Mechanisms of Thermohaline Mode Switching with Application to Warm Equable Climates

RONG ZHANG,* MICHAEL FOLLOWS, AND JOHN MARSHALL

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Manuscript received 10 August 2001, in final form 12 December 2001)

ABSTRACT

A three-box model of haline and thermal mode overturning is developed to study thermohaline oscillations found in a number of ocean general circulation models and that might have occurred in warm equable paleoclimates. By including convective adjustment modified to represent the localized nature of deep convection, the box model shows that a steady haline mode circulation is unstable. For certain ranges of freshwater forcing/vertical diffusivity, a self-sustained oscillatory circulation is found in which haline–thermal mode switching occurs with a period of centuries to millennia. It is found that mode switching is most likely to occur in warm periods of earth's history with, relative to the present climate, a reduced Pole–equator temperature gradient, an enhanced hydrological cycle, and somewhat smaller values of oceanic diffusivities.

1. Introduction

In past climates the thermohaline circulation (THC) of the ocean may have been quite different from that of today. Paleoclimatic records (e.g., Railsback et al. 1990) suggest that in warm periods of earth's history the abyssal ocean was very much warmer than that of today. A recurring theme of the paleoclimatic literature is the speculation that in these warm climates ocean deep water formation could have been triggered by evaporation from the subtropics resulting in a ''haline mode'' accounting for abnormal warmth in the subsurface ocean (e.g., Brass et al. 1982). This is very different from today's climate in which deep water formation at high latitudes brings cold water to depth in a ''thermal mode.'' Such extremes of ocean circulations have very different implications for climate and biogeochemical cycles [see, e.g., Zhang et al. (2001)].

In certain parameter regimes, ocean general circulation models (OGCMs) exhibit ''mode switching'' in which the meridional overturning circulation (MOC) switches between thermal and haline modes. Self-sustained thermohaline oscillations have been found in several OGCM studies using an idealized single basin configuration (Marotzke 1989; Wright and Stocker 1991; Weaver and Sarachik 1991a,b; Weaver et al. 1993; Winton and Sarachik 1993; Huang 1994). In a study of possible modes of the late Permian ocean circulation with a coarse-resolution OGCM (Zhang et al. 2001), we found that the haline mode (HM) was inherently unstable for fixed external forcing, periodically switching in to a transient thermal mode (TM) in which deep water formed in polar regions (''flushing events''; Marotzke 1989), and gradually returning to the HM to close the limit cycle. Such internal thermohaline oscillations might have significant implications for understanding the paleoclimatic record, such as the centuries to millennia oscillations during glacial periods (Johnson et al. 1992). Such oscillations do not appear to occur in the modern ocean, because, apparently, the surface freshwater forcing is not strong enough. Mode switching is more likely to occur, perhaps, during glacial periods in which the freshwater forcing due to ice melting at polar regions is much stronger, or during warm equable paleoclimates such as the late Permian, or mid-Cretaceous in which the buoyancy forcing due to freshwater flux may have been stronger than the air–sea heat flux.

The physical mechanism underlying such thermohaline oscillations has yet to be clearly articulated. Stommel (1961) first showed, using a highly idealized box model, that the THC can have multiple steady states when the freshwater forcing is strong enough: he found a strong stable TM circulation, a weak unstable TM circulation, and a stable HM circulation. Since Stommel's (1961) pioneering work, many box models have been constructed to study the multiple steady solutions of the THC (e.g., Rooth 1982; Huang et al. 1992). Without convective adjustment, Stommel-type box models

* Current affiliation: Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey

Corresponding author address: Dr. Rong Zhang, Geophysical Fluid Dynamics Laboratory, P.O. Box 308, Princeton University, Princeton, NJ 08542
E-mail: rzhang@splash.princeton.edu

cannot support self-sustained oscillations (Ruddick and Zhang 1996). Welander (1982) proposed a heat–salt oscillator using a model with convective adjustment between the surface and deep ocean whose temperature and salinity were fixed. It exhibits self-sustained oscillations only when warm salty water convects over cold freshwater. This kind of convection is likely to happen in low latitudes, but not in polar regions. Winton (1993) modified this model by fixing the surface temperature but allowing deep ocean temperature to vary. He obtained a self-sustained oscillation with polar convection when a nonlinear equation of state for seawater was used. Pierce et al. (1995) also obtained a self-sustained oscillation by modifying Welander's convection model to allow both surface and deep ocean temperatures and salinity to vary, and again using a nonlinear equation of state for density.

