Full Text View

Volume 30, Issue 1 (January 2000)

Journal of Physical Oceanography
Article: pp. 90–104 | Abstract | PDF (364K)

Proximity of the Present-Day Thermohaline Circulation to an Instability Threshold

Eli Tziperman

Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel

(Manuscript received March 3, 1998, in final form February 16, 1999)

DOI: 10.1175/1520-0485(2000)030<0090:POTPDT>2.0.CO;2

ABSTRACT

The relation between the mean state of the thermohaline circulation (THC) and its stability is examined using a realistic-geometry primitive equation coupled ocean–atmosphere–ice global general circulation model. The main finding is that a thermohaline circulation that is 25% weaker and less dominated by thermal forcing than that of today's ocean is unstable within this coupled GCM. Unstable initial ocean climates lead in the coupled model to an increase of the THC, to strong oscillations, or to a THC collapse.

The existence of an unstable range of weak states of the THC provides a natural explanation for large-amplitude THC variability seen in the paleo record prior to the past 10 000 years: A weakening of the THC due to an external forcing (e.g., ice melting and freshening of the North Atlantic) may push it into the unstable regime. Once in this regime, the THC strongly oscillates due to the inherent instability of a weak THC. Hence the strong THC variability in this scenario does not result from switches between two or more quasi-stable steady states.

Table of Contents:
- Introduction
- Coupled model runs
- Instability of a weak,
- THC oscillations
- Conclusions
- REFERENCES
- FIGURES

1. Introduction

The thermohaline circulation (THC) is driven by the density difference between the light warm and salty midlatitude water and the denser fresh and cold high-latitude surface water. The freshwater forcing, resulting in a meridional surface salinity gradient, tends to oppose the thermal forcing and weaken the meridional density gradient and thus weaken the THC itself. Numerous ocean-only model studies from simple box models (Stommel 1961) to three-dimensional models (see reviews by McWilliams 1996; Weaver and Hughes 1992;Rahmstorf et al. 1996) have demonstrated the sensitivity of the North Atlantic THC to the amplitude and form of the air–sea freshwater forcing, in particular at high-latitude water-mass formation areas. Similar results have also been found in idealized coupled models (Saravanan and McWilliams 1995). In addition to the overall sensitivity of the THC to freshwater forcing, the many previous model studies of the THC have shown, in particular, that a weak THC obtained under large freshwater forcing (i.e., a less thermally dominant THC) tends to be unstable, while a stronger one obtained under weaker freshwater forcing (i.e., more thermally dominant) is more stable (Walin 1985; Marotzke et al. 1988; Weaver et al. 1991; Tziperman et al. 1994). An unstable THC will either oscillate strongly (Weaver et al. 1991; Drijfhout et al. 1996), reverse or collapse (Stommel 1961;Bryan 1986; Walin 1985; Marotzke et al. 1988; Mikolajewicz and Maier-Reimer 1994), or increase to a larger, more stable, value (Toggweiler et al. 1996).

The North Atlantic THC has been thermally dominant for many millions of years, perhaps since the Cretaceous, when deep warm and salty water may have formed at midlatitudes (Brass et al. 1982). Still, the"braking" effect due to larger freshwater forcing may have been stronger in previous geological periods, making the THC somewhat weaker and less thermally dominant (Boyle and Keigwin 1987). The present-day THC has been remarkably stable for about 10 000 years, while past geological periods have been far more unstable, with much larger amplitude variability of the climate and of the THC in particular (GRIP 1993; Dansgaard et al. 1993; Boyle and Keigwin 1987). This difference in the character of the THC variability now and during glacial times may be explained by the suggestion that while the present day THC is stable, it may not be far from an unstable regime (Tziperman et al. 1994). This proximity to an instability threshold was suggested based on model experiments using a realistic-geometry ocean-only GCM and following the simple box model study of Walin (1985). According to this proposed scenario, factors external to the ocean, such as the melting of land glaciers or higher precipitation rates in the North Atlantic, have reduced the salinity of the northern North Atlantic, and therefore reduced the strength of the THC, making it less thermally dominant. Once weak enough to have crossed the instability threshold, the THC was destabilized, which lead to the observed strong past THC variability. Note that this mechanism for strong THC and climate variability does not require postulating a large amplitude external atmospheric random forcing but is internal to the ocean and self-sustained. A different mechanism for large amplitude THC variability was proposed by Weaver and Hughes (1994), where larger amplitude atmospheric forcing resulted in the oceanic circulation hopping between different quasi-stable steady states.

In the above discussion and throughout this paper we consider a given climate state to be stable if, when used to initialize the coupled model, it results in a small amplitude variability around this initial state, similar to the weak THC variability of the present day and during the Holocene. On the other hand, a climate state is considered *spontaneously unstable* if, when used to initialize the coupled model, it results in the coupled model experiencing a sizeable drift of the THC away from the initial state, or in large amplitude THC oscillations. A THC instability may thus result in an increase as well as a decrease, collapse, or in strong oscillations of the THC (Toggweiler et al. 1996) that resemble some of the strong climate variability seen during unstable past climates (GRIP 1993; Dansgaard et al. 1993). Because of the emphasis on a THC collapse in many previous studies, THC instability was often identified with a THC collapse. In fact, a THC instability may clearly also result in strong oscillations as well as an increase (Toggweiler et al. 1996). Note that the definition of THC stability used here is consistent with the usual hydrodynamics definition of instability, which implies only a change of the initial state, and no statement regarding the eventual fate of the mean state of the unstable system.

In addition to the above spontaneous instability, in which the only perturbation to the steady state may result from natural weather noise in the atmospheric model, one may introduce external forcing into the coupled model. One example for such external forcing, which we will use below following Manabe and Stouffer (1995), is the addition of a freshwater flux into the North Atlantic sinking area, representing glacier melt, etc. Adding such forcing amounts to a "forced" instability experiment rather than to an examination of the spontaneous instability of the model's climate. Note that some of these runs that are initialized with an unstable ocean-only state drift away and eventually stabilize on a different solution, so it is important to note that our instability definition refers to that of the initial climate state rather than to the entire coupled model run.

There have been quite a few studies in recent years that explored the difficulties with the ocean-only mixed boundary conditions formulation. We know now that a too short restoring time for the sea surface temperature (SST) in this formulation, as well as a restoring time that is not scale selective, could cause significant artifacts in model results obtained under these boundary conditions (Zhang et al. 1993; Mikolajewicz and Maier-Reimer 1994; Power and Kleeman 1994; Rahmstorf and Willebrand 1995). A detailed quantitative analysis of the atmospheric feedback using an atmospheric GCM and a comparison with the commonly used parameters in ocean-only models under simplified boundary conditions has been carried out by Rivin and Tziperman (1997b).

Given these difficulties with mixed boundary conditions and with ocean-only models of the THC, one wonders if the conclusion of Tziperman et al. (1994) that their realistic geometry global ocean-only model under mixed boundary conditions is near a stability transition threshold would hold under a more realistic atmospheric representation. The hypothesis that the present-day THC may be close to a threshold below which it is spontaneously unstable is of a quantitative character. To examine it requires, therefore, a model that provides as good a representation of the atmospheric feedbacks as is available at present. In other words, a realistic-geometry coupled ocean–atmosphere–ice general circulation model is needed. This is precisely the purpose of the present work. We use the GFDL coupled ocean–atmosphere–ice general circulation model to examine the spontaneous stability of a range of strong (thermally dominant) and weak (less thermally dominant) THC circulations. The primitive equations, realistic-geometry, low-resolution coupled ocean–atmosphere–ice GCM used here has been used for many studies of greenhouse scenarios (Manabe and Stouffer 1993), climate variability (Delworth et al. 1993), multiple climate equilibria (Manabe and Stouffer 1988), and more. While of a coarse resolution and suffering numerous known deficiencies, this model still provides the most accurate representation of the global climate system that is practical to run for the length of integrations required for this study, totaling together thousands of model years.

