The Role of Temperature Feedback in Stabilizing the Thermohaline Circulation

STEFAN RAHMSTORF AND JÜRGEN WILLEBRAND

Institut für Meereskunde, Universität Kiel, Kiel, Germany

(Manuscript received 12 October 1993, in final form 28 June 1994)

ABSTRACT

Ocean climate models traditionally compute the surface heat flux with a restoring boundary condition of the form $Q = \lambda(T^* - T_O)$. This implies an atmosphere of fixed temperature and breaks down when large-scale changes in the ocean circulation are considered, which have a feedback effect on atmospheric temperatures.

To include this important feedback, a new thermal boundary condition of the form $Q = \gamma(T^* - T_O) - \mu\nabla^2(T^* - T_O)$ is proposed. This is derived from an atmospheric energy balance model with diffusive lateral heat transport. The effects of this new parameterization are examined in experiments with the GFDL modular ocean model for two model basins. "Conveyor belt" circulation states are compared using traditional mixed boundary conditions and our new coupling. With the new coupling, a realistic temperature contrast is obtained between the North Atlantic and the Pacific, caused by free adjustment of surface temperature to the oceanic heat transport.

The results show that a temperature feedback involving horizontal heat transport regulates the overturning rate of the conveyor. A second feedback involving vertical convection of heat stabilizes the conveyor belt when freshwater anomalies are added to the North Atlantic, making it harder to interrupt convection and trigger a halocline catastrophe.

1. Introduction

The World Ocean is one of the most important components of the climate system. On timescales from weeks to millennia, the dynamics of climate is strongly controlled by the behavior of the oceans since the ocean circulation moves large amounts of heat around the planet. Recent model studies and observations indicate that variations in the ocean circulation may have been a major cause for natural climate fluctuations in the past (Broecker 1991; Weaver and Hughes 1992). Paleoclimatological research shows that the formation of North Atlantic Deep Water was repeatedly interrupted or reduced in glacial times, leading to abrupt climatic changes in the North Atlantic region (Boyle and Keigwin 1987; Broecker 1991; Keigwin et al. 1991). Results from the GISP-2 ice core in Greenland (Taylor et al. 1993) indicate Pleistocene climate changes between glacial and near-interglacial conditions in periods of less than a decade, and on occasion as quickly as three years. Rapid climate fluctuations like this have not only occurred during the past glacial but probably also in the Eemian interglacial and the previous Saale-Holstein glacial cycle (Dansgaard et al. 1993; GRIP Members 1993). These data suggest that the stable conveyor belt circulation of the past 10 000 years is an exception rather than the rule.

Corresponding author address: Dr. Stefan Rahmstorf, Institut für Meereskunde, Universität Kiel, Düsternbrooker Weg 20, D-24105 Kiel, Germany.

Future climatic response to man-made changes in the atmospheric trace gas content will also depend to a large extent on oceanic heat storage and circulation (Houghton et al. 1990; Rahmstorf 1991). The observed fluctuations in the Eemian, which was warmer than the present, raise concern whether global warming could destabilize the deep circulation. Coupled model runs with greenhouse warming scenarios (e.g., Cubasch et al. 1993; Manabe and Stouffer 1993) so far indicate a weakening of deep water formation but no abrupt collapse.

Due to its low thermal capacity and fast adjustment time, the atmosphere can be considered in thermal equilibrium on climatic scales. For many climate studies, it is therefore not necessary to include a fully prognostic atmospheric model, which is very costly to run for longer times. Instead, the coupled system can be represented by a dynamical ocean model, where the response of the atmosphere is represented by a simple diagnostic parameterization (Bretherton 1982; Hasselmann 1991). A number of studies investigating climatic changes in the ocean have followed this approach and have used ocean models in which the coupling to the atmosphere was prescribed by a surface boundary condition (e.g., Maier-Reimer and Mikolajewicz 1989; Marotzke and Willebrand 1991; Hughes and Weaver 1994; Weaver et al. 1993).

It is important to realize that such a boundary condition implies a certain crude model of the atmospheric behavior. The ocean will respond in very different ways depending on the type of "atmosphere" to which it is

coupled. For example, prescribing a fixed heat flux at the ocean surface rules out any feedback of the ocean temperature on the heat exchange with the atmosphere. This means that the (vertically integrated) equilibrium heat transport of the ocean is entirely determined by the prescribed flux since the two must balance at each point. Investigations of how the ocean circulation affects heat transport, or investigations of climate change, are then ruled out. Zhang et al. (1993) have performed some perturbation experiments with flux boundary conditions, which confirm that the circulation remains stable in this case.

To allow some feedback between ocean temperatures and surface heat flux, a restoring boundary condition is commonly used in which the surface temperature of the model ocean is continuously restored to some fixed effective temperature, which we call restoring temperature. This approach was justified by Haney (1971). In this paper, Haney points out that "the basic assumption for this formulation is that the ocean is in contact with an atmospheric equilibrium state which is constant in time." It is therefore clear that the restoring boundary condition as suggested by Haney is not intended to be used for climate variability experiments. It implies a model of an atmosphere whose temperature always remains fixed.

The purpose of this paper is to derive a new thermal boundary condition that allows for the adjustment of atmospheric temperature and to investigate its effects in an ocean model. This boundary condition is easily implemented and provides a first-order approximation to the behavior of the coupled system.

In section 2 we will discuss the atmospheric models implied by various thermal boundary conditions and derive our new boundary condition from a simple energy balance model. In the remainder of the paper we will show some experiments with a two-basin ocean circulation model using this new boundary condition. In particular, we investigate whether the thermohaline circulation remains stable or collapses after a freshwater perturbation, a process crucially dependent on the nature of the temperature feedback with the atmosphere.

2. Thermal forcing of ocean models

a. *Parameterizations of surface heat flux*

As a first step to go beyond simply prescribing a constant heat flux at the ocean surface, we can assume this heat flux Q to be dependent on sea surface temperature T_O. Taking this dependency to be linear, we can write

$$Q = Q_0 + \lambda(T_A - T_O).$$

In this form, the second term represents those heat flux components that depend on the air–sea temperature difference $(T_A - T_O)$, namely, the latent, sensible, and net longwave radiation fluxes, and the constant Q_0 represents mainly the net solar radiation (plus constant parts of the other three components). The coupling constant λ is the sum of the sensitivities of the latent, sensible, and longwave fluxes to changes in T_O. It can be estimated from commonly used bulk parameterization formulas for these terms (found, e.g., in Busch 1977; Gill 1982; Henderson-Sellers 1986) and is of the order of 50 W m^{-2} K^{-1} (see the appendix). This is a strong coupling that will keep the ocean temperature within narrow bounds.

Sometimes a similar, purely pragmatic approach is used: $Q = Q_{\text{obs}} + \lambda(T_{\text{obs}} - T_O)$, a formula which simply nudges the model temperature back to the observed sea surface temperature. In this interpretation, λ need not be seen as a value that characterizes a physical feedback but could be considered an arbitrary nudging parameter.

Both approaches are equivalent in form and can be written as

$$Q = \lambda(T^* - T_O) \tag{1}$$

(Haney 1971), where the restoring temperature T^* is defined as $(T_A + Q_0/\lambda)$ or $(T_{\text{obs}} + Q_{\text{obs}}/\lambda)$, respectively. Often simply $T^* = T_{\text{obs}}$ is used, which is a good approximation for large λ. Formally, $T_O \equiv T^*$ is the solution for vanishing heat flux Q; oceanic motions and the associated heat transport lead to deviations of T_O from the restoring temperature field $T^*(x, y)$. This restoring boundary condition has served well for many ocean modeling purposes, since it guarantees realistic ocean surface temperatures, provided that T^* is chosen appropriately and λ is large enough so that the restoring timescale is small compared to the advective timescale. However, since it is in effect a way of (more or less loosely) *prescribing* the surface temperature (by nudging it to a prescribed value), it is not suited to climate variability experiments, where the temperature of a different climate needs to be *predicted.* This problem was recognized by Schopf (1983) in the context of modeling El Niño/Southern Oscillation.

Consider a climate variability experiment that involves a change of ocean circulation and heat transport. Assume that the change will lead to a net heat transport toward a given region of the ocean, where a thermal "anomaly" develops. How will this affect the surface heat flux? Equation (1) describes the change in heat flux with the *assumption* that the atmospheric temperature remains constant. This assumption will be good if the thermal anomaly has a small spatial scale, so that winds will rapidly disperse the additional heat entering the atmosphere from the ocean. In case of a larger-scale heat anomaly, the atmosphere will be unable to disperse all the heat and will warm up in response, so that the assumption is no longer valid. In the limit of a uniform global heat anomaly, no heat can be removed in the horizontal, and the atmosphere has to warm until longwave radiation to space balances the increased heat flux from the ocean. Since the longwave budget is only weakly temperature dependent (compared to the strong ocean–atmosphere coupling),

the atmosphere will warm almost as much as the underlying ocean.

This illustrates that the response of the surface heat flux to a change in ocean temperature will depend on the spatial scale of the change, which determines the extent to which the atmosphere can remove additional heat by advection. This scale dependency is not a second-order correction but a crucial property of the atmospheric response. The heat flux sensitivity $\partial Q/\partial T_O$ varies over more than an order of magnitude between the smallest and largest scales. Bretherton (1982) has suggested relating the heat flux at a given point to the surface temperature at all other points through a Green's function to account for this nonlocality. Our approach is different and based on a parameterization of the horizontal heat transport in the atmosphere.

b. *A simple atmospheric model*

We start by considering a simplified local heat budget of the ocean–surface–atmosphere system (Fig. 1). The incoming net shortwave radiation Q_S is absorbed partly by the atmosphere and partly by the ocean. The ocean emits longwave radiation Q_B, which is partly absorbed by the atmosphere, which in turn emits longwave radiation upward (Q_U) and downward (Q_D). The air–sea heat exchange Q_C is the combined effect of latent and sensible heat flux. The horizontal heat transport divergence of the atmosphere is denoted Q_A.

The atmospheric heat budget is

$$fQ_S + eQ_B - Q_D - Q_U + Q_C + Q_A = 0, \quad (2)$$

and the balance equation at the ocean surface can be written as

$$(1-f)Q_S - Q_B + Q_D - Q_C = Q. \quad (3)$$

(See Table 1 for the meaning of all symbols.) We aim to represent the behavior of the coupled system by a simple parameterization $Q(T_O)$. However, the terms of the heat budget depend not only on T_O but on a range of atmospheric variables such as air temperature, humidity, wind speed, and cloud cover. The "zero-order" approach to close this problem, taken by Haney (1971), is to assume that all these parameters remain constant. Then the atmosphere behaves like an infinite heat sink, unaffected by the heat flux it receives from the ocean.