In this study, we construct a simple box model with convective adjustment and assume a linear equation of state. It combines the Stommel-type box model with the Welander-type convection model. The model captures the main character and essential physics of the thermohaline oscillation and instability exhibited in our model of the late Permian ocean circulation. When combined with our OGCM studies we are able to explain why the steady HM becomes unstable in a certain range of freshwater forcing and vertical diffusivity amplitude. We then go on to use the box model as an economical tool to explore parameter space, identify regimes and their stability, and the dependence of oscillation period to freshwater flux forcing and vertical diffusivity.

In section 2, we describe in detail the millennial thermohaline oscillation in a global OGCM configured with late Permian bathymetry. In section 3, we discuss the mechanism of such oscillations using a simple three-box model.

2. Self-sustained thermohaline oscillation in an OGCM of late Permian ocean

We briefly review thermohaline mode switching observed in certain parameter regimes of a global, coarse-resolution OGCM (Marshall et al. 1997a,b) configured for late Permian bathymetry (Fig. 1a; see appendix A; Zhang et al. 2001). The late Permian (about 250 million years ago) is thought to have been a period of warm equable climate (Taylor et al. 1992). In our model a quasi-steady HM circulation occurred with enhanced (relative to the modern) freshwater flux forcing [a maximum evaporation $-$ precipitation ($E - P$) of 1.3 m yr^{-1} compared to 0.6 m yr^{-1} in the present climate] and a relatively weak background vertical mixing in the ocean of magnitude $\hat{M} = 3 \times 10^{-5}\ \text{m}^2\ \text{s}^{-1}$. This should be compared to the canonical, global average value of about $5 \times 10^{-5}\ \text{m}^2\ \text{s}^{-1}$, required to bring into consistency deep water formation rates and the mean temperature structure of the modern ocean in OGCMs. In reality, oceanic diapycnal mixing is thought to be spatially inhomogeneous becoming larger near boundaries where tidally induced mixing processes may dominate (Toole et al. 1994; Marotzke 1997; Munk and Wunsch 1998).

Figure 1b shows the meridional overturning circulation of the model during the quasi-steady HM. As described in appendix A, mixed boundary conditions and a full nonlinear equation of state are used. The overturning is weak and shallow; warm, salty intermediate water is formed in the subtropics, then returns to the surface in polar regions and the Tropics. Figure 1c illustrates the meridional overturning circulation during a transient TM (a ''flushing'' event), which is strong and deep; deep water formed in the southern polar region upwells in the Tropics and Northern Hemisphere. Here the transient TM circulation shows strong hemispheric asymmetry, that is, a relative warm SST in southern high latitudes is associated with strong deep convection there and coexists with a relatively cool SST in the northern high latitudes where deep convection is absent. This result is very similar to the recent study by Haupt and Seidov (2001), in which for asymmetric surface thermal forcing, a strong asymmetric ocean circulation can sustain a warm abyss through deep convection while keeping the other Pole cool. Since here we have symmetric surface forcing, the asymmetric circulation obtained is a consequence of the asymmetric distribution of land and sea. Both studies show that the thermal mode overturning circulation is sensitive to the high-latitude freshwater flux and that a symmetric circulation with deep convection in both hemispheres is difficult to sustain.