While the dependence of the stability of the THC on the mean state of the THC is the main point of this work, the set of coupled model runs described and analyzed in this paper contains a variety of additional interesting issues regarding the stability and variability of the THC. This paper thus extends the brief report of Tziperman (1997) by providing a more detailed analysis of the instability mechanisms and the role of flux adjustments, by examining the response to instability forced by freshwater input, by providing a more detailed comparison with other coupled modeling results, and by analyzing the mechanisms of variability in both Southern and Northern Hemispheres.

In the following sections we describe the procedure used for running the coupled model (section 2), describe the results demonstrating the instability of a weak, less thermally dominant THC (section 3), discuss some interesting THC oscillations

that are seen in the North Atlantic and Southern Ocean model sectors during the unstable runs (section 4), and conclude in section 5.

2. Coupled model runs

In order to examine the *spontaneous* instability of the model THC, we need to initialize the coupled model with steady-state solutions of the ocean-only and atmosphere-only submodels that are balanced with each other. The different initial states for our model experiments also need to vary in the magnitude of the freshwater forcing and consequently in the strength of the meridional salinity gradient and of the North Atlantic THC. For this purpose, we use the same initialization procedure analyzed using a four-box model as well as using an ocean-only GCM in Tziperman et al. (1994). There, the models were run using restoring boundary conditions, switching to "mixed" boundary conditions after diagnosing the freshwater flux. To repeat the same initialization procedure used in the box model and ocean-only GCM for the coupled model, we follow the normal procedure for initializing coupled ocean–atmosphere models in general, and the GFDL coupled model in particular (Manabe et al. 1991), with a single deviation from this procedure, as follows.

First, the atmosphere-only model is run with a specified observed SST to a statistical steady state, and the monthly averages of the air–sea heat and freshwater fluxes are calculated. Denote these averages by $H^{\text{atmosphere}}(x, y, m)$ and $E - P^{\text{atmosphere}}(x, y, m)$, where x, y are the longitude and latitude and m stands for the month. Second, the ocean-only model is run, forced by the atmospheric air–sea fluxes calculated during the first stage and with additional restoring to observed SST and sea surface salinity. The total air–sea fluxes felt by the ocean model during this second phase are therefore

$$\begin{aligned} H^{\text{ocean}}(x, y, t) &= H^{\text{atmosphere}}(x, y, m) + C_T(\text{SST}^{\text{data}}(x, y, m) - \text{SST}^{\text{model}}(x, y, t)) \\ E - P^{\text{ocean}}(x, y, t) &= E - P^{\text{atmosphere}}(x, y, m) + C_S(\text{SSS}^{\text{data}}(x, y, m) - \text{SSS}^{\text{model}}(x, y, t)), \end{aligned} \tag{1}$$

(Click the equation graphic to enlarge/reduce size)

where C_T and C_S are the temperature and salinity restoring coefficients, set to 50 day^{-1} for a 50-m upper-ocean layer thickness.

When initialized with ocean-only and atmosphere-only steady states obtained by running the two models separately, the coupled model rapidly drifts from its initial conditions. Thus, the air–sea fluxes of heat and freshwater are supplemented at every time step by flux adjustments that depend on geographical location and month, but have no interannual variations. The adjustments are calculated by monthly averaging the temperature and salinity restoring terms in (1) over 500 years after the ocean-only integration reaches a steady state. The flux adjustments thus do not depend on the model state during the coupled model integration. Finally, the ocean and atmosphere models are coupled and are run together, exchanging fluxes and SST daily.

Our experiments follow this procedure except that during the ocean-only integration, each experiment uses a sea surface salinity [$\text{SSS}^{\text{data}}(x, y, m)$ in (1)] that is modified in the North Atlantic water-mass formation area. The modifications to the observed North Atlantic salinity are zonally uniform and a function of latitude θ only, and are of the form

$$S'(\theta) = \Delta S \exp[-(\theta - 55.5°)^2/4.4°^2], \tag{2}$$

where for our different model experiments, ΔS takes the values (−1, −0.5, −0.375, −0.25, 0, +0.5 ppt). Figure 1a shows the observed (Levitus 1982) zonally averaged North Atlantic surface salinity and the zonally averaged modified surface salinity fields. We denote each initial state of the coupled model by the amplitude ΔS of the restoring salinity perturbation used to obtain it in the ocean-only model runs (Fig. 1a). The resulting zonally averaged flux adjustments for all experiments are shown in Figs. 1b,c. The modifications to the SSS to which the surface model salinity is restored make it fresher or saltier in the North Atlantic water mass formation area (Fig. 1a), leading to a correspondingly weaker (less thermally dominant) or stronger (more thermally dominant) steady-state ocean-only North Atlantic THC (dash line in Fig. 1d). The restoring of the model surface temperature and salinity to the specified SST and SSS fields during the ocean-only integration allows us to obtain various ocean-only steady states, some of which may be unstable in the coupled model, where the surface fields are not restored and are free to evolve.

Flux adjustments have gotten a somewhat poor reputation, being an artificial procedure that is meant to compensate for ocean and atmosphere model errors (Neelin and Dijkstra 1995; Marotzke and Stone 1995). They cannot be avoided using presently available models, however, if one is interested in studying climates that are present-day-like because without them the coupled model drifts very rapidly to a very different new steady state. As we are interested in examining the stability of present-day climate and near-by climates, we must use these adjustments. Flux adjustments are used in the present experiments as a parameter allowing one to obtain the different balanced initial THC states (Fig. 1d). This is somewhat similar to changing the freshwater forcing in an ocean-only model (Weaver et al. 1991; Rahmstorf 1995; Tziperman et al. 1994). The issue of to what extent they affect our results is of a major importance, obviously, and will be discussed in detail in section 3d.

Once we couple the ocean and atmosphere models and given the model initialization procedure, including the flux adjustments that keep the ocean and atmosphere models balanced, we expect the model to remain near the initial climate, as

it does for the control run (Manabe and Stouffer 1993; Delworth et al. 1993). We expect the model to drift away from the initial climate only if the initial state is not stable. Unless stated explicitly otherwise, we do not add any initial perturbations to the ocean-only and atmosphere-only solutions or to the air–sea fluxes. Thus, initial states that are not maintained during the coupled integration are spontaneously unstable. We define stable initial states as those that remain within two standard deviations during the coupled model run (where the standard deviation is that of the stable control run), although the results are not sensitive to this particular choice of two standard deviations. We now proceed to describe the results of the coupled model integrations.

