future, pp 139-149, Abrantes and Mix Eds, Kluwer Academic/Plenum Publishers, New York (1999)

[Rahmstorf 1999c] Rahmstorf S., Decadal Variability of the Thermohaline Ocean Circulation, pp. 309-332, in Beyond El Nio: Decadal and interdecadal climate variability,edited by A. Navarra, Springer, New York, (1999)

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

[Rahmstorf 2000] Rahmstorf, S., The thermohaline ocean circulation - a system with dangerous thresholds? Climatic Change, 46, 247-256 (2000)

[Rahmstorf 2002] Rahmstorf S., Ocean circulation and climate during the past 120,000 years, Nature 419, 207-214 (2002)

[Roemmich and Wunsch 1985] Roemmich D.H., and Wunsch C., Two transatlantic sections: Meridional circulation and heat flux in the subtropical North Atlantic Ocean, Deep Sea Res., 32, 619-664, (1985)

[Rooth 1982] Rooth C., Hydrology and ocean circulation, Progr. Oceanogr., 11, 131-149 (1982)

[Schmittner and Stocker 1999] Schmittner, A., and T. F. Stocker, The stability of the thermohaline circulation in global warming experiments, J. Climate 12, 1117-1133 (1999)

[Scott et al. 1999] Scott J. R. , J. Marotzke, and P. H. Stone, 1999: Interhemispheric thermohaline circulation in a coupled box model, J. Phys. Oceanogr. 29, 351-365 (1999).

[Stocker and Wright 1991] Stocker T. F., and D.G. Wright, Rapid transitions of the ocean's deep circulation induced by changes in the surface water fluxes, Nature 351, 729-732 (1991)

[Stocker and Schmittner 1997] Stocker T. F. and A. Schmittner, Influence of CO_2 emission rates on the stability of the thermohaline circulation, Nature, 388, 862-865 (1997)

[Stocker 2000] Stocker T. F., Past and future reorganisations in the climate system. Quat. Sci. Rev. 19, 301-319 (2000)

[Stocker 2001] Stocker T. F., The Role of Simple Models in Understanding Climate Change In: Continuum Mechanics and Applications in Geophysics and the Environment. B. Straugham, R. Greve, H. Ehrentraut, Y. Wang (eds.) Springer Verlag, pp. 337-367 (2001)

[Stocker 2002] Stocker T. F., North-South Connections, Science 297, 1814-1815 (2002)

[Stommel 1961] Stommel H., Thermohaline convection with two stable regimes of flow, Tellus, 13, 224-230 (1961)

[Stone and Krasovskiy 1999] Stone P. H. and Y. P. Krasovskiy, Stability of the interhemispheric thermohaline circulation in a coupled box model. Dyn. Atmos. Oceans, 29, 415-435 (1999).

[Stouffer et al. 1991] Stouffer, R. J., S. Manabe, and K. Bryan, Climatic response to a gradual increase of atmospheric carbon dioxide, in Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations, The Netherlands, Elsevier Science Publishers, pp. 129-136 (1991)

[Stouffer and Manabe 1999] Stouffer, R. J., and S. Manabe, Response of a coupled ocean-atmosphere model to increasing atmospheric carbon dioxide: Sensitivity to the rate of increase. J. Climate 12, 2224-2237 (1999)

[Titz et al. 2002a] Titz, S., T. Kuhlbrodt, S. Rahmstorf, and U. Feudel, On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulation, Tellus A 54, 89-98 (2002)

[Titz et al. 2002b] Titz S. , T. Kuhlbrodt und U. Feudel, Homoclinic bifurcation in an ocean circulation box model, Int. J. Bif. Chaos 12, 869-875 (2002).

[Tziperman et al. 1994] Tziperman E., R. J. Toggweiler, Y. Feliks, and K. Bryan, 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 (1994)

[Tziperman 2000a] Tziperman E., Proximity of the Present-Day Thermohaline Circulation to an Instability Threshold, J. Phys. Oceanogr. 30, 90-104 (2000)

[Tziperman 2000b] Tziperman E., Uncertainties in thermohaline circulation response to greenhouse warming, Geoph. Res. Lett. 27, 3077-30807 (2000)

[Tziperman and Gildor 2002] Tziperman E. and H. Gildor, The Stabilization of the Thermohaline Circulation by the Temperature/Precipitation Feedback, J. of Phys. Ocean. 32, 2704-2714. (2002)

[Vidal et al. 1999] Vidal L. , R. R. Schneider, O. Marchal, T. Bickert, T. F. Stocker, and G. Wefer, Climate Dynamics 15, 909-919 (1999)

[Wang et al. 1999a] Wang X., Stone P.H., and Marotzke J., Global Thermohaline circulation. Part I: Sensitivity to atmospheric moisture transport, J. Climate 12, 71-82 (1999a)

[Wang et al. 1999b] Wang X., Stone P. H. and J. Marotzke, Global Thermohaline circulation. Part II: Sensitivity with interactive atmospheric transport, J. Climate 12, 83-91 (1999b)

[Weaver and Hughes 1992] Weaver A. J., and T. M. C. Hughes, Stability of the thermohaline circulation and its links to Climate, Trends in Physical Oceanography, Research Trends Series, Council of Scientific Research Integration, Trivandrum, India, 1, 15 (1992)

[Wiebe and Weaver 1999] Wiebe E. C. and A. J.Weaver, On the sensitivity of global warming experiments to the parametrisation of sub-grid scale ocean mixing. Climate Dyn. 15, 875-893 (1999) .

[Zhang et al. 1993] Zhang, S., Greatbatch R.J., and Lin C.A., A re-examination of the polar halocline catastrophe and implications for coupled ocean-atmosphere modeling, J. Phys. Oceanogr. 23, 287 (1993)

Quantity	**Symbol**	**Value**
Mass of Box $i = 1, 3$	M	$1.08 \cdot 10^{20}\ Kg$
Box2/Box$i = 1, 3$ mass ratio	V	2
Average water density	ρ_0	$1000\ Kg\ m^{-3}$
Specific heat per unit mass of water	c_p	$4 \cdot 10^3\ J\ °C^{-1} Kg^{-1}$
Average Salinity	S_0	$35\ psu$
Oceanic fractional area	ϵ	1/6
Thermal expansion coefficient	α	$1.5 \cdot 10^{-4}\ °C^{-1}$
Haline expansion coefficient	β	$8 \cdot 10^{-4}\ psu^{-1}$
Hydraulic constant	k	$1.5 \cdot 10^{-6}\ s^{-1}$
Atmospheric temperature restoring coefficient[a]	$\tilde{\lambda}$	$25.8\ Wm^{-2}\ °C^{-1}$

Table 1: **Value of the main model constants**

[a]Value relative to the oceanic surface fraction only

Constant	**Box 1**	**Box 2**	**Box 3**
Temperature	2.9 $°C$	28.4 $°C$	0.3 $°C$
Salinity	34.7 psu	35.6 psu	34.1 psu
Atmospheric Freshwater Flux	0.41 Sv	-0.68 Sv	0.27 Sv
Surface Heat Flux	-1.58 PW	1.74 PW	-0.16 PW
Oceanic Heat Flux	1.58 PW	-1.74 PW	0.16 PW
THC strength	15.5 Sv	15.5 Sv	15.5 Sv
Target Temperature	0 $°C$	30 $°C$	0 $°C$

Table 2: **Value of the fundamental parameters of the system at the initial equilibrium state**

Thermohaline circulation stability: a box model study - Part II: coupled atmosphere-ocean model

Valerio Lucarini
lucarini@mit.edu
and
Peter H. Stone
phstone@mit.edu
Joint Program on the Science and Policy of Global Change, MIT
Cambridge, MA 02138 USA

June 19, 2003

Abstract

A thorough analysis of the stability of a coupled version of an inter-hemispheric 3-box model of Thermohaline Circulation (THC) is presented. This study follows a similarly structured analysis on an uncoupled version of the same model presented in Part I of this paper. The model consists of a northern high latitudes box, a tropical box, and a southern high latitudes box, which respectively can be thought as corresponding to the northern, tropical and southern Atlantic ocean. We study how the strength of THC changes when the system undergoes forcings representing global warming conditions.

Since we are dealing with a coupled model, a direct representation of the radiative forcing is possible, because the main atmospheric physical processes responsible for freshwater and heat fluxes are formulated separately. Each perturbation to the initial equilibrium is characterized by the total radiative forcing realized, by the rate of increase, and by the North-South asymmetry. Although only weakly asymmetric or symmetric radiative forcings are representative of physically reasonable conditions, we consider general asymmetric forcings, in order to get a more complete picture of the mathematical properties of the system. The choice of suitably defined metrics allows us to determine the boundary dividing the set of radiative forcing scenarios that lead the system to equilibria characterized by a THC pattern similar to the present one, from those that drive the system to equilibria where the THC is reversed. We also

consider different choices for the atmospheric transport parametrizations and for the ratio between the high latitude to tropical radiative forcing. We generally find that fast forcings are more effective than slow forcings in disrupting the present THC patterns, forcings that are stronger in the northern box are also more effective in destabilizing the system, and that very slow forcings do not destabilize the system whatever their asymmetry, unless the radiative forcings are very asymmetric and the atmospheric transport is a relatively weak function of the meridional temperature gradient; in this latter case we present the analysis of the bifurcations of the system. The changes in the strength of the THC are primarily forced by changes in the latent heat transport in the hemisphere, because of its sensitivity to temperature that arises from the Clausius-Clapeyron relation.

1 Introduction

One of the main issues in the study of climate change is the fate of the Thermohaline Circulation (THC) in the context of global warming [Weaver and Hughes 1992, Manabe and Stouffer 1993, Stocker and Schmittner 1997, Rahmstorf 1997, Rahmstorf 1999a, Rahmstorf 1999b, Rahmstorf 2000, Wang et al. 1999a, Wang et al. 1999b]. The THC plays a crucial role in determining the main features of the North Atlantic climate, since the advection of warm water prevents the formation of sea ice up to very high latitudes even during winter, and the heat exchange between relatively warm water and cold air warms high latitudes [Levitus 1982, Broecker 1994, Stocker 2000]. The THC plays a major role in the global circulation of the oceans as pictured by the conveyor belt scheme [Weaver and Hughes 1992, Stocker 2001]; therefore the effect on the climate of a change of its pattern may be global [Broecker 1997, Manabe and Stouffer 1999b, Cubasch et al. 2001].

The THC is sensitive to changes in the climate since the NADW formation is sensitive to variations in air temperature and in precipitation in the Atlantic basin [Rahmstorf and Willebrand 1995, Rahmstorf 1996]. There are several paleoclimatic data sets indicating that changes in the patterns or collapses of the THC coincide with large variations in climate, especially in the North Atlantic region [Broecker et al. 1985, Boyle and Keigwin 1987, Keigwin et al. 1994, Rahmstorf 2002].

Box models have historically played a major role in the understanding of the fundamental dynamics of the THC [Weaver and Hughes 1992]: the Stommel two-box oceanic model [Stommel 1961] built the first conceptual bridge between the THC strength and density gradients. The oceanic model proposed by Rooth [Rooth 1982] introduced the idea that the driver of the THC was density difference between the high latitude basins of the northern and southern hemispheres, thus implying that the THC is an inter-hemispheric phenomenon. Both the Stommel and Rooth models allow two stable equilibria, one -the warm mode- characterized by downwelling of water in the North Atlantic, which resembles the present oceanic circulation, and the other - the cold mode - characterized by upwelling of water in the North Atlantic. Perturbations to the driving parameters of the system, i.e. freshwater and heat fluxes in the oceanic boxes, can cause transitions between the two regimes.

More recently various GCM experiments have shown that multiple equilibria of the THC are possible [Bryan 1986, Manabe and Stouffer 1988, Stouffer et al. 1991, Stocker and Wright 1991, Manabe and Stouffer 1999a, Marotzke and Willebrand 1991, Hughes and Weaver 1994], and most GCMs have shown that the GHGs radiative forcing could cause a weakening or eventually a shutdown of the THC by the inhibition of the sinking of the water in the Northern Atlantic. Large increases of the moisture flux and/or of the sea surface temperature in the downwelling regions are the driving mechanisms of such a process. Nevertheless in a recent global warming simulation run with the *ECHAM-4* model the THC remained almost

constant since the destabilizing mechanisms have been offset by the increase of the salinity advected in the downwelling region due to net freshwater export from the whole North Atlantic [Latif et al. 2000]. The weakening of the THC could greatly limit the regional warming in the North Atlantic [Rahmstorf 1997, Rahmstorf 1999a, Rahmstorf 1999b, Rahmstorf 2000].

Analysis of GCM data have also indicated that the THC is an inter-hemispheric phenomenon, and in some cases it has been found that the THC strength is approximatively proportional to the density difference between the Northern and Southern Atlantic [Hughes and Weaver 1994, Rahmstorf 1996, Klinger and Marotzke 1999, Wang et al. 1999a], thus supporting Rooth's approach. While hemispheric box models like Stommel's have been extensively studied [Weaver and Hughes 1992, Nakamura et al. 1994, Marotzke 1996, Krasovskiy and Stone 1998, Tziperman and Gildor 2002], and the stability of their THC has been assessed in the context of various levels of complexity in the representation of the coupling between the ocean and the atmosphere, the literature on inter-hemispheric box models is far less abundant [Rahmstorf 1996, Scott et al. 1999, Stone and Krasovskiy 1999, Titz et al. 2002a, Titz et al. 2002b], and relatively few box-model studies have included the effect of coupling an inter-hemispheric model of the ocean to the atmosphere [Scott et al. 1999, Stone and Krasovskiy 1999, Tziperman and Gildor 2002].

In this paper we perform a parametric study of the relevance of both spatial *and* temporal patterns of the radiative forcing and characterize the response of the system by determining the thresholds beyond which we have destabilization of the present mode of the THC and transition to an reversed circulation. This analysis follows the study reported in Part I [**?**] of this paper, where we have dealt with an uncoupled oceanic model. The model here analyzed is characterized by the presence of explicit coupling between the ocean and the atmosphere: the atmospheric freshwater and heat fluxes are expressed as functions of the oceanic temperatures. In particular this model includes the Clausius-Clapeyron effect on moisture and latent heat fluxes in a parameter study of how global warming affects the THC. In the case of GCMs, where obviously the Clausius-Clapeyron relations are included in the physical description of the atmosphere, parameter studies cannot be performed, while this strongly nonlinear effect has been left out of several interhemispheric box models and Earth Models of Intermediate Complexity (EMICS) [Stone and Krasovskiy 1999, Wiebe and Weaver 1999, Petoukhov et al. 2000]

The treatment of a wide range of temporal scales for the forcings allows us to join on naturally and continuously the analysis of quasi-static perturbations, which have been usually addressed with the study of the bifurcations of the system [Rahmstorf 1995, Rahmstorf 1996, Stone and Krasovskiy 1999,Scott et al. 1999,Titz et al. 2002a,Titz et al. 2002b], with the study of the effects of very rapid perturbations [Rahmstorf 1996,Scott et al. 1999,Wiebe and Weaver 1999], which are usually framed in different ways and contexts.