We want to take the closure of the heat budget one important step further by allowing the air temperature to react to changes in ocean temperature, while still neglecting the effects of changes in the other atmospheric variables. Instead of specifying T_A we determine it using the atmospheric heat budget [Eq. (2)]. The atmosphere has been assumed to be in steady state, which is a very good assumption for the purpose of long-term ocean modeling. Given an ocean model that calculates the sea surface temperature fields, Eqs. (2) and (3) allow us to compute the heat flux at the ocean surface, provided that we can parameterize the terms

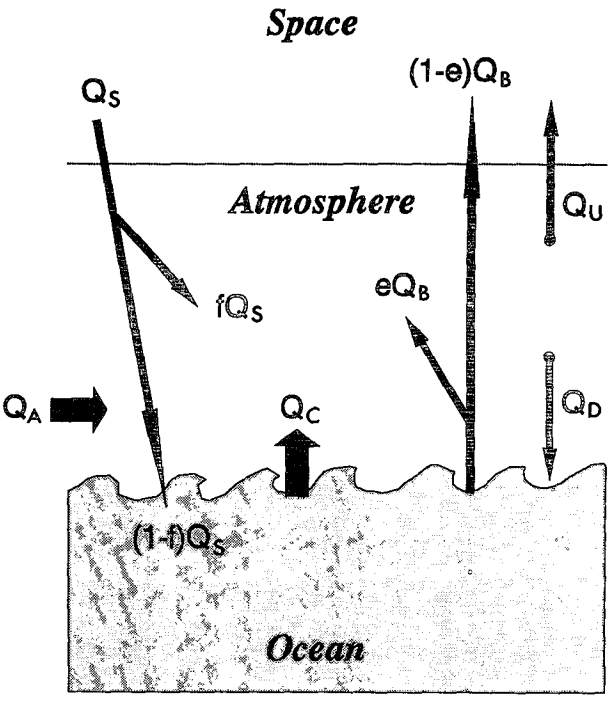

FIG. 1. Local heat balance of the ocean surface and the atmosphere.

in the heat budget as functions of the oceanic and atmospheric temperatures T_O and T_A. With T_A we mean the temperature appearing in the bulk formulas for air–sea heat exchange, that is, the temperature of the turbulent boundary layer near the ocean surface. Some longwave radiation terms of the heat budget depend on the temperature higher up in the atmosphere but can be parameterized as function of the surface temperature (Budyko 1969). The dependence of the outgoing longwave radiation on surface temperature, $d = \partial Q_U/\partial T_A$, is an important parameter of the climate system, sometimes called the climate sensitivity.

We then linearize the temperature dependence of the heat budget terms (except Q_A, which is dealt with below) with respect to a reference temperature T_{ref} so that (2) and (3) take the form

$$A + BT_O - CT_A + Q_A\{T_A\} = 0 \quad (4)$$

$$D - ET_O + FT_A = Q, \quad (5)$$

where A–F are constants or prescribed functions of latitude (or more generally, of space and time), listed in Table 1, and T_A and T_O are from now on expressed as deviations from the reference temperature. Physically speaking, A–F contain the solar input to the system, Q_S, as well as two crucial coupling constants, which determine the temperature dependencies of the heat budget: the strong air–sea coupling c and the weaker radiative coupling d, with $c \gg d$ (for a full discussion of the heat budget and definition of lower-case parameters see the Appendix). The atmospheric transport term Q_A is a nonlocal function of T_A. We could rep-

TABLE 1. Parameters of the heat budget. Values in parentheses are global average values for the global heat budget example calculated in the appendix.

Parameter	Value	Explanation
Q_S	(240 W m^{-2})	net shortwave radiation at top of atmosphere
Q_B	(392 W m^{-2})	oceanic longwave back radiation
Q_D	(316 W m^{-2})	downward atmospheric longwave radiation
Q_U	(220 W m^{-2})	upward atmospheric longwave radiation
Q_C	(100 W m^{-2})	air–sea heat flux (sensible plus latent heat)
Q_A		divergence of atmospheric heat transport
a	0.85	empirical constant giving downward longwave radiation at T_{ref}
b	0.55	empirical constant giving upward longwave radiation at T_{ref}
c	43 W m^{-2}K^{-1}	empirical coupling constant for air–sea heat flux
d	3 W m^{-2}K^{-1}	sensitivity of atmospheric longwave radiation
e	0.95	fraction of oceanic longwave radiation absorbed by atmosphere
f	0.27	fraction of net shortwave radiation absorbed by atmosphere
k	0.9×10^{13} W K^{-1}	constant of temperature diffusion in atmosphere
$\mu = kBF/C^2$	0.8×10^{13} W K^{-1}	constant of temperature diffusion in boundary condition Eq. (11)
ϵ	0.98	emissivity of ocean
σ	5.73×10^{-8} W m^{-2}K^{-4}	Stefan–Boltzmann constant
T_{ref}	273 K	reference temperature for linearization
$r = 4\epsilon\sigma T_{\text{ref}}^3$	4.6 W m^{-2}K^{-1}	sensitivity of ocean back radiation
$R = \sigma T_{\text{ref}}^4$	318.3 W m^{-2}	longwave radiation constant
$A = fQ_S + (e\epsilon - a - b)R$	(−84.5 W m^{-2})	footnote a
$B = er + c$	47.3 W m^{-2}K^{-1}	footnote a
$C = 2d + c$	49.0 W m^{-2}K^{-1}	footnote a
$D = (1 - f)Q_S - (\epsilon - a)R$	(133.8 W m^{-2})	footnote a
$E = r + c$	47.6 W m^{-2}K^{-1}	footnote a
$F = d + c$	46.0 W m^{-2}K^{-1}	footnote a

[a] Parameters of linearized heat budget (4) and (5).

resent it for example by a diffusive ($Q_A = k\nabla^2$) or advective ($Q_A = -\mathbf{U}\cdot\nabla$) operator. For fixed T_A (5) reduces to the traditional restoring (1) characterized by the strong coupling constant c.

The atmospheric budget (4) can be formally solved for T_A and inserted in (5) so that the flux entering the ocean surface becomes

$$Q = D - ET_O + F(C - Q_A)^{-1}\{A + BT_O\}. \quad (6)$$

On the time and space scales important for ocean climate experiments the air temperature is dominated by the local coupling to the ocean temperature and not by atmospheric heat transport, that is, $C \gg Q_A$, so that we can expand the operator as

$$(C - Q_A)^{-1} = \frac{1}{C} + \frac{Q_A}{C^2} + \cdots \quad (7)$$

so that to first order in Q_A/C (6) becomes

$$Q = D + \frac{FA}{C} + \left(\frac{FB}{C} - E\right)T_O + \frac{F}{C^2}Q_A\{A + BT_O\}. \quad (8)$$

We now define a restoring temperature T^* by evaluating this heat budget for $Q \equiv 0$:

$$0 = D + \frac{FA}{C} + \left(\frac{FB}{C} - E\right)T^* + \frac{F}{C^2}Q_A\{A + BT^*\}. \quad (9)$$

This allows us (assuming we use a linear operator for Q_A) to express the surface heat flux in terms of the deviation of the sea surface temperature T_O from the fixed restoring temperature T^* [by subtracting (9) from (8)]:

$$Q = \left(\frac{FB}{C} - E\right)(T_O - T^*) + \frac{BF}{C^2}Q_A\{T_O - T^*\}. \quad (10)$$

For a diffusive heat transport in the atmosphere ($Q_A = k\nabla^2$), we obtain a restoring boundary condition for the ocean model of the form

$$Q = \gamma(T^* - T_O) - \mu\nabla^2(T^* - T_O) \quad (11)$$

(first suggested by Willebrand 1993). The first term has the same form as the Haney restoring term, but it has a different meaning. It represents the relaxation of a heat anomaly by longwave radiation. The coupling constant γ is in this case much weaker than the air–sea exchange coupling, namely, 2 or 3 W m^{-2} K^{-1} (see the appendix). The second, diffusive term reflects the heat transport induced in the atmosphere by the oceanic heat transport and surface exchange. It allows the model atmosphere to respond to changes in the ocean circulation by dispersing some of the heat it receives from the ocean. This atmospheric model is simple but captures the essential feedbacks needed for long-term ocean climate studies.

The restoring temperature T^* differs from the one used in Haney's approach, since the atmospheric temperature is not fixed any more, and both T_A and T_O will deviate by several degrees from T^* in response to oceanic heat transport; T^* is the temperature toward which the ocean is forced or the equilibrium temperature that it would reach in the absence of currents and the associated heat transport. The T^* field can therefore be calculated from an atmospheric circulation model bounded by a thermally insulating "swamp" instead of an ocean. Within the framework of our linearized heat budget T^* can in principle be calculated from (9), but we chose to use a simple cosine function of latitude (see section 3).

Returning to the discussion at the end of section 2a, we note that the magnitude of the second term in (11) depends on the scale of the temperature anomaly. For a temperature anomaly of magnitude ΔT and a characteristic spatial extent Δx, this term is of the order $\mu \Delta T/(\Delta x)^2$. It goes to zero for global-scale anomalies and becomes increasingly larger for smaller scales. That it keeps increasing for ever smaller scales, instead of leveling out at the small-scale limit, is due to the expansion (7), the validity of which breaks down for small-scale anomalies that are easily removed by advection. It means that the smallest disturbances cannot survive because they are diffused too quickly. This could be prevented by applying a cutoff value beyond which the term is not allowed to increase, and which corresponds to the small-scale limit. In a coarse-resolution model the grid scale provides an automatic cutoff. We chose μ so that at a grid scale of $\Delta x \approx 400$ km the coupling sensitivity is 50 W m^{-2} K^{-1} (the value for the small-scale limit); that is, $\mu = 50$ W m^{-2} K^{-1} $\times$ (400 km)^2 $= 8 \times 10^{12}$ W K^{-1}. In order to get a realistic lifetime of SST anomalies at 1000-km scales, the heat flux sensitivity at these scales should be 10–30 W m^{-2} K^{-1} (Anderson and Willebrand 1992). Our choice of μ (and γ) corresponds to a coupling of 11 W m^{-2} K^{-1} at this scale, at the low end of this range.

In the temperature equation for the top level of the ocean model ($\partial T/\partial t = Q/h\rho c_p$) the diffusion term acts with a diffusivity of $\nu_0 = \mu/h\rho c_p$; h is the level thickness. For a surface layer of 25 m, we get $\nu_0 = 8 \times 10^4$ m^2 s^{-1}. This apparent diffusivity caused by the atmospheric heat transport is two orders of magnitude larger than the oceanic diffusivity used in our model, but we have to bear in mind that it does not act on surface temperature per se but only on deviations from the restoring temperature. In a similar way we can calculate the *atmospheric* diffusivity implied in Eq. (11); it is $\nu_A = k/h\rho c_p$, where this time we have to divide by the heat capacity of the active layer of the atmosphere. For the whole atmosphere, with a heat capacity of 10^7 J m^{-2} K^{-1} (Gill 1982, p. 22), $\nu_A = 0.9 \times 10^6$ m^2 s^{-1}. If only the lower 5 km (with just under one-half the heat capacity) of the atmosphere is involved in diffusing thermal anomalies, then $\nu_A = 2 \times 10^6$ m^2 s^{-1}. Gill (1982, p. 591) gives a typical value of eddy diffusivity in the atmosphere of 2×10^6 m^2 s^{-1}.

The analysis of the simplified ocean–atmosphere heat budget presented here is useful for understanding the properties of the atmosphere implied by the use of different restoring approaches. Using the small-scale limit, that is, traditional Haney restoring, implies a kind of "infinite heat capacity atmosphere," which does not change its temperature in response to oceanic heat fluxes. The large-scale limit has been labeled a "zero heat capacity atmosphere" (Zhang et al. 1993), but is perhaps better referred to as the "fixed heat transport" case, where the atmosphere can absorb additional heat from the ocean only at the rate at which it can radiate it to space. The new boundary condition presented in Eq. (11) implies an energy balance model of the atmosphere with a diffusive horizontal heat transport. The effect that this approach has on ocean models will be investigated in the model calculations presented below.