3. Instability of a weak, less thermally dominant THC

Our coupled model experiments, and their description in this section, are divided into two categories. First, there are the model experiments carried without any added external forcing, which thus study the spontaneous stability of more and less thermally dominated THC states (section 3a). Second, there are the model experiments that study the forced instability of the THC under the influence of additional external freshwater forcing added during the integration of the coupled model (section 3b). The results are then discussed in section 3c and the issue of flux adjustments is analyzed in detail in section 3d.

a. Spontaneous instability runs

When ocean-only models in previous studies were initialized by restoring their SST and SSS and were then switched to some simple atmospheric representation, these models invariantly found that less thermally dominant and thus weaker THC initial states are unstable, leading to a switch to a different THC state. The existence of such a stability threshold between strong and stable THC and weak and unstable THC can be understood by analyzing a simple meridional box model as done in Tziperman et al. (1994). It is remarkable how consistent the qualitative behavior is of the realistic-geometry coupled ocean–atmosphere GCM used here with that of these simpler ocean-only models and, in particular, with the box model prediction. Namely, the more thermally dominant initial states with a strong initial THC (runs [$\Delta S = -0.25$], control [$\Delta S = 0$] and [$\Delta S = +0.5$], Figs. 2a–c) are stable. The coupled model runs starting at these initial conditions remain mostly within two standard deviations from the initial state, where the standard deviation is defined from the control run's THC variability that roughly represents a small, Holocene-like, stable variability. On the other hand, the less thermally dominant and thus weaker initial THC states (runs [$\Delta S = -1$], [$\Delta S = -0.5$], [$\Delta S = -0.375$], Figs. 2d–f) are spontaneously unstable within the coupled model, as the THC drift from the initial conditions in these runs clearly exceeds two standard deviations. We call this a spontaneous instability because no external perturbation is applied to the coupled model once the coupled model integration begins.

The threshold between stable and spontaneously unstable initial states in the coupled model lies between 14 Sv (Sv $\equiv 10^6$ m^3 s^{-1}) (stable state [$\Delta S = -0.25$], Fig. 2c) and 13 Sv (unstable state [$\Delta S = -0.375$], Fig. 2d). The unstable initial states are thus characterized by a North Atlantic THC that is about 25% or more weaker than the stable control-run initial state. This is not a negligible difference, especially with the present-day variability being significantly smaller than this difference, possibly of the order of 10% of the mean North Atlantic THC.

To more fully characterize the difference between the stable and unstable initial states, consider the north–south sections of temperature, salinity, and meridional overturning (Fig. 3). The formation of North Atlantic Deep Water in the unstable ocean-only steady state [$\Delta S = -0.375$] is weaker and shallower than in the stable control run. As a result, the deep North Atlantic basin is filled with more Antarctic Bottom Water, which is fresher and colder. The differences between the stable initial condition in the control run [$\Delta S = 0$] and unstable initial state [$\Delta S = -0.375$] are clearly significant, yet are relatively modest when compared to climate variability on a glacial–interglacial timescale. We therefore feel it is justified to state that the model THC representing the present-day climate is "not far" from an instability threshold in this coupled model. Even such a statement that is restricted to the stability behavior of the model only and not of the real climate system is, of course, somewhat uncertain. One major source of such uncertainty is the issue of flux adjustment, which is treated in detail in section 3d.

Note that none of our runs displays a stable, present-day-like small amplitude THC variability within the unstable THC regime of below 14 Sv. When forced by 2 × CO_2 (Manabe and Stouffer 1993) and by freshwater input into the North Atlantic Ocean (Manabe and Stouffer 1995), the coupled model THC has recovered within 150–200 years. This defines an adjustment time by which we expect the THC to respond to various initial coupling shocks or perturbations. The inability of the THC in our runs to recover to its original weak initial states within the quite longer 500 years of integration seen in Fig. 2 indicates that the weak THC initial states may indeed be considered to be in an unstable THC regime, at which the model is unlikely to be found with small amplitude variability. We emphasize that the unstable states are defined not only as being of a weak THC, but as weak THC states that are affected by a large meridional surface salinity gradient that makes the THC less thermally dominant.

1) INSTABILITY MECHANISM

Two different main THC instability mechanisms have been discussed in the literature. The first is the linear advective salinity feedback of Walin (1985; see also Marotzke et al. 1988; Tziperman et al. 1994). The second mechanism involves changes to the inherently nonlinear process of convection, first seen by Bryan (1986) in his "polar halocline catastrophe" that caused the shutoff of convection in the North Atlantic water-mass formation area. Later studies identified convective mechanisms in different cases that do not necessarily lead to a shutoff of the THC. The instability mechanism leading to the climate drift in our unstable coupled model runs seems to be composed of two distinct phases of instability, starting with a

linear advective salinity feedback and ending as a nonlinear convective instability as that of Lenderink and Haarsma (1994). As will be seen below, a detailed understanding of the instability mechanism is crucial for assessing the sensitivity of our results to factors such as the flux adjustments, hence the detailed analysis that follows.

As an example of the instability mechanism in our runs, we concentrate on the instability of the unstable run [ΔS = −0.375], which is the unstable run nearest to the stable control run. During the first 50 years or so of this run, an advective linear salinity feedback (Walin 1985; Marotzke et al. 1988; Tziperman et al. 1994) leads to an increase of salinity and density in the water-mass formation area and thus to an increase in the formation rate of deep water southeast of Greenland. As seen in Figs. 2a–f, the instability leads to an increase of the THC in our spontaneously unstable runs. Figure 4 shows the rate of deep-water formation in the three sites of deep-water formation in this model: southeast of Greenland, in the Norwegian Sea, and in the Labrador Sea. During the first 50 years of the coupled model run, there is a slow and gradual increase in the water mass formation southeast of Greenland from 10 to around 13 Sv, while the two other sites are not affected. This increase is governed by the linear salinity feedback, as can be seen in Fig. 5, which shows that together with the increase in the rate of sinking (and hence of the THC), there is also an increase in the sinking area density due to the advection of high salinity water from the midlatitudes (solid curve in Fig. 5). The advected water is also warmer and this tends to reduce the density of the sinking area surface water (dash curve in Fig. 5). However, due to the cooling of this warm water by the atmosphere in the northern North Atlantic, the negative feedback to the temperature increase is weaker and is overcome by the salinity feedback, exactly as proposed by the simple box model analysis of Walin (1985).

Runs with stronger initial THC and weaker initial meridional salinity gradient are stable with respect to this linear instability stage for the following reason (Walin 1985; Marotzke et al. 1988; Tziperman et al. 1994). The salinity equation for a meridional box model that shows this linear instability, linearized about the initial steady state, has two main terms: first, a destabilizing term that represents the advection of the mean salinity gradient by the perturbation circulation and, second, a stabilizing term that corresponds to the advection of the perturbation salinity gradient by the mean THC. A strong initial THC makes the second stabilizing term dominant and rapidly advects away any developing salinity perturbations. Also, the destabilizing term is smaller in the stable runs because their initial meridional salinity gradient is smaller.

Initially, in the ocean-only steady states, the Labrador Sea does not form deep water and penetrative deep convection does not occur there. At the second stage of the instability, however, the increased sinking southeast of Greenland modifies the properties of the deep water inside the neighboring Labrador Sea (Fig. 6), starting deep-water formation and convective mixing there as well, and leading to a further increase in the THC strength. The stratification of the Labrador Sea prior to the onset of convection is of cold freshwater over warm and salty water. When an initial convective mixing event brings the warm salty water to the surface, this water is quickly cooled by the cold atmosphere above the Labrador Sea, leaving at the surface a cold and salty, and hence dense, water mass. This dense water enforces the initial convection and leads to self-sustained convection (Lenderink and Haarsma 1994).