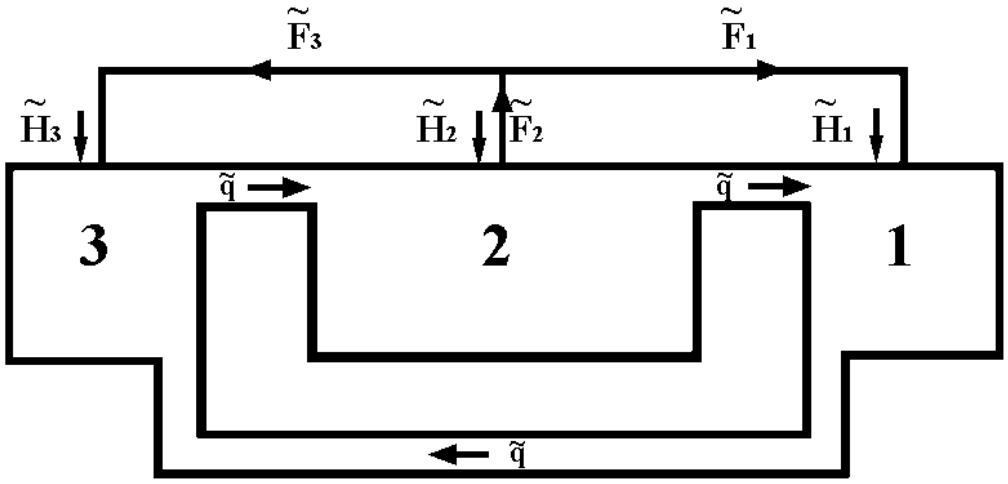

Figure 1: Schematic picture of the interhemispheric box model

Only relatively few studies have explicitly addressed how the spatial [Rahmstorf 1995, Rahmstorf 1996, Rahmstorf 1997] or temporal [Stocker and Schmittner 1997] patterns of forcing determine the response and allow the investigation of the stability properties of the system in the context of more complex models. Obvious limitations due to computing time did not allow an extensive exploration of the parameter space of the forcings applied, in terms of total change realized, rate of change, and spatial pattern of the perturbation.

Our paper is organized as follows. In the second section we provide a brief description and the general mathematical formulation of the dynamics of the three-box model used in this study. In the third section we describe the parametrization of the coupling between the atmosphere and the ocean. In the fourth section we analyze the feedbacks of the system and describe the forcings we apply. In the fifth section we analyze some relevant model runs. In the sixth section we treat the stability properties of the system. In the seventh section we perform a parametric sensitivity study. In the eight section we present a study of the bifurcations of the system. In the nineth section we present our conclusions.

2 Model description

The three-box model consists of a northern high latitude box (box 1), a tropical box (box 2), and a southern high latitude box (box 3). The volume of the two high latitudes boxes is the same, and is $1/V$ times the volume of the tropical box. We choose $V = 2$, so that box 1, box 2 and box 3 respectively can be thought as describing the portions of an ocean like the Atlantic north of $30°N$, between $30°N$ and $30°S$, and south of $30°S$. At the eastern and western boundaries of the oceanic boxes there is land; our oceanic system spans $60°$ in longitude, so that it covers $\epsilon = 1/6$ of the total surface of the planet. The boxes are

5000 m deep, so that the total mass $M_{i=1,3} = M$ of the water contained in each of the high-latitude boxes is $\approx 1.1 \cdot 10^{20}\ Kg$, while $M_2 = V \cdot M_{i=1,3} = 2 \cdot M$. The tropical box is connected to both high latitude boxes; moreover the two high latitude boxes are connected by a deep current passage (which bypasses the tropical box) containing a negligible mass. The three boxes are assumed to be well mixed, so that the *polar halocline disaster* [Bryan 1986, Zhang et al. 1993] is automatically excluded from the range of phenomena that can be described by this model.

The physical state of the box i is described by its temperature T_i and its salinity S_i. The box i receives from the atmosphere the net flux of heat $\tilde{H}_i$ and the net flux of freshwater $\tilde{F}_i$; the freshwater fluxes $\tilde{F}_i$ globally sum up to 0, since the freshwater is a globally conserved quantity. The box i is subjected to the oceanic advection of heat and salt from the upstream box through the THC, whose strength is $\tilde{q}$. The atmosphere has a negligible heat capacity and water content compared to the ocean, and its only function is to serve to transport heat and moisture; for time scales longer than a few months this is a reasonable approximation [Gill 1982]. The land is also assumed to have negligible heat capacity and water content. The oceanic box $i = 1, 3$ receive a fraction $1/\gamma_i$ of the total net moisture transported by the atmosphere from the tropics to high latitudes in the northern ($i = 1$) and southern hemisphere ($i = 3$) respectively; this includes the fraction that directly precipitates over the oceanic boxes $i = 1, 3$ and the fraction that precipitates over land and runs off to the oceanic box $i = 1, 3$.

The fractional catchment area $1/\gamma_i$ can range in our system from 1 (all of the atmospheric moisture exported from the tropics to the high latitudes ends up respectively in the box $i = 1, 3$) to 1/6 (the box$i = 1, 3$ receives only the moisture transported from the tropics that precipitates on the ocean surface). The remaining fraction $(1 - 1/\gamma_i)$ of the total atmospheric moisture exported from the tropics returns back to the oceanic box 2 by river runoff or underground flow. This latter fraction does not effect the moisture budget of the oceanic boxes $i = 1, 3$ but does effect their heat budget, since the process of condensation occurs over the high latitude regions $i = 1, 3$, so that latent heat is released to the atmosphere and is immediately transferred to the oceanic box $i = 1, 3$. Estimates for γ_i for the Atlantic range from 1.5 to 3 [Marotzke 1996]. We set $\gamma_1 = \gamma_3$ in order to keep the geometry of the problem entirely symmetric, and choose $\gamma_1 = \gamma_3 = 2$.

In figure 1 we present a scheme of our system in the northern sinking pattern: note that the arrows representative of the freshwater fluxes are arranged in such a way that the conservation law is automatically included in the graph. The dynamics of the system is described by the tendency equations for the heat and the salinity of each box. We divide the three heat tendency equations by $c_p \cdot M_i$, where c_p is the constant pressure specific heat of water per unit of mass, and in the salinity tendency equations we neglect the contribution

of the freshwater fluxes in the mass balance, so that virtual salinity fluxes and freshwater fluxes are equivalent [Marotzke 1996, Rahmstorf 1996, Scott et al. 1999]. We obtain the following final form for the temperature and salinity tendency equations for the three boxes [Scott et al. 1999]:

$$\dot{T}_1 = \begin{cases} q(T_2 - T_1) + H_1, & q \geq 0 \\ q(T_1 - T_3) + H_1, & q < 0 \end{cases} \tag{1}$$

$$\dot{T}_2 = \begin{cases} \frac{q}{V}(T_3 - T_2) + H_2, & q \geq 0 \\ \frac{q}{V}(T_2 - T_1) + H_2, & q < 0 \end{cases} \tag{2}$$

$$\dot{T}_3 = \begin{cases} q(T_1 - T_3) + H_3, & q \geq 0 \\ q(T_3 - T_2) + H_3, & q < 0 \end{cases} \tag{3}$$

$$\dot{S}_1 = \begin{cases} q(S_2 - S_1) - F_1, & q \geq 0 \\ q(S_1 - S_3) - F_1, & q < 0 \end{cases} \tag{4}$$

$$\dot{S}_2 = \begin{cases} \frac{q}{V}(S_3 - S_2) - F_2, & q \geq 0 \\ \frac{q}{V}(S_2 - S_1) - F_2, & q < 0 \end{cases} \tag{5}$$

$$\dot{S}_3 = \begin{cases} q(S_1 - S_3) - F_3, & q \geq 0 \\ q(S_3 - S_2) - F_3, & q < 0 \end{cases} \tag{6}$$

where $q = \rho_0 \cdot \tilde{q}/M$, $H_i = \tilde{H}_i/(c_p \cdot M)$, and $F_i = \rho_0 \cdot S_0 \cdot \tilde{F}_i/M$.

We impose that the average salinity is a conserved quantity; this means that $\dot{S}_1 + V\dot{S}_2 + \dot{S}_3 = 0$, which implies that:

$$F_2 = -\frac{1}{V}(F_1 + F_3). \tag{7}$$

This conservation law holds at every time, and so rules out the possibility of including in our study the effects of the melting of continental ice sheets.

On the other hand we do not impose any strict conservation law for the transient behavior of the global heat budget of the ocean, since we essentially want to include the effects of radiative imbalances. We note that the system can (asymptotically) reach an equilibrium

only if its feedbacks can drive it towards a state in which the following equation holds:

$$H_2 = -\frac{1}{V}(H_1 + H_3); \tag{8}$$

The strength of the specific THC, as in the Stommel model, is parametrized as proportional to the difference between the density of the water contained in the box 1 and the density of the water contained in the box 3. Given that the water density can approximately be expressed as:

$$\rho(T, S) \approx \rho_0(1 - \alpha T + \beta S); \tag{9}$$

where α and β are respectively the thermal and haline expansion coefficients, set respectively to $1.5 \cdot 10^{-4}\ ^\circ C^{-1}$ and $8 \cdot 10^{-4}\ psu^{-1}$, and ρ_0 is a standard reference density, we obtain for the normalized THC strength q the following relation:

$$q = \frac{k}{\rho_0}(\rho_1 - \rho_3) = k(\alpha(T_3 - T_1) + \beta(S_1 - S_3)), \tag{10}$$

where k is the hydraulic constant,chosen to obtain a reasonable value of the THC strength. The northern (southern) sinking state of the circulation is therefore characterized by $q > (<)0$.

Considering the case $q > 0$, we obtain a diagnostic relation for steady state q by setting to 0 equations 3 and 6, and using equation 9:

$$q_{eq} = (k(\alpha((H_3)_{eq} + \beta(F_3)_{eq})))^{1/2}; \tag{11}$$

in the case $q < 0$, a diagnostic relation for the steady state q can be obtained using the same procedure as above but setting to zero equations 1 and 4:

$$q_{eq} = -(k((\alpha(H_1)_{eq} + \beta(F_1)_{eq}))^{1/2}. \tag{12}$$

The *equilibrium* value of the THC strength is determined by the equilibrium values of the heat and freshwater fluxes of the box where we have upwelling of water. These results generalize the expressions given by Rahmstorf [Rahmstorf 1996].

The transient evolution of q is determined by its tendency equation:

$$\begin{aligned} \dot{q} \quad &= -q^2 + kq(\alpha(T_1 - T_2) - \beta(S_1 - S_2)) + \\ &\quad +k(\alpha(H_3 - H_1) + \beta(F_3 - F_1)) \qquad q \geq 0, \end{aligned} \tag{13}$$

$$\begin{aligned}\dot{q} &= q^2 + kq(\alpha(T_3 - T_2) - \beta(S_3 - S_2)) + \\ &\quad + k(\alpha(H_1 - H_3) + \beta(F_1 - F_3)) \qquad q < 0,\end{aligned} \tag{14}$$

We observe that the difference between the forcings applied to the freshwater and heat fluxes into the two boxes 1 and 3 directly effect the evolution of q; the presence of terms involving the gradient of temperature and salinity between box 2 and box 1 (box 3) if $q > 0$ ($q < 0$) breaks the symmetry of the role played by the two high latitude boxes. In our experiments we perturb an initial equilibrium state -which is the same for all the experiments in all of the three versions of the model analyzed- by changing at various rates the parameters controlling the fluxes H_i and/or F_i and observe under which conditions we obtain a reversal of the THC. The reversal of the THC causes a sudden cooling and freshening in the box 1 and a sudden warming and increase of the salinity in the box 3, because the former loses and the latter receives the advection of the tropical warm and salty water.

3 Parametrization of the coupling between the atmosphere and the ocean

This model closely follows the Scott et al. [Scott et al. 1999] model except for the addition of the effect of the Clausius-Clapeyron equations on moisture and latent heat transport; it differs greatly from the model presented in the Part I paper since it incorporates explicit coupling between the ocean and the atmosphere. The coupled model is such that the main atmospheric physical processes responsible for freshwater and heat fluxes have distinct formulations, therefore the simulation of global warming scenarios can be obtained more realistically by forcing the radiative flux term alone.

The rescaled atmospheric fluxes of heat and freshwater in $1-6$ are parameterized as functions of the box temperatures T_i. We choose simple but physically plausible functional forms which are based on the large scale processes governing the transfer of heat and freshwater through the atmosphere. We want to capture the dependence of atmospheric transport from the tropics to the high-latitudes on the temperature gradient, considering that baroclinic eddies contribute most of the meridional transport around 30 $^\circ N$ [Peixoto and Oort 1992], and the dependence of the moisture content of the atmosphere on the average temperature, as well as the dependence of the outgoing long wave radiation on the temperature. The net freshwater fluxes F_i are then parameterized as [Stone and Miller 1980, Stone and Yao 1990]:

$$F_1 = \frac{C_1}{(\frac{T_2+T_1}{2})^3} e^{-\frac{L_v}{R_v \frac{T_2+T_1}{2}}} (T_2 - T_1)^n, \tag{15}$$

$$F_3 = \frac{C_3}{(\frac{T_2+T_3}{2})^3} e^{-\frac{L_v}{R_v \frac{T_2+T_3}{2}}} (T_2 - T_3)^n, \tag{16}$$

where L_v is the unit mass latent heat of vaporization of water (taken as constant), R_v is the gas constant, and C_1 and C_3 are coefficients we have to calibrate. The exponential functions are derived from the Clausius-Clapeyron law, while the value of the exponent n determines the strength of the temperature gradient feedback; F_2 is obtained using equation 7.
The surface heat fluxes H_i are decomposed in three components describing physically different phenomena:

$$H_i = SH_i + LH_i + RH_i \tag{17}$$

where SH_i and LH_i are respectively the convergence of the atmospheric flux of sensible and latent heat in the box i, while RH_i describes the radiative balance between incoming solar radiation and outgoing longwave radiation. The convergence of atmospheric transports must globally sum up to zero at any time, since the atmosphere is closed. The convergence of the latent heat fluxes LH_i is respectively proportional to the convergence of the freshwater fluxes F_i:

$$\begin{aligned} LH_i &= \gamma_i \frac{L_v}{c_p \cdot S_0 \cdot \rho_0} F_i, \quad i = 1,3 \\ LH_2 &= -\frac{1}{V}(LH_1 + LH_2) \end{aligned} \tag{18}$$

where the constant relating the rescaled freshwater and the latent heat flux has been obtained considering that for the physical fluxes are related by $L\tilde{H}_i = \gamma_i \cdot L_v \cdot \tilde{F}_i$.