The onset of the flushing event is triggered by intense local convection in the southern polar region, induced by meridional transports of warm salty surface water by a large eddy. Figure 2 shows the sea surface temperature (SST) and salinity at year 4155, just before the switch from HM to transient TM. A large-scale eddy, warmer and saltier than the ambient polar surface water, is formed near the western boundary of the superocean in the southern polar region, just where strong local convection occurs. The eddy propagates toward the eastern boundary of the superocean at the latitude of strong deep convection in the TM and, we believe, plays a role in flipping the circulation into a strong transient TM. When the convection region becomes less dense due to precipitation, convection ceases and the circulation returns to the HM.

The transient TM is asymmetric, with deep convection occurring at southern high latitudes. Hence, we focus our diagnostics on the southern ocean of the model. We divide the southern ocean into three distinct regions over which we consider time series of average properties. These regions are low-latitude surface (14°–36.6°S, 0–50 m), high-latitude surface (36.6°–70.3°S, 0–50 m), and deep ocean (14°–70.3°S, 50–4000 m). In section 3 we will describe a simple three-box model

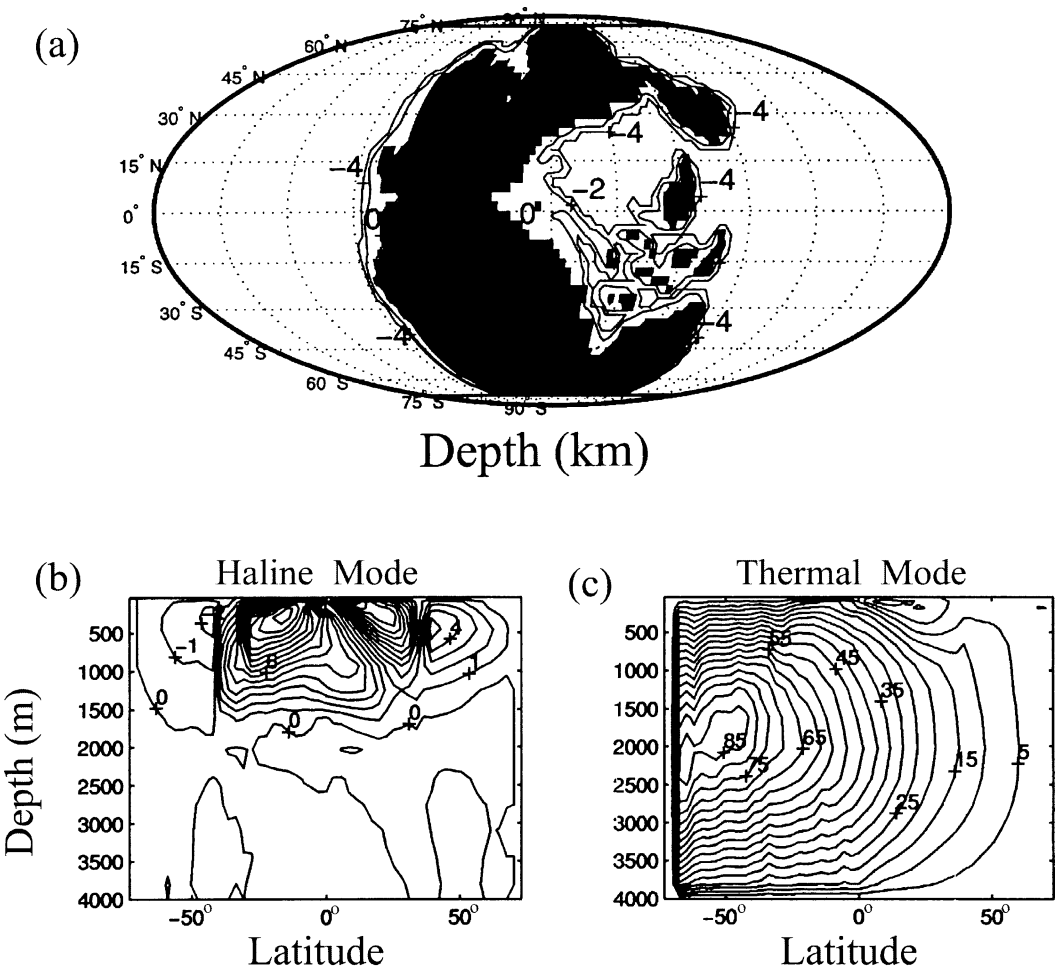

FIG. 1. OGCM studies of the late Permian: (a) late Permian bathymetry (ocean depth in km), (b) overturning streamfunction of HM at year 3975, (c) overturning streamfunction of the transient TM at year 4275. Contour interval is in Sverdrups (Sv; where $1\text{Sv} \equiv 10^6\ \text{m}^3\ \text{s}^{-1}$), positive contour means counterclockwise flow, negative contour means clockwise flow.

inspired by the diagnosis of these three regions in the OGCM.