The increase in THC seen in our unstable runs is clearly due to the presence of the inactive Labrador Sea water-mass formation area at the same latitude of the active formation regions in the Norwegian Sea and southeast of Greenland. Such an increase clearly cannot be seen in simpler 2D models (Stommel 1961; Marotzke et al. 1988; Tziperman et al. 1994) that, once unstable, often lead to a collapsed or reversed THC. These models, being 2D, cannot support an inactive deep-water formation site at the same latitude of the active formation sites. The initial linear instability mechanism in our runs, which can be represented by a simple 2D meridional model, can explain both a weakening or a strengthening of the THC as it is symmetric with respect to the sign of the initial perturbation. However, such simple 2D models cannot represent our second instability phase, which occurs due to the existence of an inactive water-mass formation site at the same latitude of an active source—obviously an effect which is three-dimensional (Rahmstorf 1995). Now, the eventual fate of the THC in simple 2D models is to evolve toward available stable THC states, which are often in these models a reversed, salinity-dominated THC. In our coupled GCM, on the other hand, the eventual fate of the unstable THC is determined by the nonlinear effects and by the activation of the Labrador Sea water-mass formation area. Because of this difference, one may not expect the simple 2D Stommel-like box model to reproduce the final state of the unstable THC, but only explain the initial instability mechanism. It is clearly important to understand the limits of these box models when using them to interpret the coupled model results (Toggweiler et al. 1996).

There is one point of uncertainty regarding the comparison between the box model and coupled GCM stability behavior. One might naively expect with a linear instability mechanism that an unstable THC should decrease about half of the time in response to a random atmospheric perturbation. However, the initial trend of all spontaneously unstable runs here is of an increased THC. It is possible that this set of coupled model runs is just a too small sample, and it is also possible that there is still something we simply do not understand here.

b. Forced instability runs

As noted, all the unstable initial states in the experiments described above (Figs. 2d–f) lead to an increase of the THC. One wonders if the unstable initial states found in the previous sections can also lead to a THC collapse under somewhat different circumstances. In order to examine this issue, we have repeated the experiment of Manabe and Stouffer (1995), injecting 1 Sv of freshwater into the North Atlantic deep-water formation area for 10 years, beginning from our [ΔS = +0.5] initial state, which was found to be spontaneously unstable within the coupled model.

Figure 2g shows the comparison of this forced instability run with the original run of Manabe and Stouffer (1995). The THC reduces in both cases to about 6 Sv after 10 years of freshwater input, although their initial states are different. Once the freshwater input is stopped, the Manabe and Stouffer run recovers to the original strength of the THC, while our

run leads to a collapsed THC that does not recover until the end of the integration, at year 300. We conclude that our spontaneously unstable initial states are also more prone to a forced THC collapse than stronger THC runs such as the control run (Manabe and Stouffer 1995). It seems that whether the instability leads to an increased THC or a collapsed THC may simply depend on type of perturbation to the initial steady state.

Note that while the THC amplitude at year 10 is the same in both model runs, the further evolution is very different. Clearly, there is more to the climate state than just the THC amplitude. The two runs start from different steady-state solutions for the temperature, salinity, and velocity fields. After 10 years of freshwater flux, these fields are still significantly different (Fig. 7) and thus the different stability behavior of the two runs after year 10. This demonstrates again that an unstable state is characterized not only by a weak THC, but also in being less thermally dominant, which is expressed in the structure of the 3D temperature and salinity fields as well.

c. Discussion

1) RELATION TO OTHER COUPLED RUNS

We briefly discuss other relevant coupled model results that can shed additional light on our results here. The ×2 and ×4 increased CO_2 runs of Manabe and Stouffer (1993) resulted in a temporarily decreased and collapsed THC, correspondingly. According to the terminology used here, one would expect that these runs should have crossed the instability threshold. However, no abrupt behavior and no strong THC oscillations are seen in these runs as they cross the threshold of 14 Sv for the North Atlantic THC. Explaining this seeming contradiction will help place our results here in a more general perspective.

The major difference between the runs shown here and the greenhouse runs is that in the later ones, the THC was reduced on a timescale of a few tens of years so that the temperature and salinity fields may have not yet adjusted to the evolving THC. Therefore, although the THC seemed to have crossed the instability threshold, the actual ocean state, which is composed of both the THC and the temperature and salinity fields, did not evolve beyond the stability threshold. This is reminiscent of the comparison (Fig. 7) between the two forced instability runs of Fig. 2g, which seem to have the same state at year 10, yet one recovers and one collapses. This difference in the forced instability case was explained in section 3b in a similar fashion.

Thus, had the CO_2 concentration been made to change very slowly in the greenhouse runs, then one could argue that at any given time the model is in equilibrium and so the instability analysis regarding a threshold applies. For the relatively rapid CO_2 changes applied in the greenhouse runs, however, the climate system may pass through the threshold without having more variability, as the threshold was passed in a transient phase rather than existing at an equilibrium phase.

This interpretation of the greenhouse runs is further strengthened by comparing our results to the experiments of Manabe and Stouffer (1995) in which they subjected the northern North Atlantic Ocean to a very slow freshwater input. In this experiment, the slow freshwater injection caused the THC to decrease slowly, slower than in the greenhouse runs, so that the temperature and salinity fields had time to adjust to the changing amplitude of the THC. In addition, the experiment is forced purely by freshwater forcing so that (unlike in the greenhouse experiments) the freshwater forcing is not balanced by changing thermal forcing and can make the THC less thermally dominated, as required in our stability crossing scenario. Examining the results of this slow freshwater injection model experiment, one can indeed see that as the THC reduces to near the stability transition THC value in this model, of about 14 Sv, the amplitude of variability increases, according to our proposed hypothesis of the existence of a stability threshold. This strengthens the case for the existence of a stability transition point in this coupled model.

d. Flux adjustments

We found that initial states of the THC that are less thermally dominant are unstable within the coupled model and resulted in a climate drift. The unstable weak THC initial states were also different from the stable ones in the amplitudes of their flux adjustment fields. Could the different stability properties be related to the different flux adjustment fields rather than to the different initial state of the THC and large-scale temperature and salinity fields? The effort to address this important question is the purpose of this subsection.

We have identified four lines of argument that, we think, lead to the conclusion that the flux adjustment fields in the unstable runs are not the reason for the unstable behavior. These arguments strengthen, in our opinion, the conclusion that the instability results from the weakness of the initial THC and the large initial meridional salinity gradient.

Figures 1b,c show the zonally averaged and annually averaged flux adjustments in the North Atlantic Ocean. Note that the curves representing the stable run [$\Delta S = -0.25$] (thin dash line) and the unstable run [$\Delta S = -0.375$] (thick solid line) are nearly indistinguishable. Thus, the first argument is that the difference in the adjustment fields between the stable and unstable runs is not large, reducing the possibility that this difference is the source of different stability behavior. This argument by itself clearly does not *eliminate* the possibility that the model stability regime is determined by the flux adjustments and that the threshold crossed in the unstable experiments is of a critical flux adjustment amplitude rather than a critical THC strength amplitude.