We note that the Clausius-Clapeyron equations had been included in the description of atmosphere-ocean coupling in earlier studies, but those dealt with hemispheric and not inter-hemispheric models [Nakamura et al. 1994, Tziperman and Gildor 2002]. We observe that in density units the following relation holds:

$$(\frac{LH_i}{F_i})_\rho = \frac{\alpha \cdot \gamma_i \cdot L_v}{\beta \cdot c_p \cdot S_0} \approx 6, \quad i = 1,3; \tag{19}$$

We underline that this value of the ratio depends on our assumed values for γ_i, α and β. For a more realistic equation of state for the density, the ratio could be smaller but is still larger than 1 for a very large temperature range.

The convergence of the sensible heat fluxes SH_i are parameterized as a constant times the n^{th} power of meridional temperature gradient [Stone and Miller 1980, Stone and Yao 1990] :

$$\begin{aligned} SH_i &= D_i(T_2 - T_i)^n, \quad i = 1,3 \\ SH_2 &= -\frac{1}{V}(SH_1 + SH_2) \end{aligned} \tag{20}$$

Finally, the radiative part RH_i is parameterized as in [Wang and Stone 1980, Marotzke and

Stone 1995, Marotzke 1996]:

$$RH_i = A_i - B_i T_i = B_i(\frac{A_i}{B_i} - T_i) = B_i(\vartheta_i - T_i), \quad i = 1, 2, 3. \tag{21}$$

With respect to the box i, A_i describes the net radiative budget if $T_i = 0\ ^\circ C$, B_i is an empirical coefficient, which, if albedo is fixed, as in this model, is a measure of the sensitivity of the thermal emissions to space to surface temperature, including also the water vapor and clouds feedbacks, and $\vartheta_i = A_i/B_i$ is the radiative equilibrium temperature.

In order to obtain for the coupled model here presented an equilibrium control solution identical to that of the uncoupled model presented in Part I, we need to carefully choose the constants in the atmospheric parametrization.

3.1 Choice of the constants

Using the parametrization developed by Marotzke [Marotzke 1996], we set for our model $B_1 = B_2 = B_3 = B = 5.1 \cdot 10^{-10}\ s^{-1}$. We can then obtain the the following restoring equation for the global average temperature T_M:

$$\dot{T}_M = B(\vartheta_M - T_M); \tag{22}$$

where the parameter B introduces a time scale of $\approx$ 60 years for the radiative processes [Marotzke 1996, Scott et al. 1999]. The ϑ_i are chosen following the parametrization presented in [Marotzke 1996] and are such that the average global radiative temperature $\vartheta_M \equiv (\vartheta_1 + V \cdot \vartheta_2 + \vartheta_3)/(2+V) = 15^\circ C$, which is also the value of the global average temperature at equilibrium, as in the uncoupled model in Part I. In physical terms, B corresponds to the property of the system that a global radiative forcing of 1 Wm^{-2} (which results in an effective radiative forcing of 6 Wm^{-2} in the oceanic surface fraction, since we assume that land and atmosphere have no heat capacity) causes an increase of the average temperature of the system T_M of $\approx 0.6\ ^\circ C$ when equilibrium is re-established; this property can be summarized by introducing a climatic temperature/radiation *elasticity* parameter $\kappa_M \approx 0.6\ ^\circ C W^{-1} m^2$. Therefore, considering that it is estimated that in the real Earth system the doubling of CO_2 causes an average radiative forcing of $\approx 4\ Wm^{-2}$, we can loosely interpret the parameter κ_M as indicating a model *climate sensitivity* of $\approx 2.5\ ^\circ C$.

Substituting in 15 and 16 the equilibrium T_i of the control run, we can derive the C_i such that we obtain $F_1 = 13.5 \cdot 10^{-11}\ s^{-1}$ and $F_3 = 9 \cdot 10^{-11}\ s^{-1}$. The latent heat fluxes are then obtained using 18, while the coefficients D_i for the sensible heat fluxes in 20 are derived by requiring that the total heat flux H_i in 17 of the coupled and of the uncoupled model match at the equilibrium solution. The relative magnitude of the latent and sensible

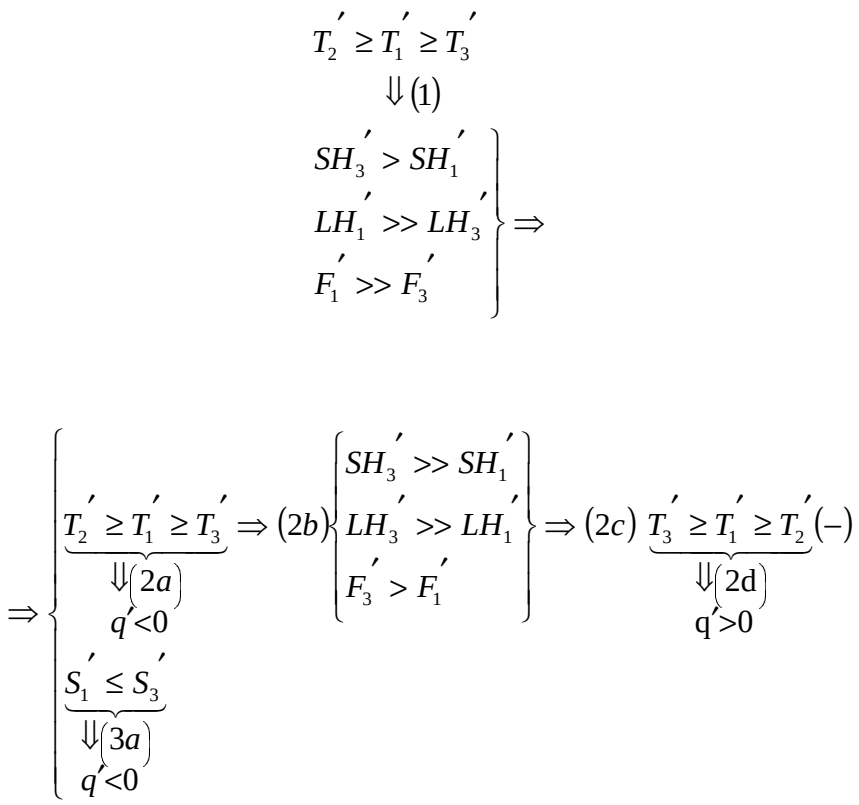

Figure 2: Atmospheric feedbacks scheme of the coupled model

heat fluxes LH_i and SH_i at equilibrium depend on the choice of $\gamma_1 = \gamma_3$ [Marotzke 1996]. We present in table 2 the data relative to the values of the total surface heat flux and of the latent, sensible, and radiative heat components for the control run. While the sensible and latent heat fluxes into box 1 are roughly the same, in the case of box 3 the sensible heat flux is almost three times as large as the latent heat flux; the figures we obtain with our model are broadly in agreement with estimates [Peixoto and Oort 1992].

The parameter n determines the efficiency of the atmospheric transports in terms of its sensitivity to the meridional temperature gradient: we consider for n the values $[1, 3, 5]$, which include the domain proposed in [Held 1978] (2 to 5), but also include simple diffusive representation ($n = 1$).

4 Radiative forcings applied and feedbacks scheme of the system

In order to analyze the stability of the coupled model to radiative forcings, we increase the radiative equilibrium temperatures ϑ_i: such changes can be thought of as a natural representation of global warming scenarios. In the case of a coupled model, given that the

fluxes depend both on averages and gradients of temperatures of neighboring boxes, changes in the parameter ϑ_2, which in the first approximation controls T_2, need to be considered in order to perform a complete and sensible study. As in the previous analysis, we alter the driving parameters ϑ_i by using a linear increase:

$$\vartheta_i(t) \quad = \quad \begin{cases} \vartheta_i(0) + \vartheta_i^t \cdot t, & 0 \leq t \leq t_0 \\ \vartheta_i(0) + \vartheta_i^t \cdot t_0. & t > t_0 \end{cases} \quad i = 1, 2, 3 \tag{23}$$

We make this choice because a linearly increasing radiative forcing approximately corresponds in physical terms to an exponential increase of the concentration of greenhouse gases [Shine et al. 1995, Stocker and Schmittner 1997]. The role of ϑ_2 is analyzed by considering three cases $\Delta\vartheta_2/\Delta\vartheta_1 = [1.0, 1.5, 2.0]$, where we have used the definition $\Delta\vartheta_i = \vartheta_i^t \cdot t_0$. This allows for the fact that in global warming scenarios the net radiative forcing increase is larger in the tropics than in mid-high latitudes [Ramanathan et al. 1979], because of differences in the specific humidity. For a given choice of $\Delta\vartheta_2/Delta\vartheta_1$, each forcing can uniquely identified by the triplet $[t_0, \Delta\vartheta_3/\Delta\vartheta_1, \Delta\vartheta_1]$ and by the triplet $[t_0, \Delta\vartheta_3/\Delta\vartheta_1, \vartheta_1^t]$. Although only weakly asymmetric or symmetric forcings are representative of physically reasonable conditions, we consider general asymmetric forcings, in order to get a more complete picture of the mathematical properties of the system. We find that in most cases the destabilization of the THC occurs only if $\Delta\vartheta_1 \geq \Delta\vartheta_3$, therefore we perform the analysis of the atmospheric feedbacks of the system - which are obviously coupled to the previously described oceanic feedbacks- under this condition. We observe that the first order effect of such a forcing is to weaken the THC, because the following relation holds:

- Increase in $\vartheta_2 \geq$ Increase in $\vartheta_1 \geq$ Increase in $\vartheta_3 \Rightarrow$ Increase in $H_2 \geq$ Increase in $H_1 \geq$ Increase in $H_3 \Rightarrow$ Increase in $T_2 \geq$ Increase in $T_1 \geq$ Increase in $T_3 \Rightarrow q$ decreases.

We underline that a larger radiative forcing in the tropical box may decrease the stability of the circulation because it would cause advection of warmer water from box 2 to box 1.

The relative complexity of the coupled model makes the analysis of the feedbacks much more difficult than in the uncoupled model presented in the Part I paper. In figure 2 we present a scheme describing the main atmospheric feedbacks; the symbol ' indicates that we are dealing with small perturbations from the equilibrium value.

If T_2 increases more than T_1, which itself increases more than T_3, both the absolute value of the meridional temperature gradients and the average temperatures of box 1 and box 2, and of box 2 and box 3 increase, so that all the poleward atmospheric transports increase (step 1). The total effect is that box 1 heats up and freshens more than boxes 2 and 3 (step 2 and 3), so that the oceanic circulation weakens (steps $2a$ and $3a$). The strong increase of T_1 reduces $T_2 - T_1$ and so hinders the efficacy of the transport between these two boxes;

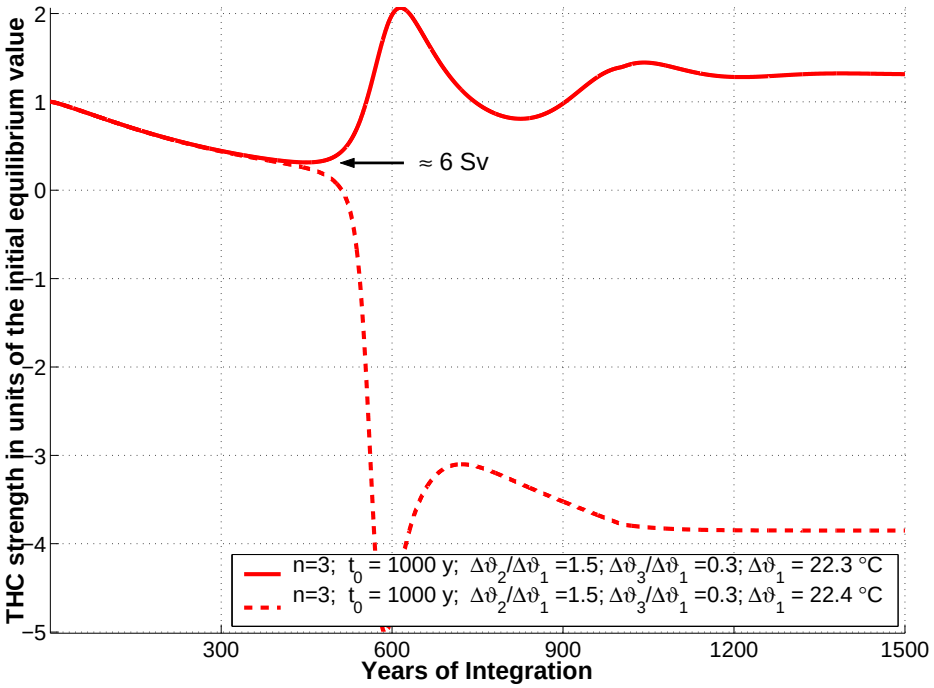

Figure 3: Evolution of the THC strength under a super- and sub-critical radiative flux forcing - Coupled Model

this causes a great reduction of the freshwater flux and of the latent heat (and sensible heat) flux into box 1. On the contrary, for box 2 we obtain that the effect of the decrease of the meridional temperature gradient is relatively small and is compensated by the increase in the capacity of the atmosphere to retain moisture, so that the freshwater and total heat flux from box 2 to box 3 increase (step 2b). This causes a freshening and heating of box 3 relative to box 1 and box 2, so that the oceanic circulation is enhanced (step $2d$); we get temperature variations that are opposite to the initial variations, so that the thermic part of the total atmospheric feedback (step $2c$) is negative. Large values of n make the negative feedback in step $2b$ associated to the decrease of $T_2 - T_1$ more powerful; forcings with large $\Delta\vartheta_2$ can generally decrease the stability of the system because warmer tropical air can transport more freshwater and latent heat to the high-latitudes, and the effect is usually larger for the fluxes into box 1. Small $\Delta\vartheta_3/\Delta\vartheta_1$ will generate in steps $2a$ and $3a$ a very intense weakening of the THC and possibly its reversal. Therefore the whole process seems altogether dominated by the variations in the latent heat and freshwater fluxes.