3. The ocean model

a. Model description

The circulation model used here is in most respects identical to that of Marotzke and Willebrand (1991, hereafter MW) (Fig. 2). It consists of two basins of 60° width, which represent the Pacific and the Atlantic and will be referred to as such, although they are identical in geometry. The basins are of uniform depth (4500 m) and are linked in the south by a "circumpolar

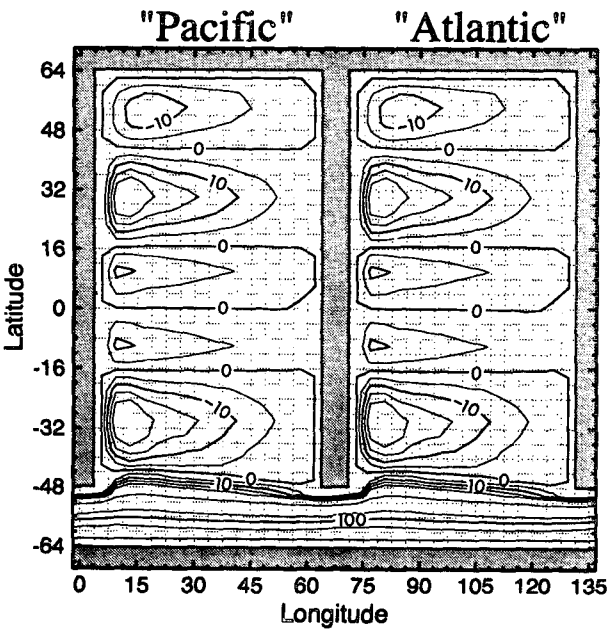

FIG. 2. Geometry of the two-basin ocean model. Dotted lines indicate the model grid. Contours are of the barotropic streamfunction, contour interval is 5 Sv, and above 20 Sv the interval is 20 Sv.

current" with a cyclic boundary condition joining 0° and 135° longitude. This highly symmetrical setup is designed to isolate the effects of the surface forcing from those of topography. The latter are considered to be important for the deep circulation, but are not the subject of this study.

The numerical code is the GFDL modular ocean model (MOM; Cox 1984; Pacanowski et al. 1991, 1993), with a grid resolution of 3.75° × 4° and 15 levels in the vertical. The model uses the rigid-lid approximation at the surface, a free-slip boundary condition at the bottom, and no slip at the lateral walls. The time step is 1.5 hours for all time-dependent experiments, with integration periods of typically some decades. For long-term integrations designed to prepare equilibrium states and typically lasting 3000–5000 model years, split time stepping was used (Bryan 1984), extending the time step for the tracer equations to 1 day. Constant horizontal and vertical diffusivities ($K_H = 10^3$ m^2 s^{-1}, $K_V = 5 \times 10^{-5}$ m^2 s^{-1}) and viscosities ($A_H = 2.5 \times 10^5$ m^2 s^{-1}, $A_V = 10^{-4}$ m^2 s^{-1}) were applied.

As in MW, the transport around Antarctica was prescribed because the dynamics of the ACC are not easily represented in coarse-resolution models and are not of interest for the present study. Important is only that the ACC is there, working as a connection between the two basins. A transport value of 140 Sv was used (compared to 200 Sv in MW).

While MW removed the nonlinear advection terms from the momentum equations, we used the full primitive equations since no significant time saving arises from this approximation [Hughes and Weaver (1994) estimate the saving as perhaps 10%]. Another minor difference between MW and our configuration is a somewhat different distribution of depth levels, with a top level of thickness 50 m in MW and 25 m in this study. These differences in configuration led to no significant differences in the deep circulation.

We use a convection scheme that completely removes all static instability at each time step. Weaver et al. (1993) argue that the computational expense of such a scheme outweighs its advantages, and they prefer the implicit diffusion (IVD) scheme. However, our implementation of the complete convection scheme is about twice as fast as the IVD scheme, typically adding 25% to the overall computational expense of the model, compared to 52% for the IVD scheme or 90% for seven iterations of the (incomplete) standard GFDL scheme (Rahmstorf 1993).

b. *Forcing and spinup*

Two types of model runs will be discussed in this paper: control runs with traditional temperature restoring (which we will refer to as HR, for Haney restoring), and the experiments with the new energy balance boundary condition, labeled EB. In both cases the standard spinup procedure was followed. The model was spun up from an isothermal, isohaline ocean at rest for 5000 years with restoring of both salinity and temperature, leading in both cases to a highly symmetrical circulation state with deep water formation in both northern and southern basins (a ++++ state in MW notation, where a + or − sign mark presence or absence of deep water formation in North Pacific, South Pacific, South Atlantic, and North Atlantic, in this order). From these equilibria the freshwater fluxes were diagnosed.

For the restoring salinity S^* we used the same simple cosine function of latitude as MW, varying from 36 psu at the equator to 33 psu at the northern and southern model boundaries. This representation does not include the effects of net precipitation in the intertropical convergence zone but can be justified for such an idealized two-basin model designed for process studies. The diagnosed freshwater fluxes were zonally averaged since they are derived from a spinup with zonal-average restoring salinity. Any zonal structure in these fluxes would result from the model salinities being forced back to zonal average values, so it has no physical basis. The diagnosed fluxes are shown in Fig. 3 and have a net precipitation over the Northern Hemisphere and net evaporation over the Southern Hemisphere. As shown by MW and confirmed by our own results, this can lead to a model preference for deep water formation in the south, while symmetrical rainfall in both hemispheres (i.e., no cross-equatorial salt transport) leads to a preference for deep water formation in the north. We chose to keep the hemispheric asymmetry for the following reasons. First, the model state that was used to diagnose these fluxes was a symmetrical (++++) state with overturning cells in both basins in both hemispheres rather than a cell extending across the equator, so the cross-equatorial freshwater transports

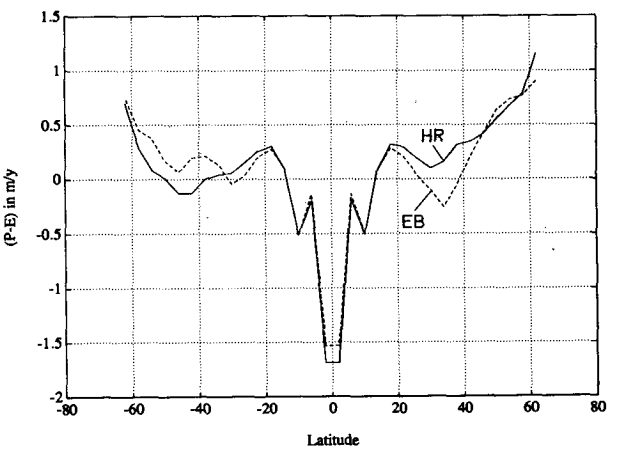

FIG. 3. Diagnosed zonal average freshwater fluxes for the Haney restoring (HR) and energy balance (EB) runs. The HR flux has more precipitation in the northern midlatitudes.

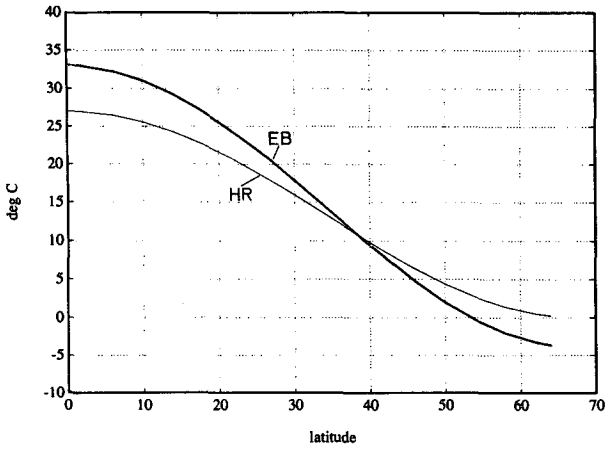

FIG. 4. Restoring temperatures T^* for the Haney restoring (HR) and energy balance (EB) runs. Note the greater contrast between low and high latitudes of the EB restoring temperature.

of 0.2 Sv for the HR spinup and 0.03 Sv for the EB spinup seem to arise not from this particular state but may be a more generally valid feature caused by the north–south asymmetry in the geometry. Second, Broecker et al. (1990) show that the real ocean has a similar, even larger, cross-equatorial freshwater transport in the same direction.

The restoring temperature T^* used for the HR spinup is the same as in MW, a cosine function varying from 0°C at the northern model boundary to 27°C at the equator. For the EB runs with the new boundary condition, T^* has a somewhat different physical interpretation and a larger equator to pole contrast. We kept the same functional form, however, and varied the extremes at the equator and high latitudes until we found a spinup state with a similar SST contrast between low and high latitudes as in the control (HR) spinup. This was achieved using a T^* varying from −4°C to 33°C. It is the oceanic heat transport that then moderates the SST contrast. This free adjustment of temperatures in response to the oceanic heat transport is the very process we wanted to capture with the new boundary condition. The restoring temperatures for HR (Haney restoring) and EB (energy balance) experiments are shown in Fig. 4.

The thermal coupling constant λ for the HR runs was chosen to be the same as in MW, namely 80 W m^{-2} K^{-1} (this corresponds to a damping timescale of 30 days for their 50-m surface model level, and 15 days for our 25-m surface level). In the EB runs with the scale-dependent boundary condition, we used $\gamma = 3$ W m^{-2} K^{-1} (damping timescale 1 year for a 25-m layer), which in this case represents the large-scale limit.

The second coupling constant in (11) was chosen as $\mu = 8 \times 10^{12}$ W K^{-1} for most runs, as discussed in section 2. We also performed some experiments with different atmospheric diffusion in x and y directions to reflect the fact that the atmosphere has a preference for zonal transport. Zonal surface wind speeds are typically an order of magnitude larger than meridional velocities. In these "anisotropic" runs we used $\mu_x = 1.5 \times 10^{13}$ W K^{-1} and $\mu_y = 1.5 \times 10^{12}$ W K^{-1}, thus keeping the total coupling sensitivity roughly the same as before, while favoring zonal diffusion of the heat released by the ocean by a factor of 10 over meridional diffusion.

The wind forcing consisted of a simple latitude-dependent zonal wind stress as pictured in MW. The barotropic flow (shown in Fig. 2) is almost entirely determined by this wind forcing due to the flat bottom and coarse resolution of the model and does not change between different model runs.

As mentioned before, the spinup states under restoring boundary conditions were ++++ states with deep water formation in both northern and southern basins and corresponding overturning cells. In the control spinup, the southern cells were much weaker (2 Sv) than the northern cells (12 Sv), presumably because the ACC gap inhibits meridional flow by preventing the buildup of zonal pressure gradients (Gill and Bryan 1971; England 1993). In the EB spinup, in contrast, the deep circulation cells were more balanced, with 7 Sv in each basin and hemisphere (Fig. 5). The reason for this is a negative temperature feedback that regulates the strength of the overturning: weak overturning leads to a reduced poleward heat transport, cold high-latitude surface temperatures, and thus enhanced density and deep water formation. In the control run (HR), strong coupling to a fixed restoring temperature suppresses this important feedback.

As reported by MW, the ++++ state is unstable upon switching to mixed boundary conditions, that is, replacing the salinity restoring by prescribing the diagnosed freshwater fluxes. This can be made plausible

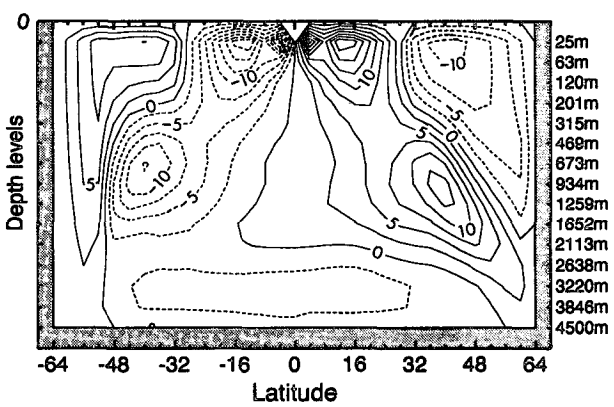

FIG. 5. Meridional overturning (Sv) for the symmetrical spinup state with salinity restoring and the new thermal boundary condition. Overturning is identical in both basins, so that for each basin separately the values are one-half of the global streamfunction shown here. Note the exponentially stretched vertical scale (tickmarks are model levels).

by looking at the deep circulation in the model as a competition between northern and southern deep water trying to fill each basin. Greater density is an advantage in this competition; density depends on the conditions in the source region but also on mixing along the way. In a ++++ state the northern and southern deep waters are balanced, but with mixed boundary conditions this situation becomes unstable due to the positive salt feedback: as soon as the overturning in one hemisphere gets stronger, more salt is advected toward its deep water formation areas, making the deep water denser and reinforcing the overturning. This side "wins" and suppresses the deep water formation of the other side by filling the deep basin with denser water than the other hemisphere can form.