Figure 3 shows the time series of mean temperature $\hat{T}_i$, and salinity $\hat{S}_i[i = l, h, d; l,$ low-latitude surface; $h,$ high-latitude surface; $d,$ deep ocean; and the caret ($\hat{}$) is a dimensional quantity] in the OGCM over a period of 8000 yr. The period of oscillation is about 3300 yr and two cycles are captured. Note that $\hat{T}_d$ increases during the persistent, quasi-steady HM; $\hat{T}_d$ falls with the onset of deep convection at high southern latitudes. The cycle is asymmetric; the quasi-steady HM lasts much longer than the transient TM period.

Figure 4 shows the time series of the nondimensional (without a caret) vertical density difference between surface regions and deep ocean

$$\Delta\rho_{ld} = \frac{\hat{\rho}_l - \hat{\rho}_d}{\hat{\rho}_0 \alpha \Delta \hat{T}_A}, \qquad \Delta\rho_{hd} = \frac{\hat{\rho}_h - \hat{\rho}_d}{\hat{\rho}_0 \alpha \Delta \hat{T}_A},$$

based on a linear equation of state for density. Here $\hat{\rho}_i(i = l, h, d)$ is the density of each region, $\hat{\rho}_0$ is the mean ocean density, $\Delta\hat{T}_A = \hat{T}_{Al} - \hat{T}_{Ah}$ is the polar–equator surface air temperature difference, and α is the thermal expansion coefficient. During the unsteady HM, the mean trend of $\Delta\rho_{hd}$ increases gradually (Fig. 4b), that is, $d(\Delta\rho_{hd})/dt > 0$, until $\Delta\rho_{hd}$ reaches the threshold—which we call ε—for the onset of polar convection, that is, $\Delta\rho_{hd} = \varepsilon \approx -1.1$. Then, suddenly, strong polar convection begins, $\Delta\rho_{hd}$ jumps to a very high value, and the circulation switches to the transient TM. Polar convection begins even though the zonal mean density structure is still statically stable ($\Delta\rho_{hd} = \varepsilon < 0$), because, as discussed above, convection only occurs in a localized, statically unstable region (Fig. 2). The warm, salty surface eddy (Fig. 2) is cooled quickly, locally destabilizing the water column. As the abyssal temperature increases, the deep ocean density $\hat{\rho}_d$ decreases until it becomes almost the same as the density of the large-scale eddy $\hat{\rho}_{\text{eddy}}$ in the high-latitude surface. Thus when $\Delta\rho_{hd}$ reaches ε, convection occurs, and $\varepsilon \approx (\hat{\rho}_h - \hat{\rho}_{\text{eddy}})/(\hat{\rho}_0 \alpha \Delta\hat{T}_A) < 0$.