5 Analysis of selected model runs

In figure 3 we show the time evolution of the THC strength q (we have chosen $n = 3$) for two different radiative forcings lasting 1000 years: the solid line describes the subcritical and

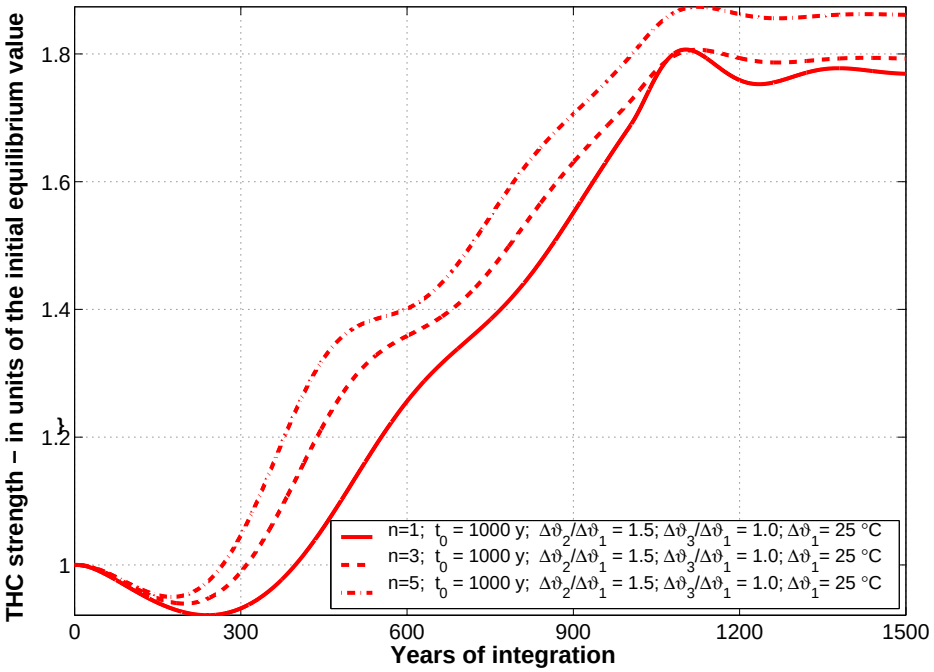

Figure 4: Influence of the atmospheric transport parametrization on the evolution of the THC strength for a given, non-destabilizing, radiative forcing - Coupled Model

the dashed line the supercritical case. In the subcritical case, the minimum value of THC strength (which is $\approx 6\ Sv$) is reached at $t \approx 450\ y$; after that, and so still within the phase of increasing forcing, q oscillates with a period of $\approx 400\ y$. This means that the negative feedbacks overcome the external forcing and stabilize the system. We underline that in a somewhat similar pair of runs with the uncoupled model (see figure 7 in the Part I paper) the system closely followed the forcing up to the end of its increase; this relevant change in behavior of the system is essentially due to the conceptual difference between the target temperature restoring coefficient in the uncoupled model and the radiative temperature restoring coefficient B in the coupled model. We observe that at the newly established equilibrium -for $t \geq 1200\ y$- the THC strength is larger than before the perturbation, as observed in the uncoupled case for freshwater flux forcings; this means, considering that at equilibrium the value of q for a northern sinking equilibrium is given by 11, that the radiative forcing causes an increase in the Southern Hemisphere polewardmoisture flux F_3.

In order to explore the relevance of the sensitivity of the atmospheric transports in determining the intensity of the feedbacks of the system, we present in figures 4, 5, and 6 the effects for choices of $n = 1, 3, 5$ of a subcritical radiative forcing that lasts 1000 years and is such that the radiative equilibrium temperature increases by 2.5 $^\circ C$ per century in both the high-latitude boxes and by 3.75 $^\circ C$ per century in the tropical box; this corresponds to changes in the radiative forcings of $\approx 4\ Wm^{-2}$ per century and $6\ Wm^{-2}$ per century respectively and so to a globally averaged increase of the radiative forcing of $\approx 5\ Wm^{-2}$ per

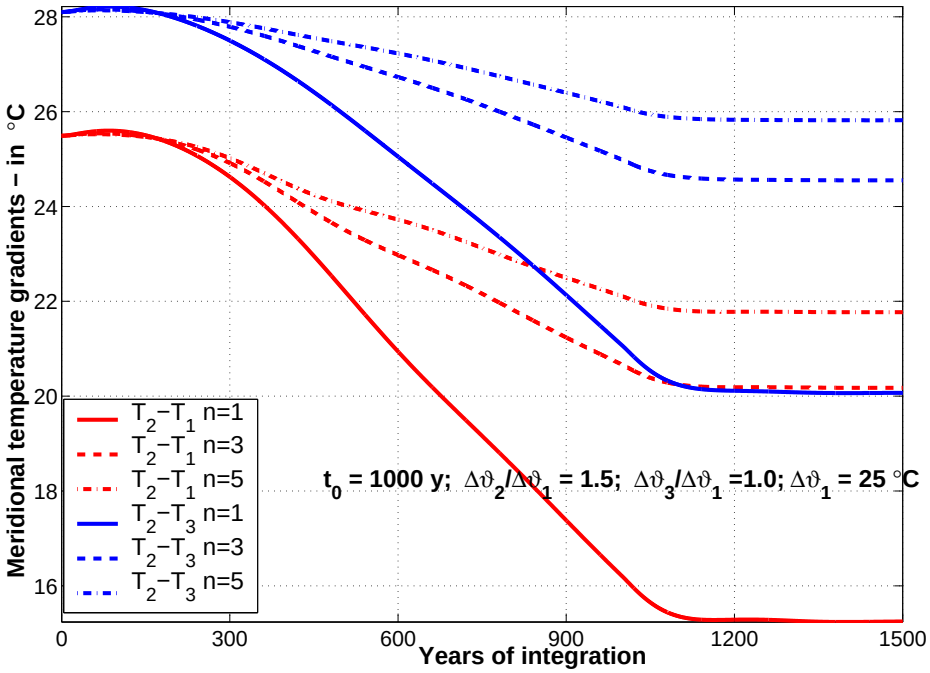

Figure 5: Influence of the atmospheric transport parametrization on the evolution of the Pole-to-Equator temperature gradients for a given, non-destabilizing, radiative forcing - Coupled Model

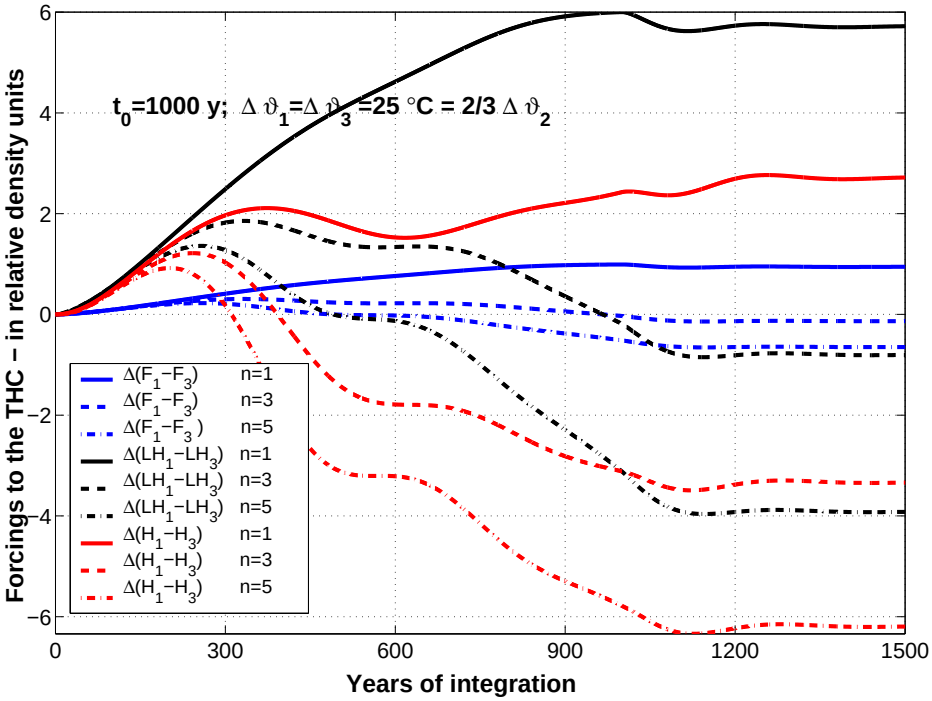

Figure 6: Influence of the atmospheric transport parametrization on the relative relevance of freshwater, latent heat and total atmospheric heat flux forcings for a given, non-destabilizing, radiative forcing - Coupled Model

century. In Figure 4 we show the evolution of the THC circulation: in all the three cases the decline stops around 200-250 years after the beginning of the forcing, when the negative feedbacks of the system described in figure 2 become relevant. The robustness of having an initial THC decrease under global warming conditions, as discussed in [Tziperman 2000b], is then confirmed also for our model.

We see that the extent of the decline and its delay are negatively correlated with the value of the exponent for the eddy transport power law, thus suggesting that a more sensitive atmospheric transport is more effective in stabilizing the system. In Figure 5 we show that there is eventually a decrease of the meridional temperature gradients in both hemispheres; this means that the atmospheric transports must become so efficient that they overcome the larger radiative forcing in the tropics, the decrease of the oceanic transport, and redistribute the heat all over the globe. Therefore, as we will see later, if we force the system more intensely at the tropics and keep the high latitudes forcings unchanged, the system response generally is not qualitatively different from what it is if the forcing were stronger everywhere. The evolution of the latent heat and freshwater flux differences between box 1 and box 3 seems then the most important part of the atmospheric processes that influences both the destabilization and the stabilization of the THC.

Similar conclusions have been reached by other studies considering either relatively sophisticated [Stocker and Schmittner 1997, Schmittner and Stocker 1999] or extremely simplified models [Tziperman and Gildor 2002]: the evolution of latent heat and freshwater meridional fluxes under global warming scenarios depends on two competing effects: the reduced efficiency in the atmospheric transport due to decreased meridional temperature gradient and increased capability of the atmosphere to retain moisture for higher average temperatures. The second effect dominates for larger climate changes, as proved by positive correlation between average global temperature and accumulation in the glaciers in paleoclimatic data or simulations [Charles et al. 1994, Manabe and Stouffer 1994, Krinner and Genthon 1998, Kitoh et al. 2001], and it is reasonable to expect that similarly the Clausius-Clapeyron effect dominates the changes in the latent heat and freshwater fluxes in global warming scenarios[Tziperman and Gildor 2002].

5.1 Comparison between the relevance of latent heat vs. relevance of freshwater flux changes

In figure 6 we present in density units the contributions of the processes that tend to destabilize the THC: the change from the initial value (indicated by the symbol Δ) of the differences between the the freshwater fluxes, latent heat fluxes and total surface heat fluxes into box 1 and into box 3. We see that the contribution to the destabilization of the THC of the

freshwater flux difference is quite small relatively to total heat flux difference's; the main reason is that the increase of the latter is entirely due to the increase in the latent heat flux difference, which gives, as shown equation 19, a contribution to the increase of buoyancy flux into box 1 that is about six times as large as the contribution due to the freshwater flux. The other components of the total heat flux tend to stabilize the system but their contribution in this sense is relatively negligible. Therefore the largest contribution to the destabilization is thermic and driven by the latent heat. This observation holds for all the three cases analyzed. In the longer term, when the THC recovers, the evolution of the total heat fluxes' difference between box 1 and box 3 changes notably with n, since it increases from the initial value for $n = 1$, while it decreases for $n = 3$ and $n = 5$. It is clear from the figures that most of this difference can be attributed to the difference of the latent heat fluxes between the two boxes 1 and 3 alone.

Two studies with coupled atmosphere-ocean GCMs have obtained contrasting results about the relative relevance of heat *vs.* freshwater as destabilizing mechanisms; in [Mikolajewicz and Voss 2000] it is shown that the heat flux change is the most important destabilizing mechanism and that this change is dominated by the latent heat flux, just as in our model. By contrast, in [Dixon et al. 1999] it is concluded that changes in the moisture flux were the main destabilizing agents. However, calculations of the contributions of changes in both the heat and moisture fluxes to the change in the density flux in the model used in [Dixon et al. 1999] show that it is in fact dominated by the heat flux changes [Huang et al. 2003]. Indeed this is true of all the seven coupled GCMs analyzed in the context of the Climate Intercomparison Project (CMIP) [Huang et al. 2003]. This apparent contradiction may be expalined by the results presented in [Kamenkovich 2003], where it is found that, even though the decrease in the thermohaline circulation may be initiated by changes in the moisture flux, changes in the heat flux induced by atmospheric feedbacks nevertheless contribuite strongly to the decrease.

6 Critical Perturbations

Once the parameters of atmospheric transport efficiency n and of ratio of the the box 2-to-box 1 radiative forcing are chosen, it is possible to uniquely identify each forcing applied using three coordinates, descriptive of the spatial and temporal pattern of the perturbation. In this section we present the set of critical forcings, which divide the forcings that disrupt the present pattern of the THC from those that drive the system to a northern sinking state qualitatively similar to the initial unperturbed state. We consider an average *model+forcing* case, which is characterized by $n = 3$ and $\Delta\vartheta_2/\Delta\vartheta_1 = 1.5$. In figure 7 we present the manifold of those critical forcings using the coordinate system $[t_0, \Delta\vartheta_3/\Delta\vartheta_1, \Delta\vartheta_1]$, while in

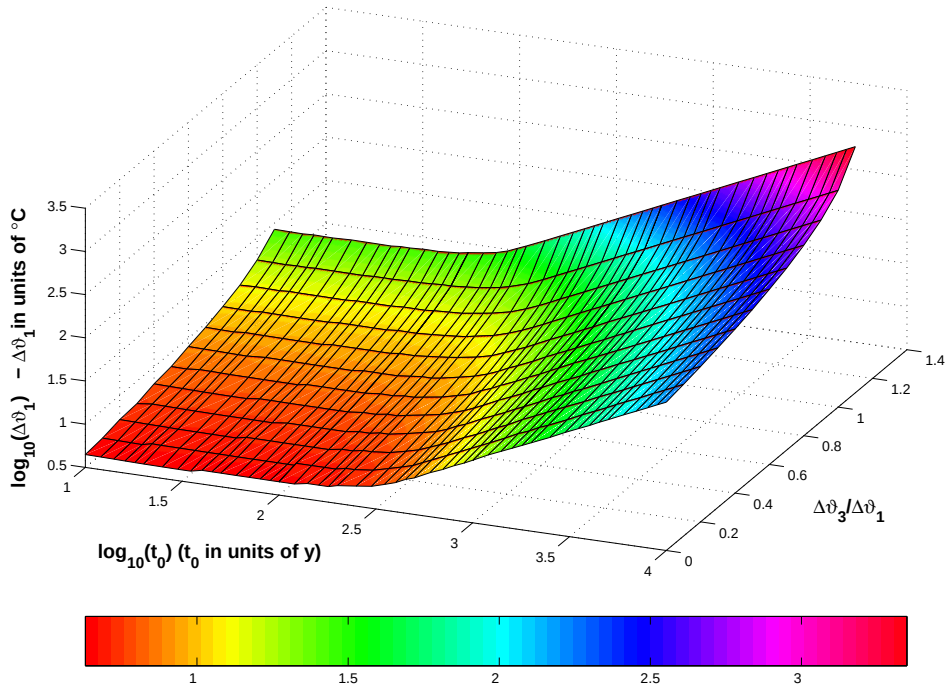

Figure 7: Critical values of the total increase of the radiative temperatures - Coupled Model