Switching to mixed boundary conditions led to a transition of the control spinup to a southern sinking state (−++−), which is preferred because of the net precipitation over the Northern Hemisphere. If this precipitation bias is removed by using symmetrical freshwater fluxes over both hemispheres, starting again from the control spinup (HR), then the transition is to a northern sinking state (+−−+), presumably because meridional flow in southern high latitudes is hindered by the ACC gap. The freshwater flux diagnosed from the EB spinup has not such a strong rainfall excess in the Northern Hemisphere (0.03 Sv) as the one from the control (HR) spinup (0.2 Sv, see also Fig. 3), so that the EB spinup flips to a northern sinking state upon switching to mixed boundary conditions, even if the hemispheric asymmetry in the fluxes is retained.

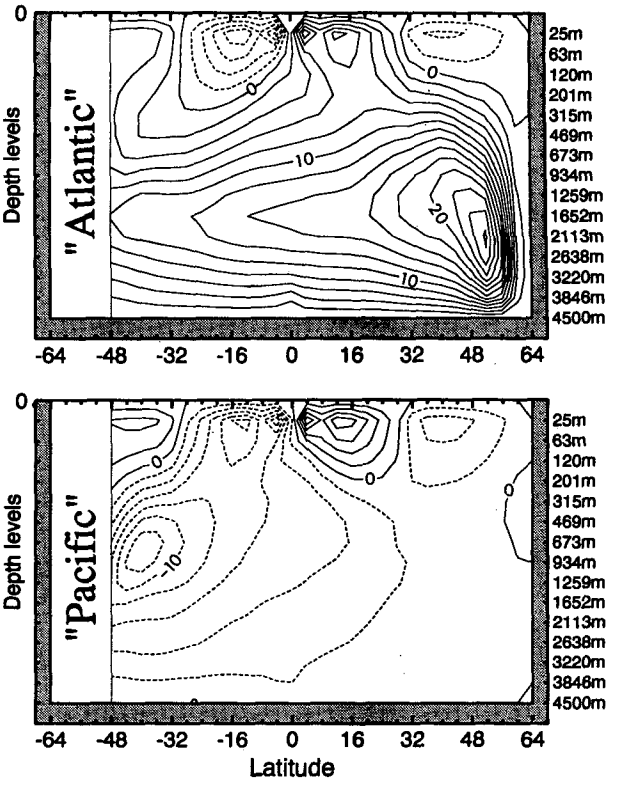

FIG. 6. Meridional overturning (Sv) in both basins for the conveyor equilibrium of the HR run. Contour interval: 2 Sv.

4. Conveyor belt circulation with traditional and new thermal coupling

In this paper we are mainly interested in conveyor belt circulation states (−+−+) since this is the present state of the World Ocean (Gordon 1986). Conveyor states were prepared from the two spinup states by switching to mixed boundary conditions, while at the same time applying a temporary freshwater flux of +0.18 m yr^{-1} in the North Pacific and −0.18 m yr^{-1} in the North Atlantic (north of 44°N) for 500 years. Once the conveyor is established, this excess flux is removed, and the integration is continued to equilibrium for another 4500 years with the diagnosed zonal average fluxes for HR and EB, respectively. The conveyor states obtained like this exist under the same boundary forcing as the northern, respectively southern, sinking states mentioned above; it is only the different initial condition (at the end of the 500-yr transition period) that causes the different states.

We will now compare the conveyor belt equilibrium states achieved with the traditional strong temperature restoring (HR) and with the new energy balance coupling (EB) and discuss their differences. The meridional overturning of the control conveyor (HR) in the Atlantic and Pacific basin is shown in Fig. 6. This state is very similar to the conveyor equilibria discussed in MW. Near the surface in both basins wind-driven Ekman cells are visible, which are a constant feature of all model runs. More interesting for this work are the deep thermohaline cells. The Atlantic is filled with such a cell of 26 Sv, driven by deep water formation at the northern model boundary. The Pacific is dominated by a much weaker reverse cell with deep water coming from the Southern Ocean. The figure cannot show whether this deep water forms in the Southern Ocean or is derived from the outflow from the Atlantic basin.

Before we answer this question, we should add a word of caution about the interpretation of zonally averaged overturning charts. A common misconception is that the circulation follows the contours of the streamfunction and that the deep water can be traced back along those contours. In Fig. 6, this would give the false impression that the North Atlantic deep water originates at 30°N; in fact it is ventilated by deep convection at the northern edge of the basin. Water originating at the surface near 30°S in the Pacific appears to "sink" at least to intermediate depths. However, a slightly tilted horizontal circulation gyre can look like a deep vertical overturning cell in the zonal-averaged picture, even if no water particles actually sink all the way down in its downward branch, and it does not ventilate the deep ocean. If we look for deep-water for-

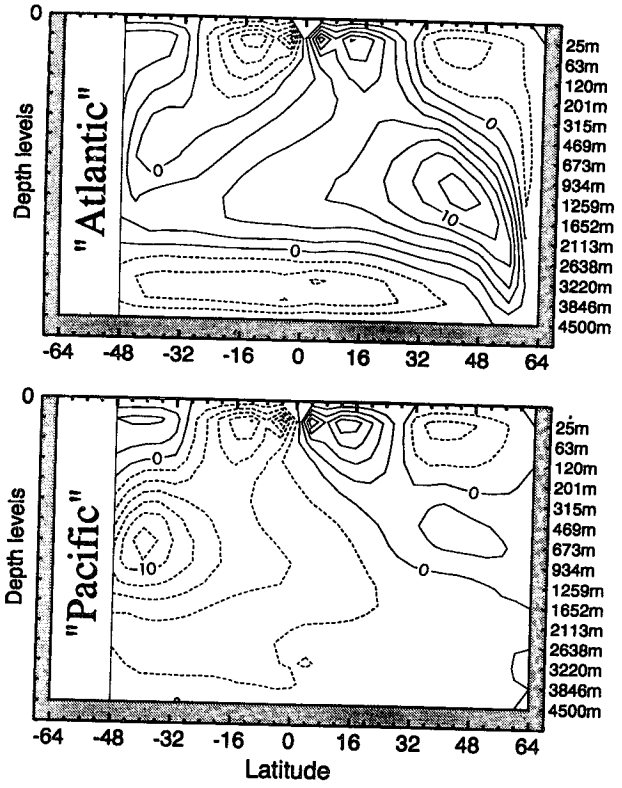

FIG. 7. As in Fig. 6 but of the EB run with the new thermal boundary condition.

mation areas, therefore, meridional overturning charts are not enough, and we must combine their information with plots of convection, water properties, and velocities.

There is a deep convection region at the southern boundary of the model domain, not only in the Pacific sector but also in the Atlantic sector. Deep convection releases buoyancy from the water column and is therefore always associated with an outflow of dense water at the bottom. Less dense water is drawn in farther up in the water column (but not necessarily at the surface). In the ACC region meridional flow is inhibited by the lack of pressure gradients, so that water from the convection region cannot flow straight into the deep Pacific and strong meridional temperature and salinity gradients persist. Based on property and velocity plots (not shown), we conclude that the deep water in the Pacific basin of our model is a mixture of water from the deep Atlantic outflow with water ventilated by convection at the edge of the Antarctic continent. The two are mixed in the ACC channel with the help of lateral diffusion and large-scale meanders.

The conveyor circulation obtained with the new boundary condition is shown in Fig. 7 for comparison. Again we see a positive deep circulation cell in the North Atlantic and a reverse cell centered in the South Pacific, but this time they are of equal strength (13 Sv). As in the spinup state discussed above, this must be a result of the negative temperature feedback regulating the overturning rate. This is supported by trials to produce a different conveyor circulation. Instead of kick-starting the conveyor with a 0.18 m yr^{-1} anomaly (see the beginning of this chapter), we have also tried 0.01, 0.5, and 2.0 m yr^{-1}. While the first of these did not yield a conveyor but a northern sinking state like the case with zero perturbation, the other two led to a conveyor with 13 Sv overturning just like the standard case. With the traditional HR boundary condition, in contrast, conveyor states of widely differing strength can be obtained (Hughes and Weaver 1994; Rahmstorf 1994).

Another striking difference to the HR run is the existence of an Antarctic bottom water cell underlying the North Atlantic deep water cell in much of the Atlantic. The reason for this becomes apparent when we examine the sea surface temperature (SST) map of this model state (Fig. 8). The EB state differs from the HR state (where SST is closely tied to the restoring temperature T^* and thus essentially zonal) in that SST responds more freely to advective heat transport. This leads to a more realistic effect of the wind-driven gyres on SST, visible as deviations from a purely zonal structure. More importantly, the heat transport by the overturning circulation has a marked effect on temperature, leading to a much warmer North Atlantic than North Pacific. Average SST north of 48°N is 3.4°C in the Pacific and 8.7°C in the Atlantic. The observed temperature contrast between Atlantic and Pacific (Levitus 1982) is around 4°C. Correct prediction of this tem-

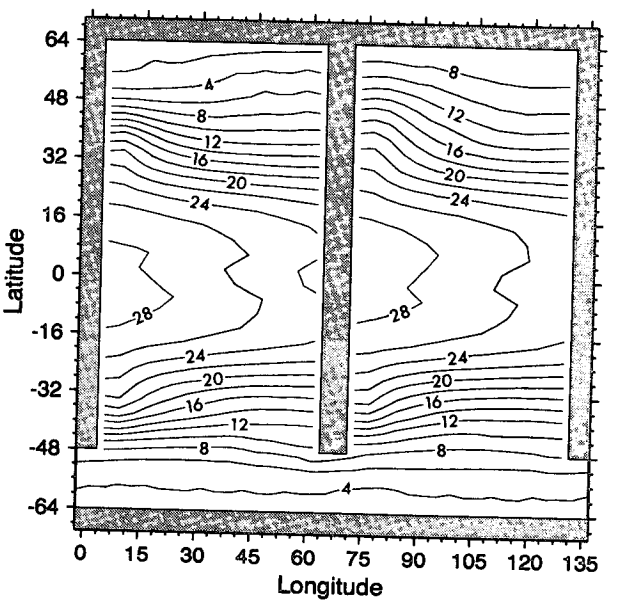

FIG. 8. Sea surface temperature map of the conveyor equilibrium of the EB run with the new boundary condition. Contour interval: 2°C.

perature contrast, caused by the conveyor circulation, is a crucial test for the new boundary condition. The model overestimates this temperature contrast somewhat; the North Atlantic appears too warm. This could mean that the coupling to T^* in the new boundary condition is slightly too weak. Alternatively, the overestimation could be caused by a more fundamental limitation of our boundary condition: the effect of continents cannot be included. We did not allow the atmospheric diffusion to operate across the land barriers, which results in all the heat brought by the conveyor being trapped within the ocean basin.

The effects of heat transport described above also make the model Southern Ocean significantly colder than the North Atlantic (but slightly warmer than the North Pacific). There is a southern overturning cell in the Pacific of the same strength as the northern cell in the Atlantic, but it cannot heat the Southern Ocean efficiently. There are two reasons for this. The ACC mixes the water zonally around the globe, so that the cell has to warm an ocean area twice the size of the corresponding model area in the North Atlantic. In addition, the ACC acts like a wall to the cell; the cell cannot extend all the way south because in the ACC area no zonal pressure gradient can establish itself to support meridional flow. The colder (and for similar reasons fresher) Antarctic water is heavier than the North Atlantic deep water formed by the model, and thus penetrates into the Atlantic as a bottom water cell, underlying the North Atlantic Deep Water not unlike in the real ocean.

Temperature and salinity sections along the western side of the Atlantic basin reflect this flow pattern (Fig. 9). Cold freshwater from the ACC region penetrates northward into the Atlantic at the ocean bottom and crosses the equator. A tongue of high salinity water extends south at depth 2000–2500 m, associated with deep water outflow from the Atlantic basin. Above it is a salinity minimum reminiscent of intermediate water, which is formed by convection off the middle land barrier ("South America") in a similar way as in the model of England et al. (1993).