A second argument concerns the initial instability mechanism, which was shown in section 3a(1) to be the linear salinity

feedback of Walin (1985). It is simple to see that the linear instability mechanism during the first 40–50 years of our coupled runs does not depend on the steady component of the air–sea fluxes: the mean fluxes simply do not appear in the linearized equations that describe this instability mechanism (Walin 1985; Marotzke et al. 1988; Tziperman et al. 1994). The flux adjustments are not a function of the model state and have no interannual variations, and thus, from the point of view of the ocean model, are part of the mean air–sea fluxes. Therefore the linear instability mechanism, which dominates the first 40–50 years of our unstable runs, is not likely to depend on the flux adjustment fields. Note that in the linear instability mechanism the mean THC and the mean meridional salinity gradient do influence the stability behavior (Tziperman et al. 1994). Both of these mean quantities are maintained to some degree by the mean fluxes, including the flux adjustments, so there is an indirect link between the adjustments and the stability behavior. However, the dependence of the stability behavior on the mean states of the THC and meridional salinity gradient is precisely what we want to investigate in this paper.

As a third specific argument concerning the role of flux adjustments in the THC instability, we examine the contribution of the heat and freshwater flux adjustments to the buoyancy flux in the Labrador Sea. Figure 8 shows that this buoyancy flux is larger in the unstable experiments than in the stable ones. Therefore the buoyancy flux due to the flux adjustments tries, in fact, to further increase the buoyancy of the surface water in the Labrador Sea in the unstable runs. This influence acts to prevent convective mixing in the Labrador Sea and thus opposes the second, convective, phase of the instability mechanism [section 3a(1)]. Thus not only do the adjustments not cause the THC instability but they try, in fact, to oppose it. A more detailed examination of the separate buoyancy fluxes due to the heat and freshwater adjustments shows that the freshwater component of the adjustments acts to weaken the stratification of the Labrador Sea, yet is dominated by the stronger difference in the heat flux adjustments that act to strengthen the stratification (not shown).

Note that some of the difficulty with using flux adjustments arises when they are used for climate states too different from those for which they were calculated. When this occurs, there is a possible breakdown of the implicit linearization assumption involved in keeping the adjustments constant when the climate changes (Neelin and Dijkstra 1995). Our stability analysis concerns the initial behavior of the coupled model near the initial state obtained from the steady state ocean-only solutions. Thus, our fourth argument is: during the initial phase of the instability, the model states (still) do not deviate significantly from the ones for which the flux adjustments were calculated. So this specific problem of using flux adjustments away from the state for which they were calculated certainly does not affect our climate stability analysis.

Having gone through these detailed and technical arguments, we feel that the strongest arguments for the credibility of these coupled model runs are quite simpler, and there are two of them. First, we feel that a coupled ocean–atmosphere–ice GCM such as used here, with primitive equation ocean and atmosphere models, is likely to be more reliable than an ocean-only GCM with highly simplified air–sea heat flux parameterization, with fixed freshwater forcing, and with no weather noise. Thus, we tend to trust these model runs more than we do our previous ocean-only results under mixed boundary conditions (Tziperman et al. 1994), which showed the same qualitative results concerning the instability of a weak, less thermally dominant THC. Second, the stability behavior of the coupled model used here is, in fact, in qualitative agreement with a long line of ocean-only and idealized coupled model results, so we feel there is a good case for the hypothesis supported by our present results that a weak and less thermally dominant THC is unstable. There is also some evidence from previous coupled model studies indicating that the results of such models may not necessarily be sensitive to the flux adjustment procedure (Schiller et al. 1997).

An issue closely related to the flux adjustment issue is the possibility that the initial drifts of the model THC are simply due to the failure of the initialization procedure to keep the model at the initial state, rather than due to some physical instability as proposed in this paper. We feel that our analysis and resulting understanding of the initial instability mechanism reduces the possibility of such a failure of the initialization procedure as a cause of the model behavior analyzed here. The past success of the same initialization procedure, used in several different coupled models, also indicates that this may not be a problem. However, given the model complexity, there is probably no way to completely rule out such a possibility.

Finally, we acknowledge that the issue of the quantitative proximity of present-day climate to an instability threshold remains open because of uncertainties resulting from various model limitations including the use of the flux adjustments. There are some indications (e.g., Fanning and Weaver 1997) that model sensitivity might depend on the use of flux adjustments in certain cases. In spite of these uncertainties, we feel that the present results do suggest that the distance of present-day THC to an instability threshold is not large and that this is an issue worth pursuing with better and more realistic coupled models as these become available in the future.

4. THC oscillations

Past the initial instability, the runs beginning from the unstable initial states are characterized by strong North Atlantic THC oscillations (Figs. 2d–f). This seems to be the first time that such large amplitude oscillations are seen in a coupled general circulation model, following numerous ocean-only model studies that observed such large amplitude THC variability (McWilliams 1996) and coupled model runs that displayed a smaller amplitude variability (Delworth et al. 1993; Griffies and Tziperman 1995). This interesting large amplitude variability in both the North Atlantic and the Southern Ocean is now analyzed in some detail.

a. North Atlantic THC oscillations

Some of the coupled model runs initialized with unstable initial states display a strong North Atlantic THC variability (e.g., Fig. 2f). The mechanism of this variability seems to be inherently thermohaline, meaning that both temperature and

salinity play a crucial role, unlike the oscillations of Greatbatch and Peterson (1996), for example, where the oscillation can exist in a single component ocean (e.g., with a variable temperature and constant salinity). Figure 9 shows time series of temperature and salinity in the North Atlantic sinking area for the unstable run [$\Delta S = -1$]. Note the small phase lag between the salinity that is leading the temperature in the oscillation, as is characteristic of thermohaline oscillations (Griffies and Tziperman 1995). Although the THC variability in some of these runs (most notably [$\Delta S = -1$] in Fig. 2f) is very strong and possibly self-sustained (see discussion in section 4b), the mechanism seems to be basically the following 2D mechanism of Griffies and Tziperman (1995).

The oscillation starts with the THC initially increasing due to the advective salinity feedback mechanism, discussed in section 3a(1), which leads to an increased salinity in the North Atlantic sinking area and thus to an increased density there and further increased sinking rate and THC. This increase is eventually stabilized by the increase in water temperature in the sinking area due to warm water advection from midlatitudes by the enhanced THC. Once the temperature increase dominates the salinity increase, the density of the sinking area begins to decrease, and this results in a decreasing phase of the oscillation. Finally, the time lag between the sinking area temperature and salinity and those at midlatitudes provides the phase change (zero crossing) mechanism for the oscillation, and the oscillation proceeds in a symmetric way with opposite signs. A more detailed description of the mechanism may be found in Griffies and Tziperman (1995) or in Rivin and Tziperman (1997a).

b. Present-day and past North Atlantic THC variability: Damped versus self-sustained

Griffies and Tziperman (1995) analyzed the THC variability in the coupled GFDL model run of Delworth et al. (1993) and proposed that it might be due to a random excitation of a damped oscillatory ocean THC mode. That is, a single perturbation to a steady-state THC (due to some random fluctuation of the air–sea fluxes, for example) should result in a decaying (damped) THC oscillation of a period determined by the ocean dynamics. A continuous random excitation (e.g., by atmospheric weather) can thus lead to a continuous random THC variability with a preferred timescale equal to the period of the damped oscillatory ocean mode.