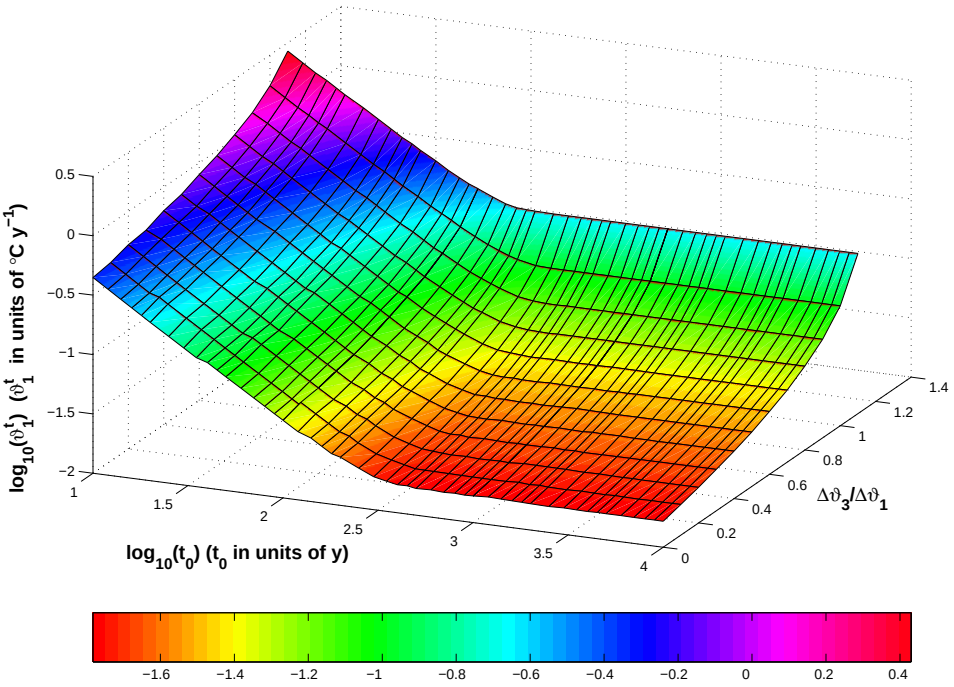

Figure 8: Critical values of the rate of increase of the radiative temperatures - Coupled Model

figure 8 we adopt the coordinate system $[t_0, \Delta\vartheta_3/\Delta\vartheta_1, \vartheta_1^t]$. There is a general agreement between the response of the coupled and of the uncoupled model, presented in the Part I paper, to destabilizing perturbations. In figure 7 we observe that the more symmetric and the slower the forcing, the less likely is the destabilization of the THC:

- for a given $\Delta\vartheta_1$, the lower is the value of the ratio $\Delta\vartheta_3/\Delta\vartheta_1$, the lower is the total change $\Delta\vartheta_1$ needed to obtain the reversal of the THC;
- for a given value of the ratio $\Delta\vartheta_3/\Delta\vartheta_1$, more rapidly increasing perturbations (larger ϑ_1^t) are more effective in disrupting the circulation.

Nonetheless, the results for the coupled model differ from those obtained in the uncoupled case for two main reasons, which are apparently contradictory.

First, observing figure 7, we see that even for strongly asymmetric perturbations having very low values of $\Delta\vartheta_3/\Delta\vartheta_1$, there is a threshold in the rate of increase of the forcing below which the reversed THC does not occur, independently of the total radiative forcing realized; in figure 8 we can observe that, for each value of $\Delta\vartheta_3/\Delta\vartheta_1$, the critical value of the rate of increase of the forcing ϑ_1^t is independent of the temporal extension of the forcing t_0. This implies that for $n = 3$ the system cannot make transitions to a southern sinking equilibrium for quasi-static perturbations, since they would require indefinitely large perturbations. If $n = 1$ there is no threshold in the rate of increase for low values of $\Delta\vartheta_3/\Delta\vartheta_1$, which implies that in these cases quasi-static perturbations can cause the collapse of the THC; in order to explore this property of the system when the atmospheric transport is less sensitive to changes in the meridional temperature gradients, in section 8 we will discuss the bifurcation diagrams.

Second, we can obtain the collapse of the THC for forcings that are larger in the southern high latitude box (note that in figures 7 and 8 the domain extends beyond 1 in the $\Delta\vartheta_3/\Delta\vartheta_1$ direction). We can understand this behavior considering the feedback mechanism previously described in figure 3. In the first place, if $\Delta\vartheta_3 \geq \Delta\vartheta_1$, it is possible to obtain destabilization of the system in steps $2a$ and $3a$ if the forcing is fast (and so able to avoid the feedback steps $2b$ and $2c$), essentially because the equilibrium control of the latent heat and freshwater fluxes into box 1 is larger, and so a smaller relative increase can result is a larger absolute increase. Moreover, if the northern meridional temperature gradient is radiatively forced to increase more than the southern, the atmospheric diffusivity feedback will enhance more the sensible heat flux into box 1 in step 1. We note that in the uncoupled case it was not possible in any case to destabilize the THC if the forcings were stronger in box 3.

Previous studies focusing on more complex models obtain a similar dependence of thresholds on the rate of increase of the forcings [Stocker and Schmittner 1997, Schmittner and Stocker 1999], while in other studies where the full collapse of the THC is not obtained,

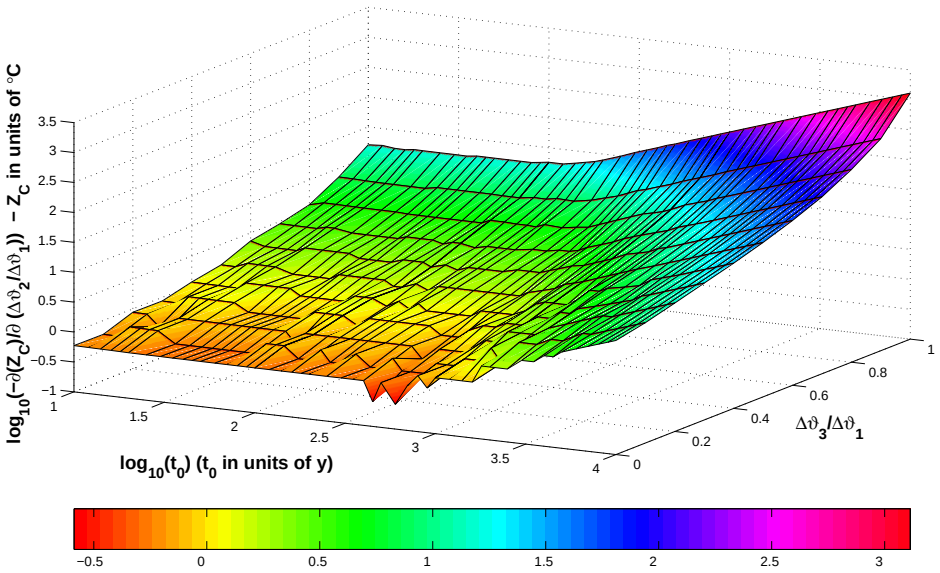

Figure 9: Sensitivity to the low-to-high latitudes radiative forcing ratio of of the critical values of the total increase of the radiative temperatures - Coupled Model

it is nevertheless observed that the higher the rate of increase of the forcing, the larger the decrease of THC realized [Stouffer and Manabe 1999]. Other studies on coupled models also show how the spatial pattern of freshwater forcing due to global warming is extremely relevant especially in the short time scales: only forcings occurring mainly in the Northern Atlantic are efficient in destabilizing the THC [Rahmstorf 1996, Rahmstorf and Ganopolski 1999, Manabe and Stouffer 2000, Ganopolski et al. 2001].

7 Sensitivity Study

From the analysis of the figures 3, 4, 5, and 6 and from the analysis of the feedbacks of the system, we have found that the difference between the latent heat fluxes into the two high-latitude boxes dominates the dynamics of the forced system and determines its stability. Given the properties and the functional form of LH_1 and LH_3 we expect that:

1. the system is less stable against radiative forcings if the tropical to high latitudes radiation forcing ratio is larger, without preference for the time scale of the forcing;

2. in the case when the radiative forcing is larger in the northern high-latitude box, the system is more stable if the atmospheric transport feedback is stronger; this effect is likely to be notable *only if* the perturbations have times scales larger than the flushing time of the oceanic boxes, as confirmed in the extreme case of quasi-static perturbations;

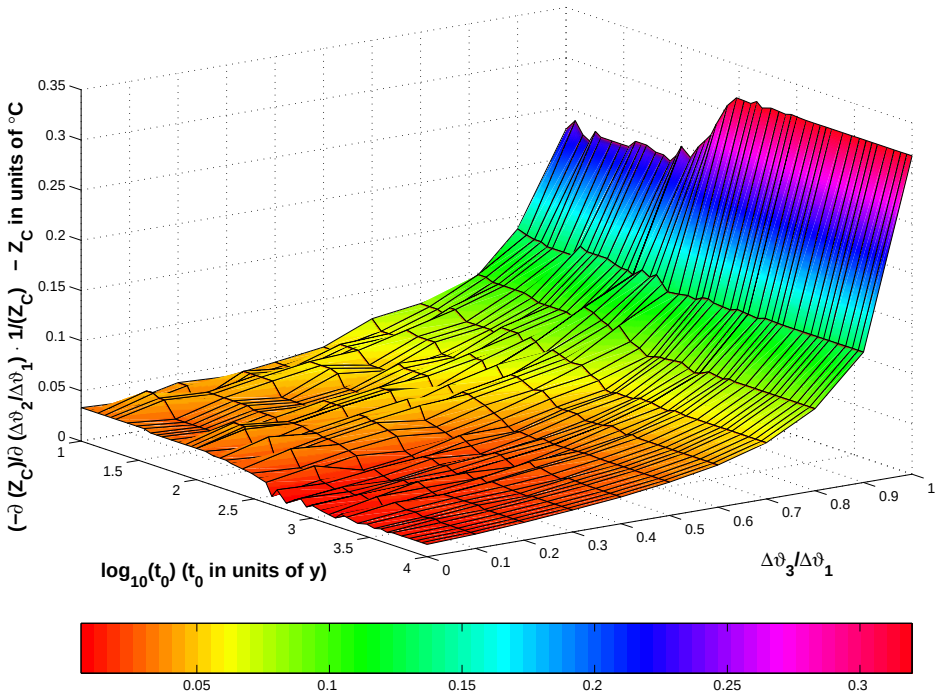

Figure 10: Weighted sensitivity to the low-to-high latitudes radiative forcing ratio of of the critical values of the total increase of the radiative temperatures - Coupled Model

3. in the case when the radiative forcing is fast and larger in the southern high-latitude box, the system is less stable if the atmospheric transport feedback is stronger.

To obtain a more quantitative measure of which processes are important, we analyze the sensitivity of the $z-$coordinate of the the manifold of the critical perturbations shown in figure 7 to changes in key parameters; we define Z_C as the the critical values of $\Delta\vartheta_1$ presented in figure 7. In particular we consider the 2-dimensional fields of the finite difference estimates of the partial derivatives $\partial(Z_C)/\partial(\Delta\vartheta_2/\Delta\vartheta_1)$ and $\partial(Z_C)/\partial(n)$. Figure 9 shows $-\partial(Z_C)/\partial(\Delta\vartheta_2/\Delta\vartheta_1)$. This manifold is very similar to figure 7: this suggests that Z_C changes roughly in proportion to its actual value when $\Delta\vartheta_2/\Delta\vartheta_1$ changes. In figure 10 we present the manifold given by the values of $-\partial(Z_C)/\partial(\Delta\vartheta_2/\Delta\vartheta_1) \cdot 1/(Z_C)$. This manifold is effectively flat for $\Delta\vartheta_3/\Delta\vartheta_1 \leq 0.7$ and is everywhere independent of the value of t_0. In physical terms, this means that the increase in the ratio between the tropical and the northern high latitudes radiative forcing changes the response of the system and favors the collapse of the THC evenly and independently of the temporal scale of the forcing itself, and is particularly effective if we use quasi-symmetric forcings. Therefore the change of the parameter $\Delta\vartheta_2/\Delta\vartheta_1$ does not have preferential effect on any of the feedbacks.

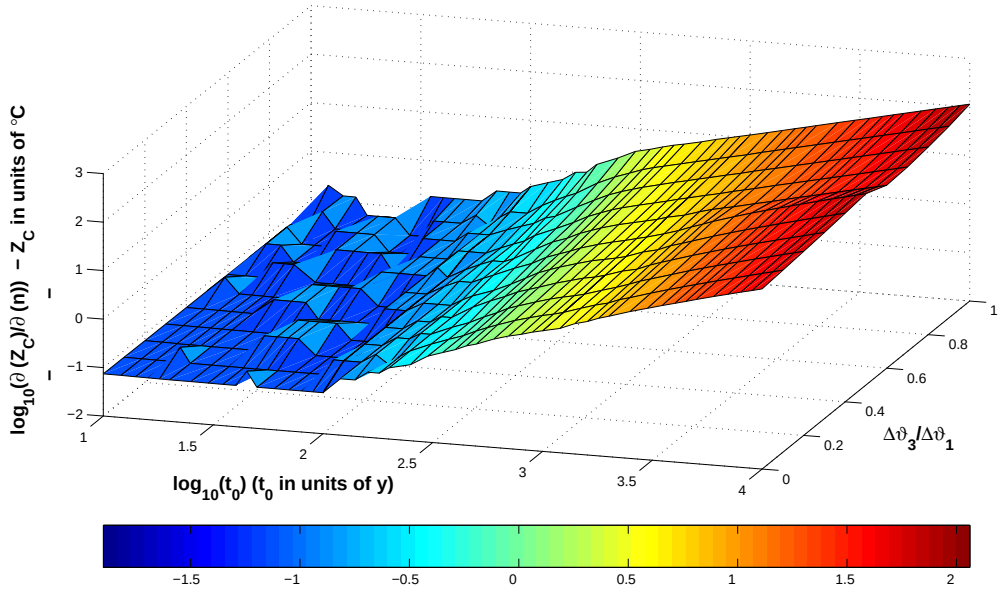

Figure 11: Sensitivity to the atmospheric transport parametrization of of the critical values of the total increase of the radiative temperatures - Coupled Model