Note that the model produces these features—warm salty "NADW," cold fresh "AABW," and a stably stratified North Pacific—completely as a consequence of the internal dynamics, not by tuning the surface forcing to produce specific water masses [as in the global model of England (1993)]. The temperature boundary conditions are fully symmetrical in both basins and both hemispheres. The freshwater forcing is zonally uniform and differs between hemispheres, but it was diagnosed from a fully symmetric restoring condition. It therefore compensates to some extent for the effect of the ACC by giving excess precipitation in the north.

Returning to the image of a "competition" between southern and northern deep water to fill the model basins, we now see that both can coexist stacked one above the other. Although the Antarctic deep water is denser, it cannot fill the whole North Atlantic and prevent convection there. This is probably because it cannot be produced in sufficient quantity because of the dynamical constraint in the ACC region (discussed above), which limits the outflow rate. By the time the AABW reaches the North Atlantic, it is so diluted that it is less dense than the freshly formed NADW, which extends all the way to the bottom in its source region.

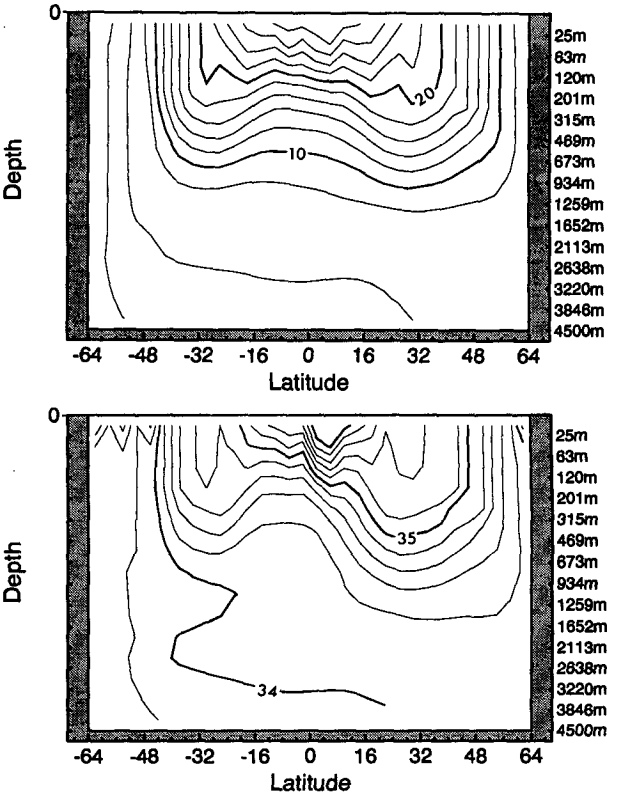

FIG. 9. Temperature section (a) and salinity section (b) along the western side of the model Atlantic at 80° longitude for the EB run.

At this point it is useful to examine the role of the diffusive term in the thermal forcing (11). It is locally the dominant contribution to surface heat flux and shows a lot of small-scale structure, with a heat loss of up to 150–200 W m^{-2} over the western boundary currents. It enables small-scale structure in the heat flux to occur (e.g., at convection cells), while the SST fields remain smooth. The radiative term, which provides the large-scale fluxes, varies smoothly between a heat input of 35 W m^{-2} in the tropical Pacific to a heat loss in high latitudes that peaks at 40 W m^{-2} in the North Atlantic. Zonal averages of the two terms are shown in Fig. 10.

A point that requires further discussion is the behavior of the diffusive term in convection regions. Figure 11 shows the convection depth at each grid cell. The tendency for convection to occur at separate points rather than in connected areas is strikingly different

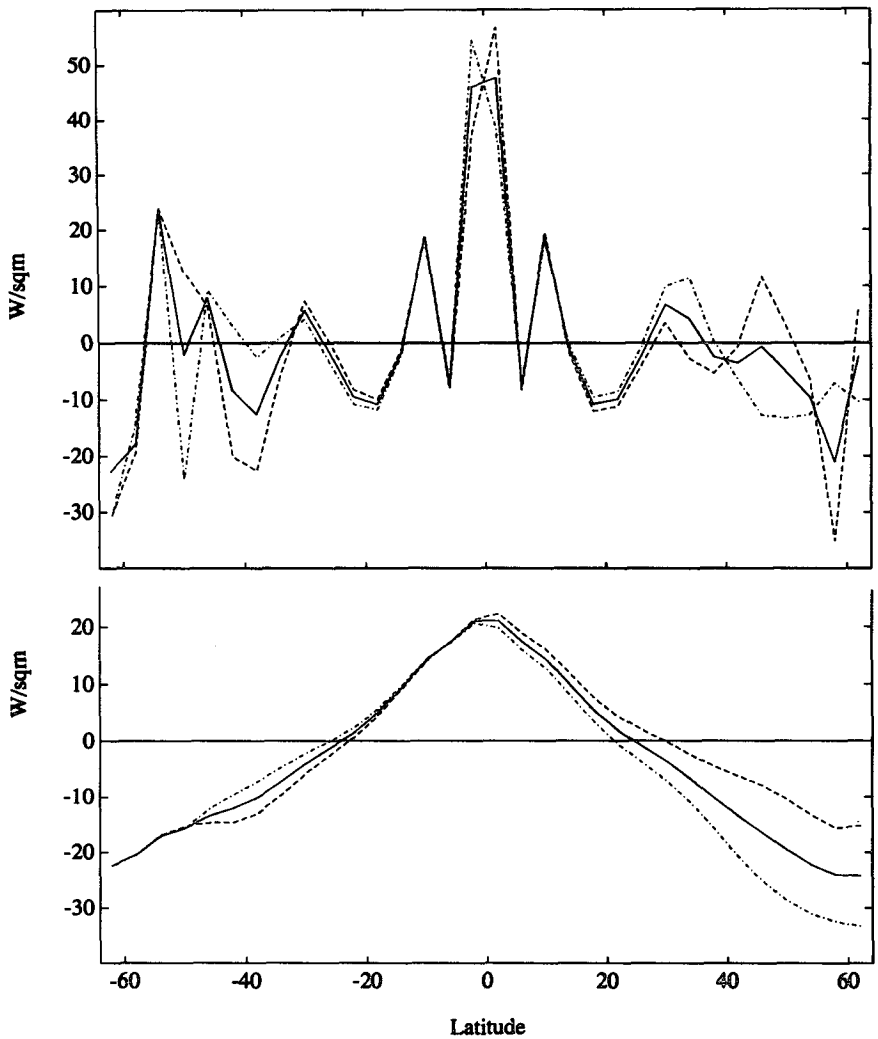

FIG. 10. Zonal average of the surface heat flux in the EB conveyor equilibrium. Upper panel: μ term only [last term of Eq. (11)]. Lower panel: γ term only [first term of Eq. (11)]. Solid line: zonal average, dashed line: Pacific only, dot-dash line: Atlantic only).

from the convection patterns in the HR runs. This is particularly visible at the edge of the Antarctic, where every third grid point convects, and at the northern boundary of the Atlantic, where every other grid point convects. The regularity of this pattern could suggest a computational problem, but analysis of various numerical stability criteria and test runs with different parameter values convinced us that this is not the case; the cause for this behavior is the atmospheric diffusion of heat released by convection. In the regions susceptible to convection, cold freshwater lies above warmer, saltier water. Once convection starts at a particular grid point, mixing the water column, the surface becomes warmer at that point. Consequently, surface heat loss is increased, a positive feedback that locks convection firmly into place once it is going. With the new diffusive boundary condition, the extra heat lost to the atmosphere at this point will not simply "disappear"; it is transported to neighboring grid points. If this heat transport is strong enough, it can suppress convection at neighboring cells, leading to individual convection cells separated by stably stratified cells. The highly regular pattern next to Antarctica is due to the regularity of the coastline, the forcing, and the flow conditions in our idealized model.

We also prepared a conveyor equilibrium using anisotropic diffusion in our new boundary condition (see section 3b), keeping all other parameters and forcing functions unchanged. The main difference to the equilibrium with isotropic diffusion was in the surface temperature distribution and consequently the temperature of the deep water. High latitudes were warmer by about one degree in both basins, and low latitudes were colder by two degrees. This is as expected from the much

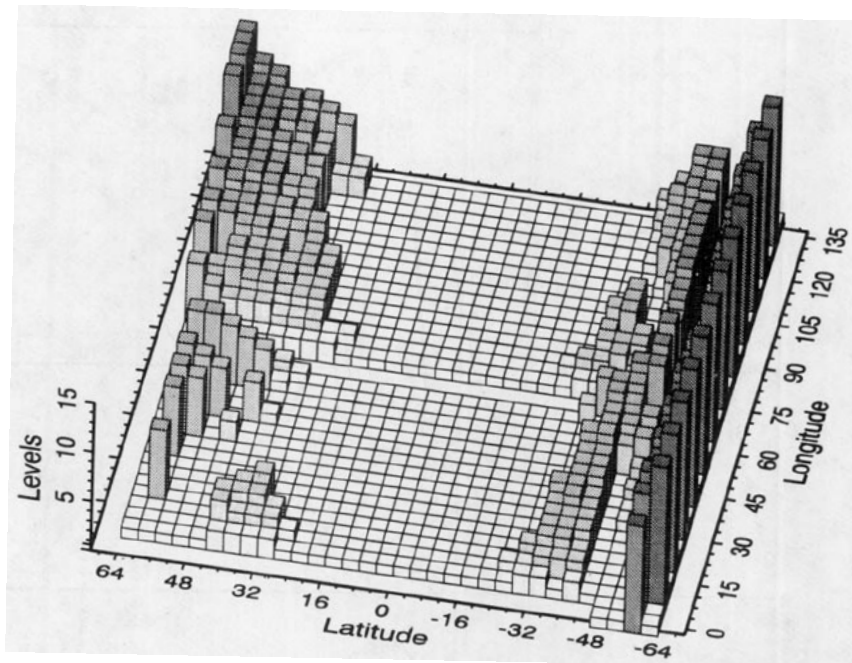

FIG. 11. Convection depth (ventilated levels) in the conveyor equilibrium with the new boundary condition. Convection depth is defined as the depth of the water column that becomes statically unstable and is then vertically mixed by the model's convection scheme during one time step; only convection connected to the surface is included. Strong convective heat flux can lead to a suppression of convection at neighboring points, hence the tendency for convection at single grid cells.

reduced meridional component of diffusion in the atmosphere. The restoring temperature T^* would have to be adjusted to this choice of μ in order to make the surface temperatures more realistic again. The overturning strength of this run, however, was 13.3 Sv—only marginally higher than in the case with isotropic diffusion (12.9 Sv). Salinity fields were hardly affected.

5. Stability of the conveyor belt

The existence of multiple equilibrium states of the ocean circulation raises the question under which conditions transitions between these states can occur. Maier-Reimer and Mikolajewicz (1989) and Marotzke (1990) show experiments where a conveyor state was disturbed by adding freshwater to the northern North Atlantic, leading to a collapse of the thermohaline cell. This collapse, triggered by the shutdown of deep convection and an accumulating freshwater cap in the former deep water formation region, is known as "polar halocline catastrophe" (Bryan 1986).

Power and Kleeman (1994) and Power et al. (1994) confirm that the polar halocline catastrophe can occur in a global ocean model with realistic topography. However, they also demonstrate that it depends on the strength of the thermal coupling to the atmosphere. By varying the restoring parameter λ [Eq. (1)] in their experiments, they show that the polar halocline catastrophe only occurs for very strong thermal coupling to the restoring temperature T^*. This corresponds to the small-scale limit, with suppressed temperature feedback. Similar results were obtained by Zhang et al. (1993), who used their idealized one-hemisphere model to study the extreme case of a fixed transport atmosphere (i.e., the large-scale limit) and found that no polar halocline catastrophe occurs in this case. Moore and Reason (1993) compare strong restoring ($\lambda = 60$ W m^{-2} K^{-1}) with a somewhat weaker thermal coupling based on flux parameterizations, again finding that the thermohaline circulation is more stable in the latter case. It is therefore interesting to investigate this matter with our new coupling, corresponding to an energy balance atmosphere.