For a large enough freshwater forcing (and therefore a large meridional salinity gradient and weak THC) the damped oscillatory mode becomes unstable and results in exponentially growing THC oscillations. Rivin and Tziperman (1997a) have shown that an appropriate nonlinearity in the ocean dynamics can saturate the growing oscillations, resulting in periodic, nonlinear self-sustained oscillations even without an external excitation. The oscillatory mechanism may be exactly that of Griffies and Tziperman (1995). Moreover, it was shown that in the presence of atmospheric random forcing, and given a limited-length time series, it is most difficult to distinguish between a self-sustained THC variability that is only randomized by the atmospheric noise and a damped variability whose very existence is due to the external excitation. While a large amplitude variability is likely to involve a nonlinear mechanism, a nonlinear self-sustained oscillation may also be of small amplitude (e.g., Chen and Ghil 1995).

The large amplitude of the THC variability in some of the unstable runs of our coupled model here raises the possibility that these oscillations may be self-sustained rather than noise driven. Given the short time series, it is difficult of course to verify this assumption. However, the THC of run [$\Delta S = -1$] (Fig. 2f) in particular lies almost entirely in the unstable THC regime below 14 Sv, and its strong variability occurs in the presence of a large meridional surface salinity gradient. These are precisely the two factors needed to obtain self-sustained THC variability (Rivin and Tziperman 1997a).

We therefore allow ourselves to speculate on the character of past versus present-day THC variability, based on the assumption that a destabilized, weak, less thermally dominant THC state can lead to self-sustained THC variability as seen in Fig. 2f. This assumption may provide a natural explanation for the unstable THC behavior, with large amplitude variability, seen in previous geological periods, both glacial and interglacial, such as the Eemian climate variability [GRIP 1993; although see McManus et al. (1994) for a reinterpretation of the paleo record suggesting that the Eemian may not have been as unstable as suspected initially]. The explanation we advocate here suggests that factors external to the ocean, such as melting of land glaciers or higher precipitation rates in the North Atlantic, have reduced the salinity of the northern North Atlantic and therefore reduced the strength of the THC. Once the meridional salinity gradient is strong enough and the THC is weak enough to have crossed the instability threshold, the THC would be destabilized as in our coupled model runs. This destabilization may have lead to the strong THC variability seen in paleo proxy records. Note that this strong variability does not require postulating a large amplitude external atmospheric random forcing (Weaver and Hughes 1994) but is internal to the ocean and self-sustained. Note also that the THC variability mechanism seen in these runs does not involve a complete switch between “on” and “off” states of the North Atlantic THC. In fact, it does not involve a switch between any two quasi-steady states of the THC altogether (Weaver and Hughes 1994). Rather, the large amplitude oscillations result solely from the inherent instability of a weak THC. Whether the correct explanation of past large-amplitude climate variability involves switches between quasi-steady states or our alternative mechanism cannot be determined with certainty at the moment given the available paleo record. We note also that presently there is no agreement regarding the character of the THC oscillations in the GFDL coupled model, and Weaver and Valcke (1998), for example, have actually suggested that these oscillations are self-sustained, in contrast to the conclusion of Griffies and Tziperman (1995), yet analyzing the same coupled model output.

c. Southern Ocean variability

While our main focus in this paper is the THC stability in the North Atlantic Ocean, some quite interesting behavior is seen in the Southern Ocean model sector as well during the unstable runs. This variability is expressed as an irregular 20-yr oscillation that affects convection, ice cover, temperature, and salinity around the circumpolar current.

Figure 10 shows the salinity and temperature anomalies in the circumpolar region during a 100-yr period. The diagonal features in the figure point to an eastward propagation of temperature and salinity anomalies and provide a clue to the mechanism of variability that results from the advection of low-salinity anomalies by and around the circumpolar current. The timescale of 20 yr seen in Fig. 10 is indeed roughly the advection time around the circumpolar current in this model. As the relatively fresh perturbations are advected around Antarctica, they enhance the stratification where they appear and cause a disruption of convection that typically occurs in the open ocean west of the Ross Sea. The convection occurs in a region where warm and salty deep water is found below the fresh and cold surface water. The disruption of convection by the low-salinity anomalies prevents the import of salty water to the surface by convection and thus further amplify the fresh anomaly, which then proceeds around the circumpolar current. The disruption of convection also affects the SST due to the decreased supply of warm deep water to the surface by convective mixing. This change to the SST, in turn, also affects the sea ice distribution around Antarctica, causing a significant signal related to the salinity oscillation seen in Fig. 10.

This mechanism of Southern Ocean variability is basically the same mechanism as seen by Pierce et al. (1995) in an ocean-only model under mixed boundary conditions. This is another indication that ocean-only models under simplified boundary conditions may capture the essential qualitative behavior that is also seen in fuller coupled models.

5. Conclusions

The main objective of this paper is to examine the relation between the mean state of the thermohaline circulation and its stability, using a realistic-geometry, primitive equations, coupled ocean–atmosphere–ice global general circulation model. Our main finding is that a thermohaline circulation that is 25% weaker and less dominated by thermal forcing than today's ocean is unstable within this coupled GCM. Unstable initial ocean climates may lead in the coupled model to an increase of the THC and/or to strong oscillations. Such unstable states of a weaker and less thermally dominated THC are also more prone to a forced THC collapse due to a temporary freshwater input than the stable initial states that have stronger THC and are more thermally dominated.

The results of the runs shown here limit the range of stable active equilibria of the THC in this model to above about 14 Sv, while the present-day THC in this model is about 18 Sv. We did not find states of the THC that are weaker than 14 Sv, are only weakly thermally dominant (i.e., have a strong salinity forcing through a large meridional salinity gradient), and are at the same time stable with a small, present-day-like variability. A distance of 25% between unstable mean states and present-day oceanic circulation is not large in the sense that the circulation, temperature, and salinity fields of such unstable states seem qualitatively similar to those of present-day. On the other hand, a 25% change required to bring the THC to within the unstable range is significantly larger than the natural variability of this model (Delworth et al. 1993) and of the actual North Atlantic over the past few thousands of years.

We have speculated that the unstable character of a weak THC may suggest a mechanism for the large-amplitude THC variability seen in proxy records for glacial periods and possibly also during the previous interglacial Eemian period (GRIP 1993; Dansgaard et al. 1993; although see McManus et al. 1994 for a reinterpretation of the paleo record, suggesting that the Eemian may not have been as unstable as suspected initially). According to the proposed mechanism, the North Atlantic water-mass formation and thus the THC prior to the Holocene may have operated in a weak and shallow mode (Boyle and Keigwin 1987) that is only weakly thermally dominant and hence unstable and with large variability. Later, the THC may have switched to a stronger and more thermally dominant mode that is therefore stable and with smaller variability, which has lasted throughout the Holocene. The switch from more to less thermally dominant THC might be due to slow-acting external factors such as continental glacier melt and growth. This mechanism for past large-amplitude THC variability is an alternative one to that suggested by Weaver and Hughes (1994) where the large variability is attributed to jumps between two or more quasi-stable modes of the THC, excited by a large amplitude weather noise. Note that in the present mechanism, the variability is internal to the ocean and self-sustained and does not require large amplitude external noise. Whether the correct explanation of past large-amplitude climate variability involves switches between quasi-stable steady states or our alternative mechanism cannot be determined with certainty at the moment given the available paleo record. It is interesting in this context that some observations suggest a recent weakening of the THC (Schlosser et al. 1991). We have also discussed some possible implications of these findings to greenhouse scenarios that predict a decrease of the THC (Manabe and Stouffer 1993).