Figure 11 presents the value of $\partial(Z_C)/\partial(n)$. We observe that this graph has almost no $\Delta\vartheta_3/\Delta\vartheta_1$ dependence, and its most striking feature is that the surface rises abruptly from a flat and low region for $t_0 \approx 300\ y$. The sensitivity of the measure of system stability to the radiative forcings with respect to the efficiency of the atmospheric transport feedback is small and positive for forcings having temporal scale shorter than the characteristic oceanic time scale of the system, while it becomes very large and positive for forcings having long temporal scales. The positive value means that a more sensitive atmospheric transport (higher values of n) stabilizes if $\Delta\vartheta_3/\Delta\vartheta_1 \leq 1$. With a more temperature-gradient sensitive atmospheric transport smaller changes in temperature gradients between the boxes are needed (see figure 5) to change *all* the atmospheric fluxes; therefore the system is able to dynamically arrange very effectively the atmospheric fluxes, so that they can counteract more efficiently the external forcings and keep the system as close as possible to the initial state. If the forcings are very fast, the enhancement of the atmospheric stabilizing mechanism is not very effective. On the contrary, for slow forcings involving time scales comparable to or larger than those of the system, the enhanced strength of the negative feedback obtained with increasing efficiency of the atmospheric transport can play a very significant stabilizing role in the dynamics of the system. This seem to be in contrast with the result that in the uncoupled model shorter relaxation times for box temperatures -which as shown in [Marotzke 1996] correspond to more sensitive atmospheric heat transports- imply less stability for the system [Tziperman et al. 1994, Nakamura et al. 1994, Rahmstorf 2000]. Actually the contrast is

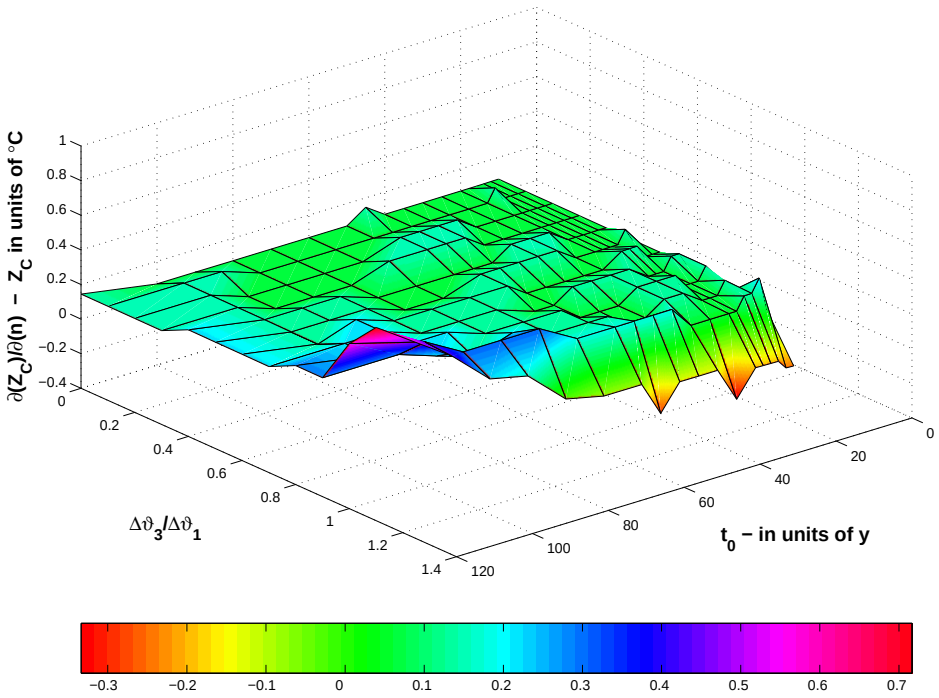

Figure 12: Sensitivity to the atmospheric transport parametrization of of the critical values of the total increase of the radiative temperatures (detail) - Coupled Model

only apparent, the point being that in an uncoupled model there is no adjustment of of moisture fluxes due to temperature changes: therefore decreasing the relaxation time make the system *less flexible* and adaptable to forcings. Our results seem to disagree with the conclusions drawn in the coupled model presented in [Scott et al. 1999]. We think that the disagreement is mainly due to the fact that in this study we are dealing with a different kind of perturbation, descriptive of global warming, and that these perturbations have a direct and strong influence on the most relevant atmospheric fluxes, because of their highly nonlinear dependence on average temperature changes; this property was not present in the parametrizations used in [Scott et al. 1999].

By contrast with the above described result, we see in figure 12 that for $\Delta\vartheta_3/\Delta\vartheta_1 > 1$ and $t_0 \leq 80\ y$, the sensitivity $\partial(Z_C)/\partial(n)$ is negative, therefore a system with larger n is more easily destabilized. The destabilizing feedbacks in the special case $\Delta\vartheta_3/\Delta\vartheta_1 > 1$ discussed in the previous section become more and more relevant with increasing n: this is especially clear for the increase of efficiency of the destabilizing changes in sensible heat for higher power laws for the diffusivity.

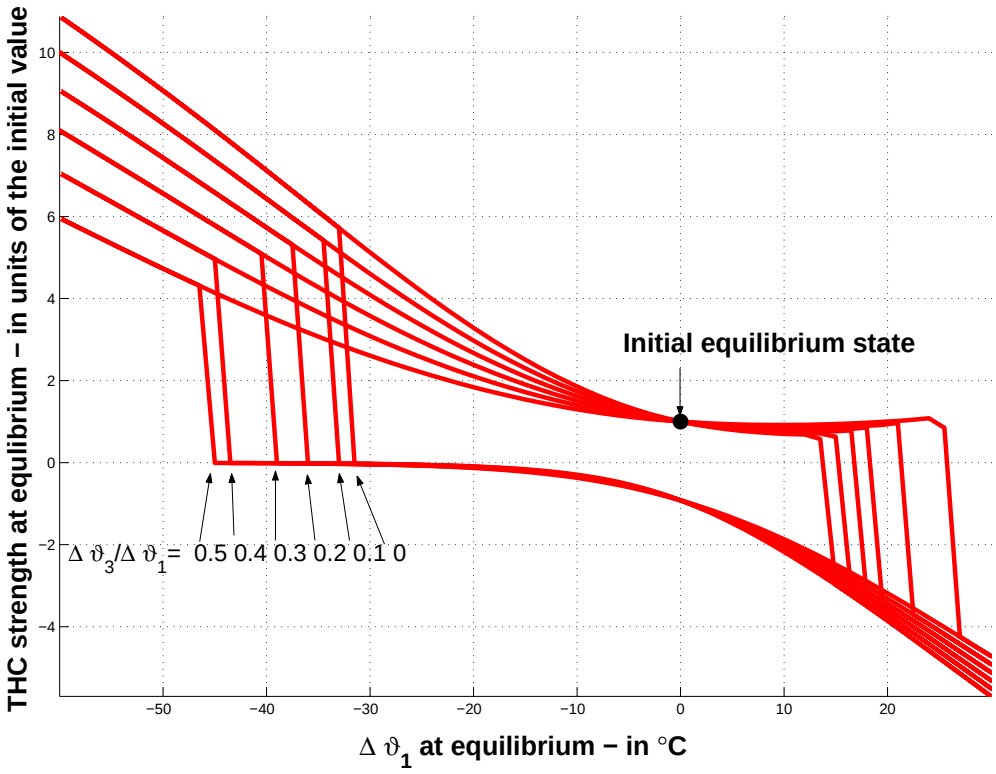

Figure 13: Bifurcation diagram of radiative temperature perturbations - Coupled Model with $n = 1$

8 Bifurcations

Choosing quasi-static perturbations to the radiative temperatures can lead to the reversal of the THC only if we select $n = 1$; this qualitative difference between the behavior of the various versions of the model is independent of the parameter $\Delta\vartheta_2/\Delta\vartheta_1$, because with a more temperature-gradient sensitive atmospheric transport the system is always able to counteract a slowly increasing destabilizing radiative forcing. With the choice of $n = 1$ and $\Delta\vartheta_2/\Delta\vartheta_1 = 1.5$, we have that quasi/static perturbations are destabilizing only if $\Delta\vartheta_3/\Delta\vartheta_1 \leq 0.5$ We present in figure 13 the bifurcation diagrams relative to $\Delta\vartheta_3/\Delta\vartheta_1 = [0, 0.1, 0.2, 0.3, 0.4, 0.5]$; we have $\Delta\vartheta_1$ as abscissa and q in units of the initial equilibrium value as ordinate, so that the initial state is the point $(0, 1)$. The abscissae of the subcritical Hopf bifurcation points on the right hand-side of the graph increase with increasing value of $\Delta\vartheta_3/\Delta\vartheta_1$ from $\approx 12°C$ to $\approx 25°C$, while the ordinates are close to 1, thus implying that when the bifurcation occurs the the THC is only slightly different from the initial equilibrium value. The bistable region is remarkably large in all cases, the total extent increasing with increasing value of $\Delta\vartheta_3/\Delta\vartheta_1$, because the abscissae of the bifurcation points in the left hand-side of the graph circuits are in all cases below $-30°C$ and so well within the unphysical region of parameter space. This means that once the circulation has reversed, the newly established pattern is extremely

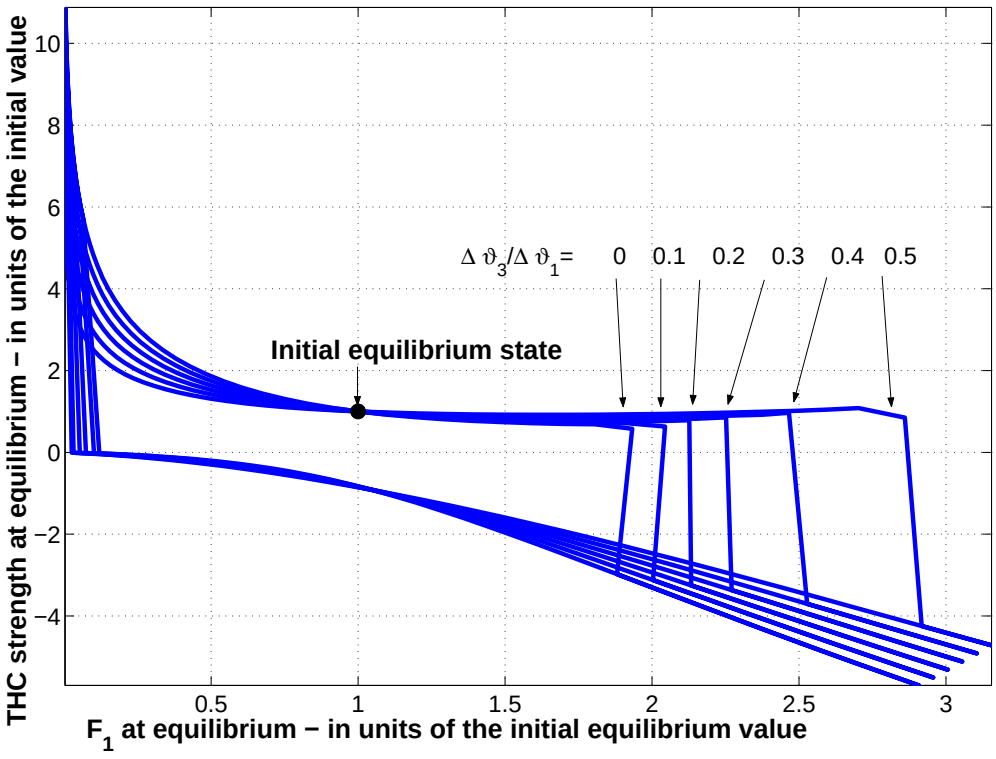

Figure 14: Bifurcation diagram of freshwater flux perturbations - Coupled Model with $n = 1$

stable and can hardly be changed again. In figure 14 we present a bifurcation graph where the abscissa is the value of the freshwater flux into box 1 when the radiatively forced system has reached a newly established equilibrium (the initial equilibrium is the point (1,1)): we can see that there is a monotonic 1 to 1 mapping between figures 13 and 14 (apart from a very limited region around the bifurcation point in the right hand side for low $\Delta\vartheta_3/\Delta\vartheta_1$ values), which suggests that in our model, for a given value of $\Delta\vartheta_3/\Delta\vartheta_1$, at equilibrium the changes of the average surface temperature and the changes in the hydrological cycle are positively correlated. We observe that in terms of freshwater flux the bistable region is very limited (almost by an order of magnitude) with respect to the uncoupled case (see figure 6 in the Part I paper), thus suggesting a caveat in the interpretation of uncoupled models' results.

9 Conclusions

In this paper we have analyzed the stability of the THC as described by a set of coupled models differing in the ratio between the radiative forcing realized at the tropics and at high latitudes and in the parametrization of the atmospheric transports.

In a coupled model a natural representation of the radiative forcing is possible, since

the main atmospheric physical processes responsible for freshwater and heat fluxes are formulated separately. Although only weakly asymmetric or symmetric radiative forcings are representative of physically reasonable conditions, we have considered general asymmetric forcings, in order to get a more complete picture of the mathematical properties of the system.

We have analyzed five combinations of the system *model+radiative forcing*, considering different combinations of the atmospheric transport parametrization and of the ratio between the high to low latitudes radiative forcing.

When the system is radiatively forced, initially the latent heat fluxes and the freshwater fluxes are strongly enhanced, thanks to the increase in the saturation pressure of the water vapor due to the warming, and the fluxes into the northern high latitude box increase. These are the main causes for the reduction of the THC strength. The strong increase of heat flux into box 1 reduces the efficacy of the atmospheric transport and so causes a great reduction of the freshwater flux and of the latent heat (and sensible heat) flux into box 1, which induces an increase in the THC. The variations of latent heat fluxes and of freshwater fluxes into the two high-latitude boxes, or better, the difference between the variations of those fluxes in box 1 and box 3, then dominate the dynamics of the forced coupled model, the main reason being, in the context of global warming experiments, the very strong dependence of these fluxes on the average temperature due to the Clausius-Clapeyron relation, whose inclusion in the fluxes' parametrization seem to play a key role in all the results obtained in our study. In our system the major role is played by the latent heat because in density terms it is stronger than the freshwater flux by a factor of ≈ 6. A qualitatively similar dominance is found in the CMIP models [Huang et al. 2003] and by [Mikolajewicz and Voss 2000], while other studies give the opposite result [Stocker and Schmittner 1997, Rahmstorf and Ganopolski 1999, Schmittner and Stocker 1999, Manabe and Stouffer 1999b, Ganopolski et al. 2001].

We obtain, with a parametric study involving the total forcing realized, its rate of increase, and its North-South asymmetry, the manifold of the critical forcings dividing the forcings driving the system to a southern sinking equilibrium from those that do not qualitatively change the pattern of the THC. We generally find that fast forcings are more effective than slow forcings in disrupting the present THC patterns, forcings that are stronger on the northern box are also more effective in destabilizing the system, and that very slow forcings do not destabilize the system whatever their asymmetry, unless the atmospheric transport is only weakly dependent on the meridional temperature gradient.