To this end, we started with the two conveyor equilibria described in the previous section, that is, the HR equilibrium with the traditional Haney restoring and the EB equilibrium with our new boundary condition. Both were subjected to the same freshwater perturbation: an additional flux of 1.7 m yr^{-1} into the North Atlantic north of 56°N (equal to an input rate of 0.16 Sv) lasting for four years. The total amount of freshwater thus introduced equals 10 times the amount estimated for the great salinity anomaly (GSA) of the 1970s (Dickson et al. 1988). Marotzke (1990) triggered a polar halocline catastrophe by adding less than one GSA *instantaneously*. In contrast, we changed the flux

during an extended time period in order to investigate which *rate* of freshwater inflow will trigger a collapse rather than which salinity change. The total amount of freshwater that needs to be added is in this case larger, since the added freshwater is constantly being mixed down and removed by the ocean circulation as long as convection is still going. Maier-Reimer and Mikolajewicz (1989) achieved a collapse at a much lower rate of freshwater input than ours (0.011 Sv), but the release was concentrated at one point (the mouth of the St. Lawrence River) and lasted longer.

The time evolution of overturning in the Atlantic is shown in Fig. 12. Curve HR1 shows the control run, and as expected the circulation collapses in response to the added freshwater, which caps off the convection. The circulation flips permanently into a southern sinking state (confirmed by asynchronous integration for a further 1500 years). Curve EB shows the experiment with the new thermal coupling. The freshwater perturbation causes a temporary oscillation but no collapse of the thermohaline cell. It is well known that the behavior of the thermohaline circulation can depend strongly on the details of the freshwater forcing, which is different in both cases (Fig. 3). To separate the effects of the different thermal boundary condition and the different freshwater fluxes used, we also performed a model run (starting with the preparation of a conveyor equilibrium) using traditional restoring as in run HR1 but the freshwater flux used for run EB. This is shown in Fig. 12 as curve HR2. In this run the circulation initially collapses as in HR1, but then recovers after a few decades to close to its former strength.

To understand these results, we must look at the corresponding time series of convection in the deep water formation area. As a measure for this we show the convective heat flux through 25-m depth, averaged over the North Atlantic north of 56°N (Fig. 13). This closely mirrors the surface heat flux (not shown), since in this region the dominant balance in the surface layer is between the heat brought up from the deeper layers by convection and the heat lost to the atmosphere. In equilibrium this heat mixed up by convection is brought to the area by the conveyor, of course.

In run HR1, convection stops as a result of the freshwater perturbation. The thermohaline cell has lost its driving force, the buoyancy flux associated with convection, and winds down within a decade. This timescale is probably related to the propagation time of Kelvin and Rossby waves across the North Atlantic (Kawase 1987; Döscher et al. 1994). This allows a southern overturning cell to establish itself permanently, helped by the positive salt feedback discussed earlier, so that the model flips into a southern sinking state.

In run EB, convection never stops. Convective heat flux is reduced because of the added freshwater, but due to the free adjustment of SST under the new boundary condition this leads to an average drop in

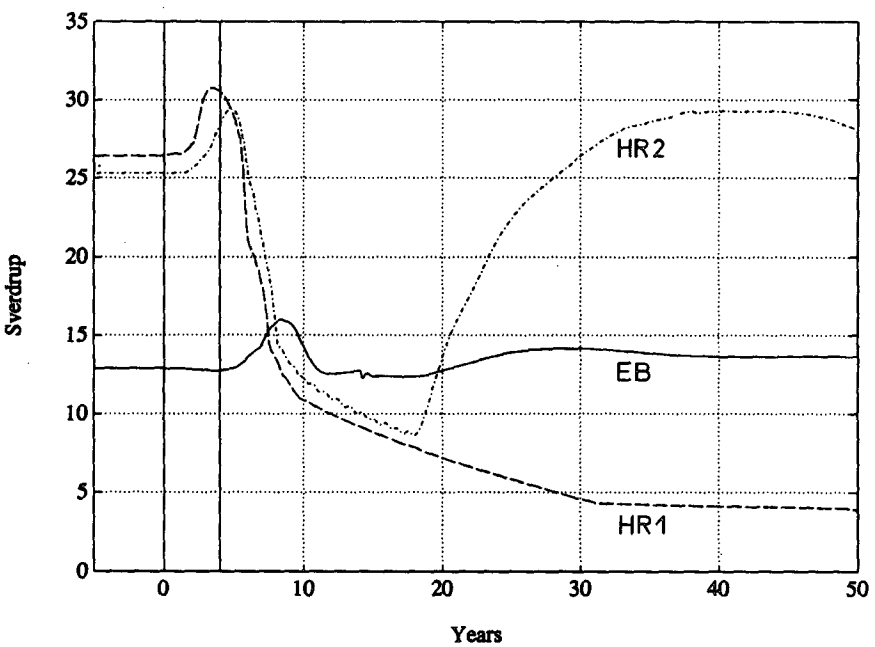

FIG. 12. Time series of overturning in the Atlantic for the runs with 1.7 m yr^{-1} flux perturbation lasting four years. The two vertical lines mark the beginning and end of the addition of freshwater. HR1: Haney restoring with corresponding freshwater flux (see Fig. 3). HR2: Haney restoring with freshwater flux of the EB run. EB: energy balance boundary condition with corresponding freshwater flux.

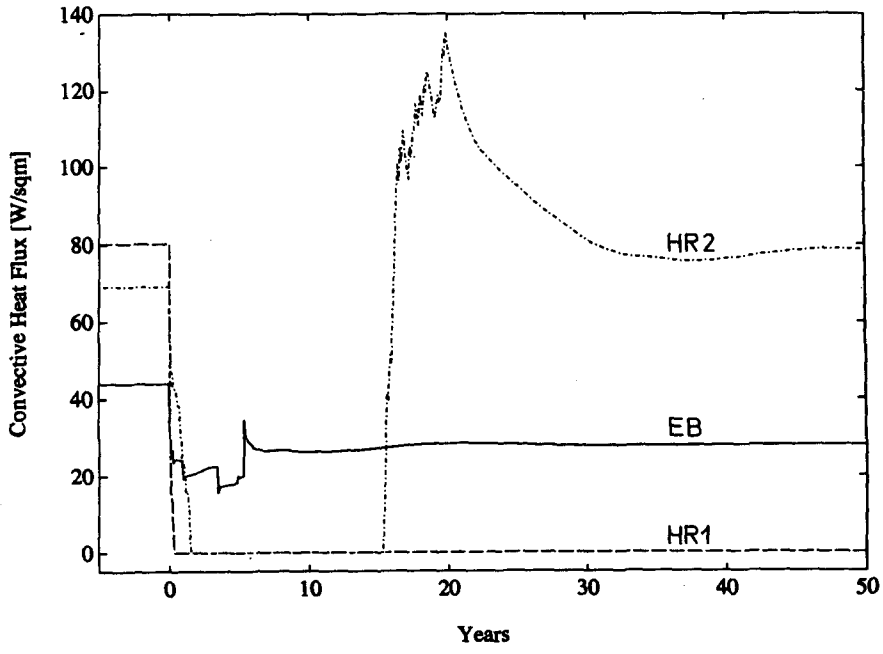

FIG. 13. Time series of convective heat flux across 25-m depth in the northern North Atlantic for the experiments shown in Fig. 12.

surface temperature by 1.2°C (curve not shown), which prevents the surface density from falling below the level where convection is interrupted.

In run HR2, convection is suppressed by freshwater as in HR1, showing that this is due to the strong temperature coupling, not to the difference in freshwater forcing. However, advection slowly warms the deep ocean in the high-latitude Atlantic and disperses the freshwater cap at the surface until convection starts again. We have already seen in section 3b that the freshwater flux used for this run is not conducive to southern sinking. The model therefore cannot use this opportunity of stalled Atlantic overturning to make a transition to a southern sinking state but recovers. The recovery is not to the same equilibrium state as before; like in the model of Hughes and Weaver (1994), we have found multiple conveyor belt states. These are the subject of separate papers (Rahmstorf 1994a,b) and are not discussed further here.

The picture that emerges is that two distinct processes govern the transition between model states. The first is the interruption of convection, caused by the positive feedback mentioned in section 4: once convection is momentarily stopped, surface heat loss and the associated buoyancy flux are reduced, freshwater accumulates at the surface, and convection remains stopped. This interruption acts as a trigger, leading to a collapse of the thermohaline circulation. The second process is a different positive feedback associated with the large-scale horizontal advection of salt. This is the feedback of the box model of Stommel (1961); salt transport to high latitudes enhances deep-water formation, which in turn enhances the salt transport (see section 6). This feedback, together with the thermohaline forcing, determines which equilibria are possible, which are preferred, and which one will emerge after the circulation was temporarily stalled in one basin due to interrupted convection.

A different kind of transition between equilibria is also possible—not triggered by interrupted convection but by a moderate *permanent* change in freshwater forcing. We can increase precipitation in the Northern Hemisphere to the point where an equilibrium with northern deep-water formation cannot be sustained because the advective salt feedback is not strong enough to outweigh the extra precipitation. Such a precipitation change need not be so strong as to interrupt convection. This kind of transition occurs if the EB conveyor is subjected to the HR freshwater forcing; the conveyor circulation then slowly winds down over a period of 1000 years, giving way to a southern sinking state (details of these experiments will be reported elsewhere). This kind of slow transition was also observed in a recent coupled global warming experiment of Manabe and Stouffer (1993) after the model atmosphere was subjected to a gradual quadrupling of CO_2.

An interesting question is whether convection can also be interrupted with the new thermal boundary condition by applying a larger freshwater perturbation than that in the experiments discussed so far in this section. A series of EB runs with progressively larger additions of freshwater to the northern North Atlantic is shown in Fig. 14. For very large freshwater input we do indeed get a temporary collapse of the Atlantic deep

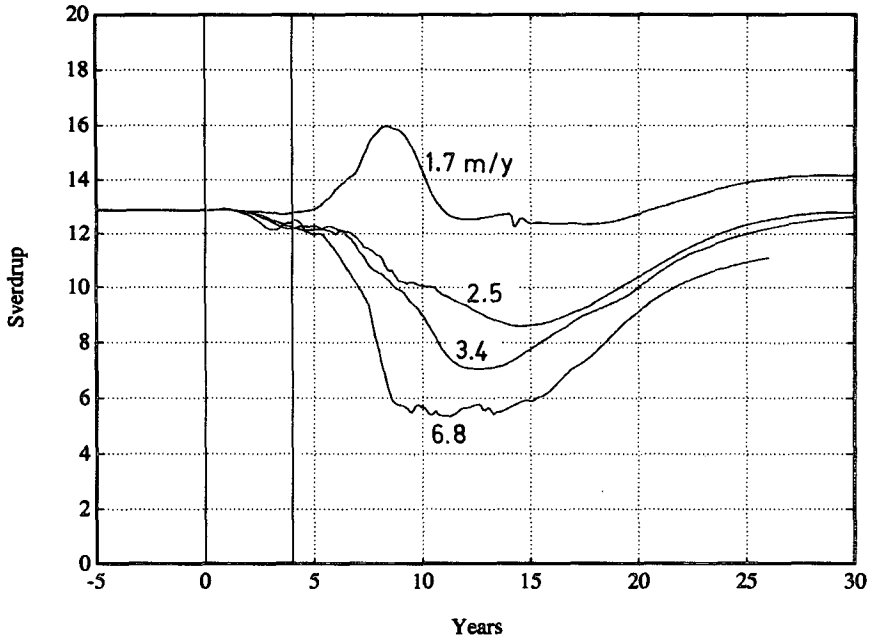

FIG. 14. Time series of overturning in the Atlantic after freshwater perturbations of different strengths (labeled in m yr^{-1}) with the new thermal boundary condition (EB).

circulation like in run HR2. This is associated with a drop in average surface temperature in the northern North Atlantic by 4°C. Similar experiments with anisotropic diffusion (as discussed at the end of section 4) show basically the same sensitivity, only the details of the transient behavior differ.