We have dealt at length with the issue of flux adjustments, trying to verify that the stability behavior of the model is not an artifact due to the artificial procedure of flux adjustment. Apart from the detailed technical arguments in this regard, we have indicated that a coupled ocean–atmosphere–ice general circulation model with a detailed and active atmospheric component and with weather noise explicitly modeled is likely to be more reliable than ocean-only models under simplified boundary conditions (e.g., "mixed" boundary conditions). We therefore believe that the results of this study may be a step forward as compared with previous ocean-only studies of the same issue. While the location of the stable THC range found here is, no doubt, model dependent, we expect the qualitative dependence of the model stability and variability behavior on the mean THC strength to carry over to the real world. Finally, a better knowledge of past THC magnitude (Yu et al. 1996) and a more precise determination of the unstable THC regime and its distance to present-day climate is clearly needed, using better coupled ocean–atmosphere models as they become available in the future.

Acknowledgments

I am grateful to S. Manabe and R. Stouffer for allowing the use of their coupled model, and to S. Griffies and R. Stouffer without whose help the completion of this work would not be possible. Thanks to the Atmospheric and Oceanic Sciences

program at Princeton and to GFDL for their hospitality during parts of this work. Many thanks to I. Held, S. Manabe, J. Marotzke, U. Mikolajewicz, R. Toggweiler, and A. Weaver for most useful comments. Thanks also to A. Weaver for his thoughtful reviews. This work is partially funded by the Israeli Academy of Sciences.

REFERENCES

Boyle, E. A., and L. Keigwin, 1987: North Atlantic thermohaline circulation during the past 20 000 years linked to high-latitude surface temperature. *Nature*, **330,** 35–40..

Brass, G. W., J. R. Southam, and W. H. Peterson, 1982: Warm saline bottom water in the ancient ocean. *Nature*, **296,** 620–623..

Bryan, F., 1986: High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature*, **323,** 301–304..

Chen, F., and M. Ghil, 1995: Interdecadal variability of the thermohaline circulation and high-latitude surface fluxes. *J. Phys. Oceanogr.*, **25,** 2547–2568.. Find this article online

Dansgaard, W., and Coauthors, 1993: Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, **364,** 218–220..

Delworth, T., S. Manabe, and R. J. Stouffer, 1993: Interdecadal variations of the thermohaline circulation in a coupled ocean–atmosphere model. *J. Climate*, **6,** 1993–2011.. Find this article online

Drijfhout, S., C. Heinze, M. Latif, and R. Maier-Reimer, 1996: Mean circulation and internal variability in an ocean primitive equation model. *J. Phys. Oceanogr.*, **26,** 559–580.. Find this article online

Fanning, A. F., and A. J. Weaver, 1997: On the role of flux adjustments in an idealized coupled climate model. *Climate Dyn.*, **13,** 691–701..

Greatbatch, R. J., and K. A. Peterson, 1996: Interdecadal variability and oceanic thermohaline adjustment. *J. Geophys. Res.*, **101,** 20 467–20 482..

Griffies, S., and E. Tziperman, 1995: A linear thermohaline oscillator driven by stochastic atmospheric forcing. *J. Climate*, **8,** 2440–2453.. Find this article online

GRIP, 1993: Climate instability during the last interglacial period recorded in the GRIP ice core. *Nature*, **364,** 203–207..

Lenderink, G., and R. J. Haarsma, 1994: Variability and multiple equilibria of the thermohaline circulation associated with deep water formation. *J. Phys. Oceanogr.*, **24,** 1480–1493.. Find this article online

Levitus, S., 1982: *Climatological Atlas of the World Ocean.* NOAA Prof. Paper No. 13, 173 pp..

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

— , and — , 1993: Century-scale effects of increased atmospheric CO_2 on the ocean–atmosphere system. *Nature*, **364,** 215–218..

— , and — , 1995: Simulation of abrupt climate change induced by freshwater input to the North Atlantic ocean. *Nature*, **378,** 165–167..

— , M. J. Spelman, and K. Bryan, 1991: Transient response of a coupled ocean–atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. *J. Climate*, **4,** 785–818.. Find this article online

Marotzke, J., and P. H. Stone, 1995: Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. *J. Phys. Oceanogr.*, **25,** 1350–1364.. Find this article online

— , P. Welander, and J. Willebrand, 1988: Instability and multiple steady states in a meridional-plane model of the thermohaline circulation. *Tellus*, **40A,** 162–172..

McManus, J. F., G. C. Bond, W. S. Broecker, S. Johnsen, L. Labeyrie, and S. Higgins, 1994: High-Resolution Climate Records from the North-Atlantic during the Last Interglacial. *Nature*, **371,** 326–329..

McWilliams, J. C., 1996: Modeling the oceanic general circulation. *Ann. Rev. Fluid. Mech.*, **28,** 215–48..

Mikolajewicz, U., and E. Maier-Reimer, 1994: Mixed boundary conditions in ocean general circulation models and their influence on the stability of the model's conveyor belt. *J. Geophys. Res.*, **99,** 22 633–22 644..

Neelin, J. D., and H. A. Dijkstra, 1995: Ocean–atmosphere interaction and the tropical climatology. Part I: The dangers of flux adjustment. *J. Climate*, **8,** 1325–1342.. Find this article online

Pierce, D. W., T. P. Barnett, and U. Mikolajewicz, 1995: Competing roles of heat and fresh-water flux in forcing thermohaline oscillations. *J. Phys. Oceanogr.*, **25,** 2046–2064.. Find this article online

Power, S. B., and R. Kleeman, 1994: Surface heat flux parameterization and the response of ocean general circulation models to high-latitude freshening. *Tellus*, **46,** 86–95..

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

— ,and J. Willebrand, 1995: The role of temperature feedback in stabilizing the thermohaline circulation. *J. Phys. Oceanogr.*, **25,** 787–805.. Find this article online

— ,J. Marotzke, and J. Willebrand, 1996: Stability of the thermohaline circulation. *The Warm Water Sphere of the North Atlantic Ocean*, W. Krauss, Ed., Borntraeger, 129–158..

Rivin, I., and E. Tziperman, 1997a: Linear versus self-sustained interdecadal thermohaline variability in a coupled box model. *J. Phys. Oceanogr.*, **27,** 1216–1232.. Find this article online

— ,and — ,1997b: Sensitivity of air–sea fluxes to SST perturbations. *J. Climate*, **10,** 2431–2446.. Find this article online

Saravanan, R., and J. C. McWilliams, 1995: Multiple equilibria, natural variability and climate transitions in an idealized ocean–atmosphere model. *J. Climate*, **8,** 2296–2323.. Find this article online

Schiller, A., U. Mikolajewicz, and R. Voss, 1997: The stability of the thermohaline circulation in a coupled ocean–atmosphere general circulation model. *Climate Dyn.*, **13,** 325–347..

Schlosser, P., G. Bonisch, M. Rhein, and R. Bayer, 1991: Reduction of deep water formation in the Greenland Sea during the 1980s:Evidence from tracer data. *Science*, **251,** 1054–1056..

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

Toggweiler, J. R., E. Tziperman, Y. Feliks, K. Bryan, S. M. Griffies, and B. Samuels, 1996: Reply. *J. Phys. Oceanogr.*, **26,** 1106–1110.. Find this article online

Tziperman, E., 1997: Inherently unstable climate behaviour due to weak thermohaline ocean circulation. *Nature*, **386,** 592–595..