These results confirm those obtained by [Tziperman and Gildor 2002] for a hemispheric coupled box model and confirm the results obtained by [Stocker and Schmittner 1997, Schmittner and Stocker 1999] in the context of EMICSs and by [Manabe and Stouffer 1999a, Manabe and Stouffer 1999b, Manabe and Stouffer 2000, Stouffer and Manabe 1999] in the context of

GCMs, thus completing the whole hierarchical ladder of model and providing evidence for the robustness of these results.

We also compute the sensitivity of the results obtained with respect to the tropical-to-high latitude radiative forcing and with respect to the efficiency of the atmospheric transport. A higher forcing in the tropics destabilizes the system evenly at every time scale and greatly favors destabilization for quasi-symmetric forcings. Increasing the efficiency of the atmospheric transports makes the system in general more stable against destabilizing radiative forcings because it allows a very effective control of all the density fluxes and provides an enhancement of the atmospheric transport's negative feedback; this effect is especially relevant for forcings' time scales comparable or larger than the system's characteristic time scale, while it is not notable for fast forcings, which bypass all the feedbacks of the system.

The bifurcation diagrams of the system confirms in more formal terms that quasi-static perturbations cannot disrupt the northern sinking pattern of the circulation unless we are considering a relatively inefficient atmospheric transport and radiative forcing with large, unrealistic North-South asymmetry; the hysteresis loops realized in these cases show that the initial equilibrium point is in the bistable region.

Therefore when analyzing the behavior of the THC in global warming scenarios with more complex coupled models, the spatial pattern of the radiative forcing should be carefully taken into account, because the THC is a highly nonlinear, nonsymmetric system, and the effect of changing the rate of increase of the radiative forcing should be explored in great detail, thus providing a bridge between the instantaneous and quasi-static changes, because the THC dynamics encompasses very different time-scales, which can be explored only if the timing of the forcing is varied.

We conclude by pointing out some possible improvements to the present model and possible extensions of this work. In this study we have put emphasis on how the atmospheric transport processes, the local radiative budget and the oceanic advection can reverse the symmetry of the pattern of the circulation in the context of a rigorously symmetric geometry. The system is likely to be very sensitive to changes in the volumes of the boxes and especially to the introduction of asymmetries between the two high-latitude boxes, which would also make the system more realistic since the the southern mid-high latitude portion of the Atlantic is considerably larger than the northern mid-high latitude portion. Asymmetries in the oceanic fractional areas would induce asymmetries in the values of B_i, thus causing the presence of different restoring times for the various boxes, while asymmetries in the freshwater catchment areas ($\gamma_1 \neq \gamma_3$) would make the relative importance of the latent heat fluxes and of the freshwater fluxes (when expressed in common density units) different in the two boxes $i = 1, 3$.

The debate of whether the heat or freshwater fluxes are the most relevant in destabilizing

the THC circulation would benefit by including a more appropriate nonlinear equation of state for the water density; the likely effect would be increasing the relative relevance of the freshwater fluxes.

The presence of the albedo feedback would also enhance any asymmetry between the two high latitudes boxes: it could be included in the model by parameterizing the radiative terms A_i as increasing functions of the temperatures T_i, along the lines of [Stocker and Schmittner 1997, Schmittner and Stocker 1999, Tziperman and Gildor 2002], considering the temperatures as proxies for the fraction of the surface covered by ice; we expect that the inclusion of an ice-albedo feedback would decrease the stability of the THC.

The introduction of a noise component in the tendency equations along the lines of [Titz et al. 2002a, Titz et al. 2002b] would increase the realism of the model, since it would introduce a parametrization of the high-frequency variability; recent studies have undertaken this strategy and shown that close to instability threshold the evolution of the THC has a very limited predictability [Knutti and Stocker 2001], and that stochastic resonance could be responsible for glacial/interglacial climate shifts [Ganopolski and Rahmstorf 2002]. It would be interesting to analyze how the intensity and the color of the noise would influence the results obtained and the conclusions drawn in this work.

Finally, the model could be improved so that it could be able to represent the main features of the whole conveyor belt, by adding other boxes, descriptive of other oceanic basins, and carefully setting the connections between the boxes, along the lines of some examples presented in [Weaver and Hughes 1992].

Acknowledgments

The authors are grateful to J. Scott for interesting discussions and useful suggestions. One author (V.L.) wishes to thank R. Stouffer for having proposed improvements to an earlier version of the manuscript, and T. Stocker and E. Tziperman for having suggested a number of relevant references. This research was funded in part by the US Department of Energy's (DOE) Climate Change Prediction Program, in part by the Alliance for Global Sustainability (AGS), and in part by the MIT Joint Program on the Science and Policy of Global Change (JPSPGC). Financial support does not constitute an endorsement by DOE, AGS, or JPSPGC of the views expressed in this article.

References

[Baumgartner and Reichel 1975] Baumgartner A., and E. Reichel, The World Water Balance (Elsevier, New York, 1975)

[Boyle and Keigwin 1987] Boyle E. A., and L. Keigwin, North Atlantic thermohaline circulation during the past 20000 years linked to high-latitude surface temperature, Nature 330, 35-40 (1987)

[Bretherton 1982] Bretherton F. P., Ocean climate modeling, Progr. Oceanogr., 11, 93-129 (1982)

[Broecker et al. 1985] Broecker W. S., D. M. Peteet, and D. Rind, Does the ocean-atmosphere system have more than one stable mode of operation?, Nature, 315, 21-26 (1985)

[Broecker 1994] Broecker, W. S., Massive iceberg discharges as triggers for global climate change, Nature 372, 421-424 (1994)

[Broecker 1997] Broecker W. S., Thermohaline circulation, the Achilles heel of our climate system: Will man-made CO2 upset the current balance? Science, 278, 1582-1588 (1997)

[Bryan 1986] Bryan, F., High-latitude salinity effects and interhemispheric thermohaline circulations, Nature, 323, 301-304 (1986)

[Charles et al. 1994] Charles, C. D., D. Rind, J. Jouzel, R. D. Koster, and R. G. Fairbanks, Glacial/interglacial changes in moisture sources for Greenland: Influences on the ice core record of climate, Science, 263, 508-518 (1994)

[Cubasch et al. 2001] Cubasch, U., G. A. Meehl, G. J. Boer, R. J. Stouffer, M. Dix, A. Noda, C. A. Senior, S. Raper, K. S. Yap, A. Abe-Ouchi, S. Brinkop, M. Claussen, M. Collins, J. Evans, I. Fischer-Bruns, G. Flato, J. C. Fyfe, A. Ganopolski, J. M. Gregory, Z.-Z. Hu, F. Joos, T. Knutson, R. Knutti, C. Landsea, L. Mearns, C. Milly, J. F. B. Mitchell, T. Nozawa, H. Paeth, J. Risnen, R. Sausen, S. Smith, T. Stocker, A.Timmermann, U. Ulbrich, A. Weaver, J. Wegner, P. Whetton, T. Wigley, M. Winton, and F. Zwiers, 9. Projections of future climate change, in Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge University Press, Cambridge, UK) pp. 526-582. (2001)

[Dixon et al. 1999] K. W. Dixon, T. L. Delworth, M. J. Spelman, and R. J. Stouffer, The influence of transient surface fluxes on North Atlantic overturning in a coupled GCM climate change experiment, Geophys. Res. Lett. 26, 2749-2752 (1999)

[Gill 1982] Gill A.E., Atmosphere-Ocean Dynamics (Academic Press, London, 1982)

[Held 1978] Held, I.M., The vertical scale of an unstable baroclinic wave and its importance for eddy heat flux parameterizations, J. Atmos. Sci. 35, 572-576 (1978)

[Ganopolski and Rahmstorf 2001] Ganopolski, A. and S. Rahmstorf, Stability and variability of the thermohaline circulation in the past and future: a study with a coupled model of intermediate complexity, in The Oceans and Rapid Climate Change: Past, Present and Future, edited by D. Seidov, M. Maslin and B. J. Haupt, AGU, Washington, pp. 261-275 (2001)

[Ganopolski and Rahmstorf 2002] Ganopolski, A. and S. Rahmstorf, Abrupt Glacial Climate Changes due to Stohastic Resonance, Phys. Rev. Lett. 88, DOI 038051 (2002)

[Ganopolski et al. 2001] Ganopolski, A., V. Petoukhov, S. Rahmstorf, V. Brovkin, M. Claussen and C. Kubatzki, 2001: CLIMBER-2: A climate system model of intermediate complexity. Part II: Model sensitivity. Clim. Dyn. 17, 735-751 (2001)

[Huang et al. 2003] Huang B., P. H. Stone, A. P. Sokolov, and I. V. Kamenkovich, 2003: The deep-ocean heat uptake in transient climate change. J. Climate, submitted.

[Hughes and Weaver 1994] Hughes T. C. M., and Weaver A. J., Multiple equilibrium of an asymmetric two-basin model, J. Phys. Oceanogr. 24, 619-637 (1994)

[Kamenkovich 2003] I. V. Kamenkovich, A. P. Sokolov, and P. H. Stone, Feedbacks affecting the response of the thermohaline circulation to increasing CO2: a study with a model of intermediate complexity, Clim. Dyn. DOI 10.1007/s00382-003-0325-5 (2003)

[Keigwin et al. 1994] Keigwin L. D., Curry W. B., Lehman S. J., and Johnsen S., The role of the deep ocean in North Atlantic climate change between 70 and 130 ky ago, Nature 371, 323-327 (1994)

[Kitoh et al. 2001] Kitoh A., S. Murakami, and H. Koide, A simulation of the Last Glacial Maximum with a coupled atmosphere/ocean GCM, Geophys. Res. Lett., 28, 2221-2224 (2001)

[Klinger and Marotzke 1999] Klinger B. A., and J. Marotzke, Behavior of double hemisphere thermohaline flow in a single basin, J. Phys. Oceanogr. 29, 382-399 (1999)

[Knutti and Stocker 2001] Knutti, R., and T. F. Stocker, Limited predictability of the future thermohaline circulation close to an instability threshold, J. Climate 15 176-186 (2001)

[Krasovskiy and Stone 1998] Krasovskiy Y. P. and P. H. Stone, Destabilization of the thermohaline circulation by atmospheric transports: An analytic solution, J. Climate 11, 1803-1811 (1998)

[Krinner and Genthon 1998] Krinner, G. and C. Genthon, GCM simulations of the Last Glacial Maximum surface climate of Greenland and Antarctica, Climate Dyn. 14, 741-758 (1998)

[Latif et al. 2000] Latif M., E. Roeckner, U. Mikolajewicz, and R. Voss, Tropical Stabilization of the Thermohaline Circulation in a Greenhouse Warming Simulation, J. Climate 13, 1809-1813 (2000)

[Levitus 1982] Levitus, S., Climatological atlas of the world ocean, NOAA Professional Paper, vol. 13, U.S. Department of Commerce, NOAA, Washington DC (1982)

[Macdonald and Wunsch 1996] Macdonald A.M., and C. Wunsch, An estimate of global ocean circulation and heat fluxes, Nature 382, 436-439 (1996)

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

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

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

[Manabe and Stouffer 1999a] Manabe S. and R. J. Stouffer, Are two modes of thermohalince circulation stable?, Tellus, 51A, 400-411 (1999)

[Manabe and Stouffer 1999b] Manabe S. and R. J. Stouffer, The role of thermohaline circulation in climate, Tellus, 51A-B(1), 91-109 (1999)

[Manabe and Stouffer 2000] Manabe, S. and R. J. Stouffer, Study of abrupt climate change by a coupled ocean-atmosphere model, Quaternary Science Reviews, 19 285-299 (2000)

[Marotzke and Willebrand 1991] Marotzke J. and J. Willebrand, Multiple equilibria of the global thermohaline circulation, J. Phys. Oceanogr. 21, 1372 (1991)

[Marotzke and Stone 1995] J. Marotzke and P. H. Stone, Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. J. Phys. Ocean., 25, 1350-136 (1995) .

[Marotzke 1996] Marotzke J., Analysis of thermohaline feedbacks, Decadal Climate Variability: Dynamics and predicatibility, Anderson D.L.T. and Willebrand J. Eds., Springer-Verlag, Berlin, pp. 333-378 (1996)

[Mikolajewicz and Maier-Reimer 1994] Mikolajewicz, U. and E. Maier-Reimer, Mixed boundary conditions in ocean general circulation models and their influence on the stability of the model's conveyor belt. J. Geophys. Res. 99, 22633-22644 (1994)

[Mikolajewicz and Voss 2000] Mikolajewicz U. and R. Voss, The role of the individual air-sea flux components in CO2-induced changes of the ocean's circulation and climate, Climate Dyn. 16, 627-642 (2000)

[Nakamura et al. 1994] Nakamura M., P. H. Stone, and J. Marotzke, Destabilization of the thermohaline circulation by atmospheric eddy transports, J. Climate 7, 1870-1882 (1994)

[Peixoto and Oort 1992] Peixoto A. and Oort B., Physics of Climate (American Institute of Physics, 1992)

[Petoukhov et al. 2000] Petoukhov, V., A. Ganopolski, V. Brovkin, M. Claussen, A. Eliseev, C. Kubatzki, and S. Rahmstorf, A climate system model of intermediate complexity. Part I: model description and performance for present climate. Climate Dyn., 16, 1-17 (2000)

[Rahmstorf 1995] Rahmstorf S., Bifurcations of the Atlantic Thermohaline circulation in response to changes in the hydrological cycle, Nature 378, 145-149 (1995)

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

[Rahmstorf 1996] Rahmstorf S., On the freshwater forcing and transport of the Atlantic thermohaline circulation, Climate Dyn. 12, 799-811 (1996)

[Rahmstorf 1997] Rahmstorf S., Risk of sea-change in the Atlantic, Nature, 388, 825-826 (1997)

[Rahmstorf 1999a] Rahmstorf S., Shifting seas in the greenhouse?, Nature, 399, 523-524 (1999)

[Rahmstorf 1999b] Rahmstorf S., Rapid Oscillation of the thermohaline ocean circulation,in Reconstructing ocean history: A window into the future, pp 139-149, Abrantes and Mix Eds, Kluwer Academic/Plenum Publishers, New York (1999)

[Rahmstorf 1999c] Rahmstorf S., Decadal Variability of the Thermohaline Ocean Circulation, pp. 309-332, in Beyond El Nio: Decadal and interdecadal climate variability,edited by A. Navarra, Springer, New York, (1999)

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

[Rahmstorf 2000] Rahmstorf S., The thermohaline ocean circulation - a system with dangerous thresholds? Climatic Change, 46, 247-256 (2000)

[Rahmstorf 2002] Rahmstorf S., Ocean circulation and climate during the past 120,000 years, Nature 419, 207-214 (2002)

[Ramanathan et al. 1979] Ramanathan, V., M. S .Lian, and R. D. Cess, Increased atmospheric CO2: zonal and seasonal estimates of the effect on the radiation energy balance and surface temperature, J. Geophys. Res., 84, 4949-4958 (1979)

[Roemmich and Wunsch 1985] Roemmich D.H., and Wunsch C., Two transatlantic sections: Meridional circulation and heat flux in the subtropical North Atlantic Ocean, Deep Sea Res., 32, 619-664, (1985)

[Rooth 1982] Rooth C., Hydrology and ocean circulation, Progr. Oceanogr., 11, 131-149 (1982)

[Schmittner and Stocker 1999] Schmittner, A., and T. F. Stocker, The stability of the thermohaline circulation in global warming experiments, J. Climate 12, 1117-1133 (1999)

[Scott et al. 1999] .Scott J. R., J. Marotzke, and P. H. Stone, 1999: Interhemispheric thermohaline circulation in a coupled box model, J. Phys. Oceanogr., 29, 351-365 (1999).