A simple box model (Fig. 15) can be used to estimate how the threshold needed to interrupt convection depends on the thermal coupling strength. The surface layer is restored to temperature T^* (in which we could include some heating by horizontal advection) and receives a freshwater flux F_S. The heat balance of this layer is between this coupling to T^* and convective coupling to the deep ocean, represented by a reservoir of constant temperature T_0: $M(T_0 - T) = -\lambda(T^* - T)$. Storage can be neglected if the layer is thin. The convective coupling strength M will be some function of the density difference $(\rho - \rho_0)$ between the surface layer and the deep ocean, with $M = 0$ if this difference becomes negative. While convection is going, we assume $M \gg \lambda$, since convection is a very efficient stirring mechanism, so that $T \approx T_0$ and $S \approx S_0$. How much surface freshening can this simple system sustain before convection stops? Convection ceases when $\rho = \rho_0$, or $\beta\Delta S = \alpha(T - T_0)$, where ΔS is the salinity change of the surface layer, and α and β are the thermal and haline expansion coefficients. When convection stops, the surface temperature T drops to T^*, and we see that the salinity change ΔS needed to interrupt convection is proportional to the difference between the temperature of the deep water that is being formed and the restoring temperature. For a given rate of heat loss Q in the deep-water formation region (which is related to the poleward heat transport of the conveyor), the critical salinity change is

$$\Delta S \approx \frac{\alpha}{\beta}\frac{Q}{\lambda},$$

that is, inversely proportional to the coupling strength λ. This explains why with our weaker, more realistic

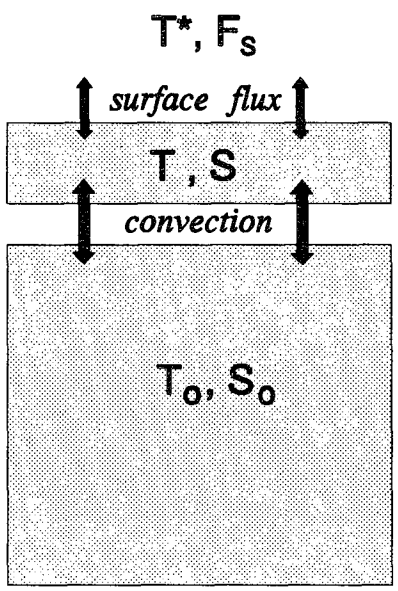

FIG. 15. Simple box model of convective mixing.

thermal coupling the amount of freshwater needed to trigger a polar halocline catastrophe is larger than with traditional Haney restoring, in spite of the conveyor (and thus Q) in our HR run being more intense, which by itself would make it more stable than the EB conveyor.

This result raises the question of how strong the "effective" coupling of our new boundary condition (11) is. Obviously the diffusive term depends on the scale of the process; but we can give an estimate for the behavior of the northern North Atlantic region during the perturbation experiments described in this chapter. Regression curves of average heat flux versus average SST north of 56°N were plotted (not shown here). In the EB experiment shown in Fig. 12, SST dropped by 1.2°C during the freshwater input, while heat loss was reduced by 25 W m^{-2}, indicating an effective coupling of $\lambda \approx 12$ W m^{-2} K^{-1}. The regression was not smooth, however; periods with a slope of 20 W m^{-2} K^{-1} were interrupted by brief times with negative slope due to the complex reordering of convection occurring both inside and outside the averaging region. For a four times stronger perturbation (6.8 m yr^{-1}, Fig. 14), where convection is widely interrupted, a good straight regression line with a slope of 9 W m^{-2} K^{-1} was obtained. These numbers contrast with the extreme case of fixed atmospheric transport considered by Zhang et al. (1993), which used a fixed coupling of 2.3 W m^{-2} K^{-1}.

6. Discussion and conclusions

Forcing an ocean model using some form of parameterization of the surface heat and freshwater fluxes always implies a certain model of an atmosphere coupled to the ocean through these fluxes. For example, using Haney-type temperature restoring implies an atmosphere of constant temperature. Therefore, this type of boundary condition is unsuitable for climatic change experiments involving large-scale changes in oceanic heat transport. If large-scale changes in the ocean circulation occur, they can have a marked effect on surface temperature, which is suppressed if fixed atmospheric temperature is assumed. In this paper we have shown that these changes in surface temperature are crucial for the behavior of the thermohaline circulation and cannot be neglected.

To allow a more realistic adjustment of surface temperatures to changes in the ocean circulation, we have proposed the thermal boundary condition (11). We have shown that this implies an atmosphere whose temperature is not fixed, but which obeys a simple heat budget. This heat budget allows for lateral diffusion of heat received from the ocean and is further characterized by a "forcing field" T^* and a radiative relaxation parameter γ, both of which can be derived from atmospheric general circulation models. This boundary condition is as easily implemented in any ocean model as the traditional restoring approach and makes a negligible contribution to processing time.

To test the new boundary condition we performed a series of experiments with a two-basin ocean circulation model, comparing runs with the traditional restoring and the new diffusive thermal coupling. An important success is the simulation of the temperature difference between North Atlantic and North Pacific high latitudes—caused entirely by the heat transport of the thermohaline conveyor belt since the geometry and forcing of the two model basins were identical. This corresponds to the simulation of two different "climates," one with deep-water formation and one without, and therefore illustrates the suitability of the new boundary condition for climate studies involving changes in oceanic heat transport.

The free adjustment of surface temperatures allowed by the new boundary condition also leads to low temperatures in the model Southern Ocean and the penetration of Antarctic bottom water into the Atlantic basin underneath the North Atlantic deep water, an important feature absent in the HR runs with the traditional restoring approach. Of course, with Haney restoring we could *specify* cold Antarctic temperatures, but in our model they are predicted as a result of the dynamics of the ocean circulation. It is perhaps surprising that such a simple model, which does not include the effects of topography (or even basin shape), ice formation, or continental influence, qualitatively displays these major features of the oceanic deep circulation.

Of particular interest are the experiments concerning the stability of the conveyor belt circulation. Our model runs demonstrate that two distinct mechanisms are important for instabilities and state transitions of the thermohaline circulation. The first is the vertical convection feedback (Fig. 15). Interruption of convection acts as a trigger for the sudden collapse of the deep circulation within a few years, a timescale reminiscent of the rapid changes observed in the Greenland ice cores. The second mechanism is the balance between the horizontal salt and heat advection feedbacks, which determines the overturning rate and leads to the existence of qualitatively different equilibrium states (southern sinking, conveyor, northern sinking) and conveyor states of different intensity. Both mechanisms are strongly affected by the free adjustment of surface temperatures allowed by our new thermal boundary condition. On one hand, the *advective* temperature feedback acts to regulate the strength of the conveyor. On the other hand, the *convective* temperature feedback makes it harder to interrupt convection and trigger a polar halocline catastrophe. We conclude that due to the temperature feedbacks, the conveyor circulation is more stable than previous model studies suggested. It is now generally recognized that these earlier models were too sensitive. For example, in the model of Marotzke (1990) the deep circulation collapsed after a

freshwater perturbation smaller than the great salinity anomaly of the 1970s.

Some authors have also studied internal oscillations of the thermohaline circulation and its response to stochastic freshwater forcing (e.g., Mikolajewicz and Maier-Reimer 1990; Weaver et al. 1991; Mysak et al. 1993; Winton and Sarachik 1993). These oscillations are not as dramatic as the instabilities discussed above but are nevertheless important—for example, for the detection of the greenhouse warming signal. We have not investigated the effect of our new boundary condition on this type of variability, but we expect that the temperature feedback would act to dampen these oscillations. Recent model experiments by Mikolajewicz and Maier-Reimer (1994) with reduced thermal coupling confirm this.

Our work highlights the importance of the stabilizing temperature feedbacks for the deep circulation and the role of deep convection as "Achilles heel" (Broecker 1991) of the conveyor. Just how vulnerable the conveyor is remains an open question. We have only considered the temperature feedback in this paper; the freshwater fluxes remained fixed. Nakamura et al. (1993) have suggested that a destabilizing feedback exists between the thermohaline circulation and atmospheric freshwater transports. To get a better idea how robust the convection and deep-water formation processes are in a changing climate, experiments with a high-resolution model of the northern North Atlantic and Arctic would be desirable, including realistic seasonal forcing and the effects of ice and bottom topography.

While our new boundary condition cannot replace a fully coupled ocean–atmosphere model, it approximates the heat balance of the atmosphere at low computational cost and allows qualitative study of climatic processes with ocean circulation models.

Acknowledgments. SR wishes to thank Sybren Drijfhout, Matthew England, Reindert Haarsma, Tertia Hughes, Geert Lenderink, Jochem Marotzke, and Peter Stone for stimulating discussions relating to this work. This work was supported through Sonderforschungsbereich 133. The model experiments were performed at the German Climate Computer Center in Hamburg.

APPENDIX

The Heat Budget of the Coupled System

In this appendix we discuss the heat budget considered in section 2b more fully. To close the heat balance [Eq. (2) and (3)] we want to express all terms as functions of T_A and T_O only. We then have two equations for two unknowns, the air temperature T_A and the surface heat flux Q; T_O is provided by the ocean model.

One way of treating the heat budget components is as follows. We can prescribe the net solar input (Q_S and f) at each point. Air–sea exchange (i.e., sensible plus latent heat flux) can be parameterized as $Q_C = c(T_O - T_A)$. The oceanic back radiation can be calculated as $Q_B = \epsilon\sigma T_O^4$, or in linearized form $Q_B = \epsilon\sigma T_{\text{ref}}^4 + 4\epsilon\sigma T_{\text{ref}}^3(T_O - T_{\text{ref}})$. Using $\sigma = 5.73 \times 10^{-8}$ W m^{-2} K^{-4}, $\epsilon = 0.98$, and $T_{\text{ref}} = 273$ K yields $r = 4\epsilon\sigma T_{\text{ref}}^3 = 4.6$ W m^{-2} K^{-1}. On average ninety-five percent of the ocean's back radiation is absorbed by the atmosphere ($e = 0.95$) so that the value of r is unimportant since air–sea exchange is dominated by c, which is an order of magnitude larger (see below).

We assume that the atmospheric longwave terms depend linearly on T_A: $Q_D = a\sigma T_{\text{ref}}^4 + d(T_A - T_{\text{ref}})$ and $Q_U = b\sigma T_{\text{ref}}^4 + d(T_A - T_{\text{ref}})$. The temperature sensitivity d of the outgoing longwave radiation is an important parameter, closely linked to the climate sensitivity at the center of the global warming debate, which determines how much the equilibrium temperature of the earth will change in response to a change in radiative forcing. (A minor difference to our d is due to the fact that we treat longwave radiation escaping to space directly from the ocean surface separately.) The IPCC (Houghton et al. 1991) gives a range of 1–3 W m^{-2} K^{-1} for this parameter. We use the same sensitivity for the downwelling radiation, but this is not important; downwelling radiation is part of the ocean–atmosphere coupling dominated by c. The factors a and b are empirical, reflecting the emissivity and temperature of the atmosphere at the altitudes where the radiation originates. From the global radiation budget given by Gill (1982), we estimate the following values for these parameters (for $T_{\text{ref}} = 273$ K, $d = 3$ W m^{-2} K^{-1}): $a = 0.85$, $b = 0.55$.

This approach leads to the linearized heat budget [Eq. (4) and (5)], with the constants A–F shown in Table 1. The constants needed for these parameterizations (a–f) can depend, for example, on latitude and season but are assumed to be unaffected by the ocean circulation.