— ,R. 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.. Find this article online

Walin, G., 1985: The thermohaline circulation and the control of ice ages. *Paleogeogr., Paleoclimatol., Paleoecol.*, **50,** 323–332..

Weaver, A. J., and T. M. C. Hughes, 1992: Stability and variability of the thermohaline circulation and its link to climate. *Trends Phys. Oceanogr.*, **1,** 15–70..

— ,and — ,1994: Rapid interglacial climate fluctuations driven by North Atlantic ocean circulation. *Nature*, **367,** 447–450..

— ,and S. Valcke, 1998: On the variability of the thermohaline circulation in the GFDL coupled model. *J. Climate*, **11,** 759–767.. Find this article online

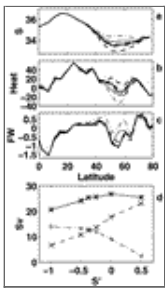
— ,E. S. Sarachik, and J. Marotzke, 1991: Freshwater flux forcing of decadal and interdecadal oceanic variability. *Nature*, **353,** 836–838..

Yu, E.-F., R. Francois, and M. P. Bacon, 1996: Similar rates of modern and last-glacial ocean thermohaline circulation inferred from radiochemical data. *Nature*, **379,** 689–694..

Zhang, S., R. J. Greatbatch, and C. A. Lin, 1993: A reexamination of the polar halocline catastrophe and implications for coupled ocean–atmosphere modeling. *J. Phys. Oceanogr.*, **23,** 287–299.. Find this article online

---

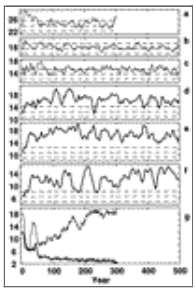
Figures



Click on thumbnail for full-sized image.

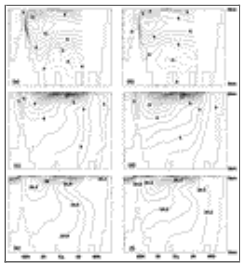
Fig. 1. (a) The zonally averaged and annually averaged North Atlantic surface salinity used to initialize the coupled model experiments. (b) and (c) Zonally and annually averaged heat and freshwater flux adjustments in the North Atlantic model sector. The line types used in (a)–(c) are consistently chosen for each experiment: [ΔS = +0.5]: thin dash–dot; [ΔS = 0]: thin solid; [ΔS =

−0.25]: thin dash;[Δ*S* = −0.375]: thick solid; [Δ*S* = −0.5]: thick dash–dot; [Δ*S* = −1]: thick dash. (d) The steady state ocean-only North Atlantic Deep Water formation rate (dash), the Southern Ocean Deep Water formation rate (dash–dot), and their sum (solid) as a function of the restoring salinity perturbation $S' = \Delta S$ in the North Atlantic.



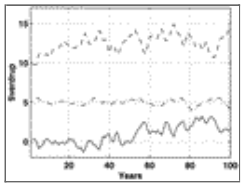
Click on thumbnail for full-sized image.

Fig. 2. The results of the coupled runs investigating the spontaneous (a)–(f) and forced (g) instability of the THC. Shown is the North Atlantic yearly averaged THC (in Sverdrup) vs time (years) from eight coupled model runs. Dash lines denote the initial state and plus/minus two standard deviation lines, where the standard deviation is defined from the control run (b). The different panels show (a) the coupled model run initialized with the stable state [Δ*S* = +0.5], (b) the stable control run (Delworth et al. 1993), (c) run initialized with the stable state [Δ*S* = −0.25], (d) run initialized with the unstable state [Δ*S* = −0.375], (e) run initialized with the unstable state [Δ*S* = −0.5], and (f) run initialized with the unstable state [Δ*S* = −1]. (g) Thick line: The North Atlantic THC for a coupled model run starting from year 10 of the run shown in (e), initialized with the unstable initial state [Δ*S* = −0.5]. Freshwater flux is added to the North Atlantic sinking area during the first ten years. The THC collapses and does not recover for 300 years, unlike the the corresponding experiment of Manabe and Stouffer (Manabe and Stouffer 1995), started from the stable control run and shown by the thin line.



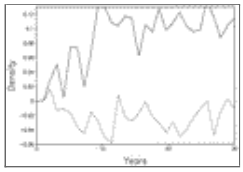
Click on thumbnail for full-sized image.

Fig. 3. North–south zonally averaged sections of Atlantic Ocean meridional overturning (a, b), temperature (c, d), and salinity (e, f) for the stable initial state of the control run (left panels) and for the unstable initial state nearest the instability threshold [Δ*S* = −0.375] (right panels).



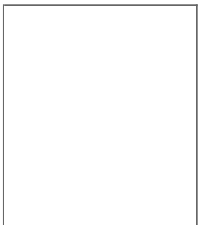
Click on thumbnail for full-sized image.

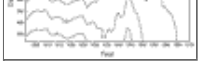
Fig. 4. Rates of deep-water formation (calculated as integrated downward velocity at level 5 of the ocean model, over the relevant area of the formation sites) as function of time for the unstable run [Δ*S* = −0.375]. Shown are the Norwegian Sea formation (dash–dot), SE of Greenland site (dash), and Labrador Sea (solid).



Click on thumbnail for full-sized image.

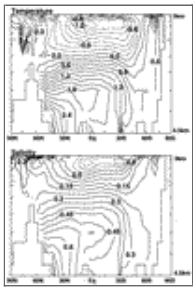
Fig. 5. Changes in surface water density in the North Atlantic sinking area due to the temperature (dash) and salinity (solid) changes, demonstrating that the instability mechanism is the linear salinity feedback.





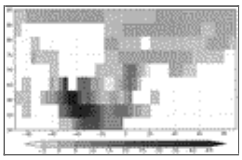
Click on thumbnail for full-sized image.

Fig. 6. A depth vs time plot of the (a) salinity and (b) temperature, averaged over the Labrador Sea area, for the unstable run [$\Delta S = -0.375$]. Note the gradual change in salinity from years 20 to 50 when the stratification below the surface water is weakened enough so that convective mixing is initialized.



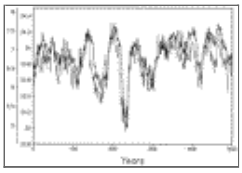
Click on thumbnail for full-sized image.

Fig. 7. Differences in the Atlantic Ocean temperature and salinity between the two forced instability experiments shown in Fig. 2g averaged over years 10–20.



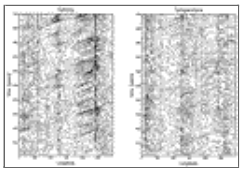
Click on thumbnail for full-sized image.

Fig. 8. A grayscale plot of the difference in buoyancy flux [kg/(m^2 yr)] due to the flux adjustments in the Labrador Sea, in the stable experiment [$\Delta S = -0.25$] and the unstable run [$\Delta S = -0.375$].



Click on thumbnail for full-sized image.

Fig. 9. Time series of the surface temperature (°C) and salinity (ppt) averaged over the North Atlantic sinking area for the unstable run [$\Delta S = -1$].



Click on thumbnail for full-sized image.

Fig. 10. A longitude vs time plot of the salinity anomaly in the circumpolar current region in run [$\Delta S = -0.375$].

Corresponding author address: Eli Tziperman, Department of Environmental Sciences, Weizmann Institute of Science, 76100 Rehovot, Israel.
E-mail: eli@beach.weizmann.ac.il

top