[Shine et al. 1995] Shine, K. P., B. P. Briegleb, A. S. Grossman, D. Hauglustaine, H. Mao, V. Ramaswamy, M. D. Schwarzkopf, R. Van Dorland, and W-C. Wang, Radiative forcing due to changes in ozone: A comparison of different codes. In Atmospheric Ozone as a Climate Gas, NATO ASI Series I, Vol. 32, Heidelberg, Germany: Springer-Verlag, pp. 373-396 (1995)

[Stocker and Wright 1991] Stocker T. F., and D.G. Wright, Rapid transitions of the ocean's deep circulation induced by changes in the surface water fluxes, Nature 351, 729-732 (1991)

[Stocker and Schmittner 1997] Stocker T. F. and A. Schmittner, Influence of CO_2 emission rates on the stability of the thermohaline circulation, Nature, 388, 862-865 (1997)

[Stocker 2000] Stocker T. F., Past and future reorganisations in the climate system. Quat. Sci. Rev. 19, 301-319 (2000)

[Stocker 2001] Stocker T. F., The Role of Simple Models in Understanding Climate Change In: Continuum Mechanics and Applications in Geophysics and the Environment. B. Straugham, R. Greve, H. Ehrentraut, Y. Wang (eds.) Springer Verlag, pp. 337-367 (2001)

[Stocker 2002] Stocker T. F., North-South Connections, Science 297, 1814-1815 (2002)

[Stommel 1961] Stommel H., Thermohaline convection with two stable regimes of flow, Tellus, 13, 224-230 (1961)

[Stone and Miller 1980] Stone P. H. and D. A. Miller, Empirical relations between seasonal changes in meridional temperature gradients and meridional fluxes of heat, J. Atmos. Sci. 37, 1708-1721 (1980)

[Stone and Yao 1990] Stone P. H. and M. S. Yao, Development of a two-dimensional zonally averaged statistical-dynamical model. Part III: The parametrization of eddy fluxes of heat and moisture. J. Climate 3, 726-740 (1990)

[Stone and Krasovskiy 1999] Stone P. H. and Y. P. Krasovskiy, Stability of the interhemispheric thermohaline circulation in a coupled box model. Dyn. Atmos. Oceans, 29, 415-435 (1999).

[Stouffer et al. 1991] Stouffer, R. J., S. Manabe, and K. Bryan, Climatic response to a gradual increase of atmospheric carbon dioxide, in Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations, The Netherlands, Elsevier Science Publishers, pp. 129-136 (1991)

[Stouffer and Manabe 1999] Stouffer, R. J., and S. Manabe, Response of a coupled ocean-atmosphere model to increasing atmospheric carbon dioxide: Sensitivity to the rate of increase. J. Climate 12, 2224-2237 (1999)

[Titz et al. 2002a] Titz, S., T. Kuhlbrodt, S. Rahmstorf, and U. Feudel, On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulation, Tellus A 54, 89-98 (2002)

[Titz et al. 2002b] Titz S. , T. Kuhlbrodt und U. Feudel, Homoclinic bifurcation in an ocean circulation box model, Int. J. Bif. Chaos 12, 869-875 (2002).

[Tziperman et al. 1994] Tziperman E., R. J. Toggweiler, Y. Feliks, and K. Bryan, 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 (1994)

[Tziperman 2000b] Tziperman E., Uncertainties in thermohaline circulation response to greenhouse warming, Geoph. Res. Lett. 27, 3077-30807 (2000)

[Tziperman and Gildor 2002] Tziperman E. and H. Gildor, The Stabilization of the Thermohaline Circulation by the Temperature/Precipitation Feedback, J. of Phys. Ocean. 32, 2704-2714. (2002)

[Wang et al. 1999a] Wang X., Stone P.H., and Marotzke J., Global Thermohaline circulation. Part I: Sensitivity to atmospheric moisture transport, J. Climate 12, 71-82 (1999a)

[Wang et al. 1999b] Wang X., Stone P. H. and J. Marotzke, Global Thermohaline circulation. Part II: Sensitivity with interactive atmospheric transport, J. Climate 12, 83-91 (1999b)

[Wang and Stone 1980] Wang W. C. and P. H. Stone, Effect of ice-albedo feedback on global sensitivity in a one-dimensional radiative-convective climate model, J. Atmos. Sci. 37, 545-552 (1980)

[Weaver and Hughes 1992] Weaver A. J., and T. M. C. Hughes, Stability of the thermohaline circulation and its links to Climate, Trends in Physical Oceanography, Research Trends Series, Council of Scientific Research Integration, Trivandrum, India, 1, 15 (1992)

[Wiebe and Weaver 1999] Wiebe E. C. and A. J.Weaver, On the sensitivity of global warming experiments to the parametrisation of sub-grid scale ocean mixing. Climate Dyn. 15, 875-893 (1999) .

[Zhang et al. 1993] Zhang, S., Greatbatch R.J., and Lin C.A., A re-examination of the polar halocline catastrophe and implications for coupled ocean-atmosphere modeling, J. Phys. Oceanogr. 23, 287 (1993)

Quantity	**Symbol**	**Value**
Mass of Box i, $i = 1, 3$	M	$1.08 \cdot 10^{20}\ Kg$
Box 2/Box i, $i = 1, 3$ mass ratio	V	2
Average water density	ρ_0	$1000\ Kg\ m^{-3}$
Specific heat per unit mass of water	c_p	$4 \cdot 10^3\ J\ {}^\circ C^{-1} Kg^{-1}$
Latent heat per unit mass of water	L_v	$2.5 \cdot 10^6\ JKg^{-1}$
Gas constant	R_v	$461\ J\ {}^\circ C^{-1} Kg^{-1}$
Average Salinity	S_0	$35\ psu$
Oceanic fractional area	ϵ	$1/6$
Box $i = 1, 3$ fractional water catchment area	$1/\gamma_i$	$1/2$
Thermal expansion coefficient	α	$1.5 \cdot 10^{-4}\ {}^\circ C^{-1}$
Haline expansion coefficient	β	$8 \cdot 10^{-4}\ psu^{-1}$
Hydraulic constant	k	$1.5 \cdot 10^{-6}\ s^{-1}$
Global climatic temperature/radiation elasticity [b]	κ_M	$0.6\ {}^\circ C\ W^{-1} m^2$

Table 1: **Value of the main model constants**

[b] Value relative to the whole planetary surface; corresponds to a *Global climate sensitivity* per CO_2 doubling of $\approx 2.5\ {}^\circ C$

Constant	**Box 1**	**Box 2**	**Box 3**
Temperature	$2.9\ {}^\circ C$	$28.4\ {}^\circ C$	$0.3\ {}^\circ C$
Salinity	34.7 psu	35.6 psu	34.1 psu
Atmospheric Freshwater Flux	0.41 Sv	-0.68 Sv	0.27 Sv
Total Surface Heat Flux	-1.58 PW	1.74 PW	-0.16 PW
-Latent Heat Flux	1.84 PW	-3.06 PW	1.22 PW
-Sensible Heat Flux	2.14 PW	-5.75 PW	3.61 PW
-Radiative Heat Flux	-5.57 PW	10.57 PW	-5.00 PW
Oceanic Heat Flux	1.58 PW	-1.74 PW	0.16 PW
THC strength	15.5 Sv	15.5 Sv	15.5 Sv
Radiative Equilibrium Temperature[a]	$-22.9\ {}^\circ C$	$52.9\ {}^\circ C$	$-22.9\ {}^\circ C$

Table 2: **Value of the fundamental parameters of the system at the initial equilibrium state**

REPORT SERIES of the MIT *Joint Program on the Science and Policy of Global Change*

1. **Uncertainty in Climate Change Policy Analysis** *Jacoby & Prinn* December 1994
2. **Description and Validation of the MIT Version of the GISS 2D Model** *Sokolov & Stone* June 1995
3. **Responses of Primary Production and Carbon Storage to Changes in Climate and Atmospheric CO_2 Concentration** *Xiao et al.* October 1995
4. **Application of the Probabilistic Collocation Method for an Uncertainty Analysis** *Webster et al.* January 1996
5. **World Energy Consumption and CO_2 Emissions: 1950-2050** *Schmalensee et al.* April 1996
6. **The MIT Emission Prediction and Policy Analysis (EPPA) Model** *Yang et al.* May 1996
7. **Integrated Global System Model for Climate Policy Analysis** *Prinn et al.* June 1996 (*superseded by No. 36*)
8. **Relative Roles of Changes in CO_2 and Climate to Equilibrium Responses of Net Primary Production and Carbon Storage** *Xiao et al.* June 1996
9. **CO_2 Emissions Limits: *Economic Adjustments and the Distribution of Burdens*** *Jacoby et al.* July 1997
10. **Modeling the Emissions of N_2O & CH_4 from the Terrestrial Biosphere to the Atmosphere** *Liu* August 1996
11. **Global Warming Projections: *Sensitivity to Deep Ocean Mixing*** *Sokolov & Stone* September 1996
12. **Net Primary Production of Ecosystems in China and its Equilibrium Responses to Climate Changes** *Xiao et al.* Nov 1996
13. **Greenhouse Policy Architectures and Institutions** *Schmalensee* November 1996
14. **What Does Stabilizing Greenhouse Gas Concentrations Mean?** *Jacoby et al.* November 1996
15. **Economic Assessment of CO_2 Capture and Disposal** *Eckaus et al.* December 1996
16. **What Drives Deforestation in the Brazilian Amazon?** *Pfaff* December 1996
17. **A Flexible Climate Model For Use In Integrated Assessments** *Sokolov & Stone* March 1997
18. **Transient Climate Change & Potential Croplands of the World in the 21st Century** *Xiao et al.* May 1997
19. **Joint Implementation: *Lessons from Title IV's Voluntary Compliance Programs*** *Atkeson* June 1997
20. **Parameterization of Urban Sub-grid Scale Processes in Global Atmospheric Chemistry Models** *Calbo et al.* July 1997
21. **Needed: A Realistic Strategy for Global Warming** *Jacoby, Prinn & Schmalensee* August 1997
22. **Same Science, Differing Policies; *The Saga of Global Climate Change*** *Skolnikoff* August 1997
23. **Uncertainty in the Oceanic Heat and Carbon Uptake & their Impact on Climate Projections** *Sokolov et al.* Sept 1997
24. **A Global Interactive Chemistry and Climate Model** *Wang, Prinn & Sokolov* September 1997
25. **Interactions Among Emissions, Atmospheric Chemistry and Climate Change** *Wang & Prinn* September 1997
26. **Necessary Conditions for Stabilization Agreements** *Yang & Jacoby* October 1997
27. **Annex I Differentiation Proposals: *Implications for Welfare, Equity and Policy*** *Reiner & Jacoby* October 1997
28. **Transient Climate Change & Net Ecosystem Production of the Terrestrial Biosphere** *Xiao et al.* November 1997
29. **Analysis of CO_2 Emissions from Fossil Fuel in Korea: 1961–1994** *Choi* November 1997
30. **Uncertainty in Future Carbon Emissions: *A Preliminary Exploration*** *Webster* November 1997
31. **Beyond Emissions Paths: *Rethinking the Climate Impacts of Emissions Protocols*** *Webster & Reiner* November 1997
32. **Kyoto's Unfinished Business** *Jacoby, Prinn & Schmalensee* June 1998
33. **Economic Development and the Structure of the Demand for Commercial Energy** *Judson et al.* April 1998
34. **Combined Effects of Anthropogenic Emissions & Resultant Climatic Changes on Atmosph. OH** *Wang & Prinn* April 1998
35. **Impact of Emissions, Chemistry, and Climate on Atmospheric Carbon Monoxide** *Wang & Prinn* April 1998
36. **Integrated Global System Model for Climate Policy Assessment: *Feedbacks and Sensitivity Studies*** *Prinn et al.* June 1998
37. **Quantifying the Uncertainty in Climate Predictions** *Webster & Sokolov* July 1998
38. **Sequential Climate Decisions Under Uncertainty: *An Integrated Framework*** *Valverde et al.* September 1998
39. **Uncertainty in Atmospheric CO_2 (Ocean Carbon Cycle Model Analysis)** *Holian* October 1998 (*superseded by No. 80*)
40. **Analysis of Post-Kyoto CO_2 Emissions Trading Using Marginal Abatement Curves** *Ellerman & Decaux* October 1998
41. **The Effects on Developing Countries of the Kyoto Protocol & CO_2 Emissions Trading** *Ellerman et al.* November 1998
42. **Obstacles to Global CO_2 Trading: *A Familiar Problem*** *Ellerman* November 1998
43. **The Uses and Misuses of Technology Development as a Component of Climate Policy** *Jacoby* November 1998
44. **Primary Aluminum Production: *Climate Policy, Emissions and Costs*** *Harnisch et al.* December 1998
45. **Multi-Gas Assessment of the Kyoto Protocol** *Reilly et al.* January 1999
46. **From Science to Policy: *The Science-Related Politics of Climate Change Policy in the U.S.*** *Skolnikoff* January 1999
47. **Constraining Uncertainties in Climate Models Using Climate Change Detection Techniques** *Forest et al.* April 1999
48. **Adjusting to Policy Expectations in Climate Change Modeling** *Shackley et al.* May 1999
49. **Toward a Useful Architecture for Climate Change Negotiations** *Jacoby et al.* May 1999
50. **A Study of the Effects of Natural Fertility, Weather & Productive Inputs in Chinese Agriculture** *Eckaus & Tso* July 1999
51. **Japanese Nuclear Power and the Kyoto Agreement** *Babiker, Reilly & Ellerman* August 1999
52. **Interactive Chemistry and Climate Models in Global Change Studies** *Wang & Prinn* September 1999

Contact the Joint Program Office to request a copy. The Report Series is distributed at no charge.

Contact the Joint Program Office to request a copy. The Report Series is distributed at no charge.