The strongest temperature dependence arises from the sensitivity c of the air–sea heat exchange. The factor c can be derived from the bulk parameterization formulas commonly used to compute air–sea heat flux, and it depends on wind speed, temperature, and humidity. Using the bulk formulas given by Busch (1977) for sensible heat and by Sill (1983) for latent heat and typical values for wind speed, temperature, and humidity ($U_{10} = 8$ m s^{-1}, $T = 18$°C, $h = 80\%$), we obtain a sensitivity of 15 W m^{-2} K^{-1} for sensible heat and 28 W m^{-2} K^{-1} for latent heat; that is, $c \approx 43$ W m^{-2} K^{-1}. It is proportional to wind speed and increases strongly for warmer temperatures and lower humidities. It is important that c is an order of magnitude larger than the radiative sensitivities, but its actual value is only important for the small-scale response of the coupled system.

It is instructive to consider the global average of the heat budget (4) and (5). In this case Q_A and Q vanish

(assuming equilibrium). Using values from Gill again, the net solar radiation received at the top of the atmosphere is $Q_S = 240$ W m^{-2}, 27% of which are absorbed by the atmosphere and 73% by the ocean (i.e., $f = 0.27$). The resulting air temperature is 15.1°C, and the sea surface temperature is 17.4°C. The individual budget terms match Gill's values to within a few watts per square meter. The magnitude of the air–sea temperature difference is of course dependent on c; for a stronger coupling it decreases.

In the final boundary condition (11), the behavior of the coupled system is condensed into three parameters, where γ gives the global climate sensitivity, T^* represents the climatic forcing field, and μ determines the lateral heat dispersion characteristics. From the derivation we see that

$$\gamma = E - \frac{FB}{C} = r + c - (er + c)\frac{d + c}{2d + c}. \quad (12)$$

Since $c \gg r$, $c \gg d$, and $(1 - e) \ll 1$ this is approximately $\gamma \approx d$ as discussed above. In physical terms: since the thermal coupling to the atmosphere is strong and only a small percentage of the ocean's back radiation escapes directly to space, the sensitivity γ is essentially the temperature dependence of the *atmosphere's* longwave emission to space. For the case considered by Haney, T_A = const, Eq. (5) gives $\gamma = E = r + c$, the sum of the sensitivity of oceanic longwave radiation, and air–sea exchange. As we have seen above, this is of the order of 50 W m^{-2} K^{-1}, more than an order of magnitude larger than d.

The field of T^* can in principle be calculated from (9). Note that (11) only requires that temperature *deviations* from T^* are dispersed in a diffusive manner; to determine the climatic forcing field T^*, we can use a more realistic representation of heat transport in Eq. (9). For example, we can split the heat transport in a component corresponding to present day temperatures and a diffusive ansatz only for the remainder and then use the observed atmospheric heat transport. Unfortunately, the atmospheric heat transport is not known to sufficient accuracy (Carissimo et al. 1985; Peixoto and Oort 1992). The uncertainty in its divergence is likely to exceed 30 W m^{-2}, which would lead to an error in T^* of 10°C. Alternatively, an atmospheric general circulation model run could be used, where the model is bounded at the bottom by a zero heat capacity "swamp." For our idealized two-basin model we decided to use a simple cosine function of latitude, which we fitted to give realistic sea surface temperatures (see section 3).

In summary, the boundary condition (11) provides us with an energy balance model in a nutshell. Neither the global sensitivity γ nor the forcing field T^*, which characterize this model, depend on the simplifying assumptions we made about the atmosphere. Instead they can both be derived from sophisticated atmospheric general circulation models. The major simplification we have made to the atmospheric response lies in the treatment of horizontal transport, where we have assumed a diffusive dispersion of temperature anomalies.

REFERENCES

Anderson, D. L. T., and J. Willebrand, 1992: Recent advances in modelling the ocean circulation and its effects on climate. *Rep. Prog. Phys.*, **55**, 1–37.

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.

Bretherton, F. P., 1982: Ocean climate modeling. *Progress in Oceanography*, Vol. 11, Pergamon, 93–129.

Broecker, W. S., 1991: The great ocean conveyor. *Oceanography*, **4**, 79–89.

——, T. H. Peng, J. Jouzel, and G. Russell, 1990: The magnitude of global fresh-water transports of importance to ocean circulation. *Clim. Dyn.*, **4**, 73–79.

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

Bryan, K., 1984: Accelerating the convergence to equilibrium of ocean-climate models. *J. Phys. Oceanogr.*, **14**, 666–673.

Budyko, M. I., 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, **21**, 611–619.

Busch, N., 1977: Fluxes in the surface boundary layer over the sea. *Modelling and Prediction of the Upper Layers of the Ocean*, Pergamon, 72–91.

Carissimo, B. C., A. H. Oort, and T. H. Vonder Haar, 1985: Estimating the meridional energy transports in the atmosphere and ocean. *J. Phys. Oceanogr.*, **15**, 82–91.

Cox, M. D., 1984: A primitive equation, 3-dimensional model of the ocean. GFDL Ocean Group Tech. Rep. 1, GFDL, Princeton University, Princeton, NJ, 143 pp.

Cubasch, U., K. Hasselmann, H. Höck, E. Maier-Reimer, U. Mikolajewicz, B. D. Santer, and R. Sausen, 1993: Time-dependent greenhouse warming computations with a coupled ocean–atmosphere model. *Clim. Dyn.*, **8**, 55–69.

Dansgaard, W., S. J. Johnsen, H. B. Clausen, N. S. Dahl-Jensen, N. S. Gundestrup, C. U. Hammer, C. S. Hvidberg, J. P. Steffensen, A. E. Sveinbjörnsdottir, J. Jouzel, and G. Bond, 1993: Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, **364**, 218–220.

Dickson, R. R., J. Meincke, S. A. Malmberg, and A. J. Lee, 1988: The "Great Salinity Anomaly" in the northern North Atlantic, 1968–82. *Progress in Oceanography*, Vol. 20, Pergamon, 103–151.

Döscher, R., C. W. Böning, and P. Herrmann, 1994: Response of circulation and heat transport in the North Atlantic to changes in thermohaline forcing in northern latitudes: a model study. *J. Phys. Oceanogr.*, **24**, 2306–2320.

England, M. H., 1993: Representing the global-scale water masses in ocean general circulation models. *J. Phys. Oceanogr.*, **23**, 1523–1552.

——, J. S. Godfrey, A. C. Hirst, and M. Tomczak, 1993: The mechanism for Antarctic Intermediate Water renewal in a World Ocean model. *J. Phys. Oceanogr.*, **23**, 1553–1560.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press, 662 pp.

——, and K. Bryan, 1971: Effects of geometry on the circulation of a three-dimensional Southern Hemisphere ocean model. *Deep-Sea Res.*, **18**, 685–721.

Gordon, A. L., 1986: Interocean exchange of thermocline water. *J. Geophys. Res.*, **91**, 5037–5046.

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

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

Hasselmann, K., 1991: Ocean circulation and climate change. *Tellus,* **43,** 82–103.

Henderson-Sellers, B., 1986: Calculating the surface energy balance for lake and reservoir modeling: A review. *Rev. Geophys.,* **24,** 625–649.

Houghton, J. T., G. J. Jenkins, and J. J. Ephraums, Eds., 1990: *Climate Change. The IPCC Scientific Assessment.* Cambridge University Press, 365 pp.

Hughes, T. M. C., and A. J. Weaver, 1994: Multiple equilibria of an asymmetric two-basin ocean model. *J. Phys. Oceanogr.,* **24,** 619–634.

Kawase, M., 1987: Establishment of deep circulation driven by deep water production. *J. Phys. Oceanogr.,* **17,** 2294–2316.

Keigwin, L. D., G. A. Jones, S. J. Lehman, and E. A. Boyle, 1991: Deglacial meltwater discharge, North Atlantic deep circulation, and abrupt climate change. *J. Geophys. Res.,* **96,** 16 811–16 826.

Levitus, S., 1982: *Climatological Atlas of the World Ocean.* NOAA Prof. Paper 13, Vol. 13, U.S. Dept. of Commerce, NOAA, Washington, D.C., 173 pp.

Maier-Reimer, E., and U. Mikolajewicz, 1989: Experiments with an OGCM on the cause of the Younger Dryas. *Oceanography,* UNAM Press, 87–100.

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

Marotzke, J., 1990: Instabilities and multiple equilibria of the thermohaline circulation. Ph.D. thesis, Christian-Albrechts-Universität Kiel, 126 pp.

——, and J. Willebrand, 1991: Multiple equilibria of the global thermohaline circulation. *J. Phys. Oceanogr.,* **21,** 1372–1385.

Mikolajewicz, U., and E. Maier-Reimer, 1990: Internal secular variability in an ocean general circulation model. *Clim. Dyn.,* **4,** 145–156.

——, and ——, 1994: 'Mixed boundary conditions' in OGCMs and their influence on the stability of the model's conveyor belt. Report No. 130, MPI, Hamburg, 24 pp.

Moore, A. M., and C. J. Reason, 1993: The response of a global ocean general circulation model to climatological surface boundary conditions for temperature and salinity. *J. Phys. Oceanogr.,* **23,** 300–327.

Mysak, L. A., T. F. Stocker, and F. Huang, 1993: Century-scale variability in a randomly forced, two-dimensional thermohaline ocean circulation model. *Clim. Dyn.,* **8,** 103–116.

Nakamura, M., P. H. Stone, and J. Marotzke, 1993: Destabilization of the thermohaline circulation by atmospheric feedback. Center for Global Change Science, Massachusetts Institute of Technology, Report 22, 14 pp.

Pacanowski, R., K. Dixon, and A. Rosati, 1991, 1993: The GFDL modular ocean model users guide. GFDL Ocean Group Tech. Report 2, GFDL.

Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate.* American Institute Physics, 520 pp.

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

——, A. M. Moore, D. A. Post, N. R. Smith, and R. Kleeman, 1994: Stability of NADW formation in a global ocean general circulation model. *J. Phys. Oceanogr.,* **24,** 904–916.

Rahmstorf, S., 1991: A zonal-averaged model of the ocean's response to climatic change. *J. Geophys. Res.,* **96,** 6951–6963.

——, 1993: A fast and complete convection scheme for ocean models. *Ocean Model.,* **101,** 9–11.

——, 1994a: Rapid climate transitions in a coupled ocean–atmosphere model. *Nature,* **372,** 82–85.

——, 1994b: Multiple convection patterns and thermohaline flow in an idealized OGCM. *J. Climate,* submitted.

Schopf, P. S., 1983: On equatorial waves and El Niño. II: Effects of air–sea thermal coupling. *J. Phys. Oceanogr.,* **13,** 1878–1893.

Sill, B. L., 1983: Free and forced convection effects on evaporation. *J. Hydraul. Eng.,* **109,** 1216–1231.

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

Taylor, K. C., G. W. Lamoray, G. A. Doyle, R. B. Alley, P. M. Grootes, P. A. Mayewski, J. W. C. White, and L. K. Barlow, 1993: The "flickering switch" of late Pleistocene climate change. *Nature,* **361,** 432–435.

Weaver, A. J., and T. M. C. Hughes, 1992: Stability and variability of the thermohaline circulation and its link to climate. *Trends in Physical Oceanography,* Research Trends Series, Council of Scientific Research Integration, 56 pp.

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

——, J. Marotzke, P. F. Cummins, and E. S. Sarachik, 1993: Stability and variability of the thermohaline circulation. *J. Phys. Oceanogr.,* **23,** 39–60.

Willebrand, J., 1993: Forcing the ocean by heat and freshwater fluxes. *Energy and Water Cycles in the Climate System,* NATO ASI Series, NATO, 215–233.

Winton, M., and E. S. Sarachik, 1993: Thermohaline oscillations induced by strong steady salinity forcing of ocean general circulation models. *J. Phys. Oceanogr.,* **23,** 1389–1410.

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.