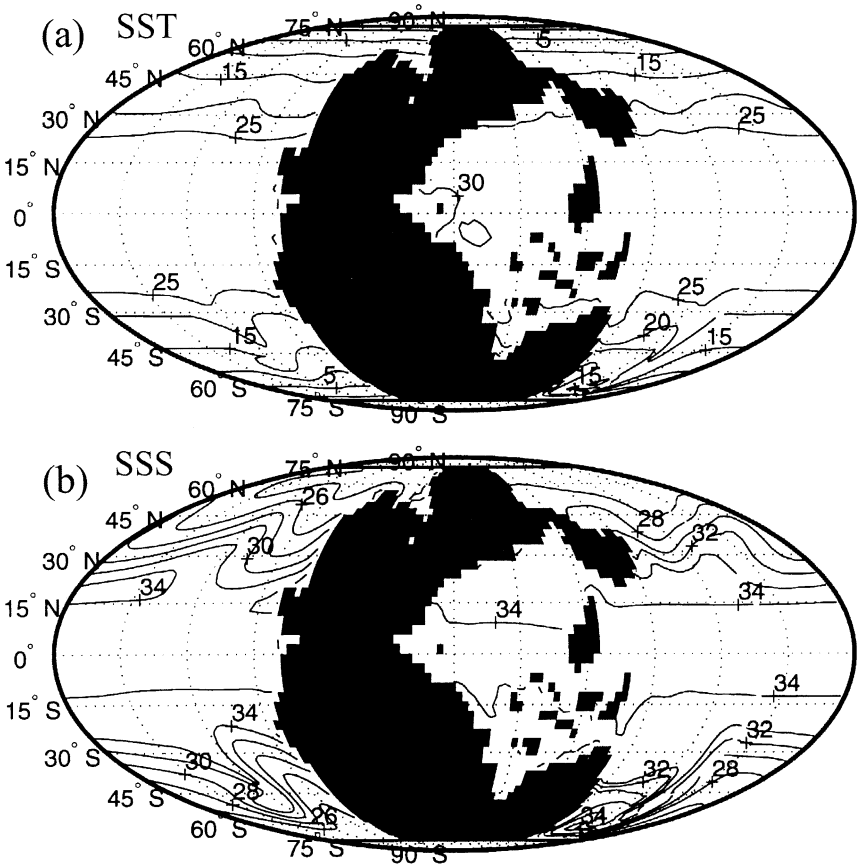

FIG. 2. SST and salinity during the switch from the HM to the TM in our simulation of the late Permian ocean circulation: (a) SST at year 4155 and (b) sea surface salinity (SSS) at year 4155.

During the switch from HM to TM, the mean polar surface temperature and salinity (Figs. 3c,d) increase significantly due to strong mixing and horizontal advection from low to high latitudes. During the transient TM, the mean trend of $\Delta\rho_{hd}$ (Fig. 4b) decreases gradually [$d(\Delta\rho_{hd})/dt < 0$], because the deep ocean density increases due to polar convection. Finally, when the freshwater forcing becomes dominant again, the density of the surface convective region becomes less than that of the deep ocean, that is, $\Delta\rho_{hd}$ reaches the threshold for the termination of polar convection: $\Delta\rho_{hd} = \eta_h \approx -0.3$ (Fig. 4b), convection ceases and $\Delta\rho_{hd}$ drops sharply.

3. Self-sustained thermohaline oscillation in a simple box model

a. Box model description

To better understand the mechanism of the thermohaline switching reviewed in section 2, we developed a simple three-box model inspired by study of the OGCM results. For simplicity the three-box model only represents a single hemisphere and combines the Stommel-type box model (Fig. 5a) with the Welander-type convection model (Fig. 5b). It includes a low-latitude surface box (from latitude 12° to 35°), a high-latitude surface box (from latitude 35° to 70°), and a deep ocean box (Fig. 5c). The horizontal diffusivity between surface boxes is $\hat{K}$, which represents lateral eddy mixing. The vertical diffusivities between the surface and deep ocean in low and high latitudes are, respectively, $\hat{M}_l$ and $\hat{M}_h$, and are functions of the vertical density difference.

Let $\hat{T}_l$, $\hat{T}_h$, $\hat{T}_d$ be the mean temperature of the low-latitude surface box, high-latitude surface box, and deep ocean box, respectively; and let $\hat{S}_l$, $\hat{S}_h$, $\hat{S}_d$ be the mean salinity of the low-latitude surface box, high-latitude surface box, and deep ocean box, respectively. At the air–sea surface, there is a net mean freshwater flux F into the high-latitude box, transported from the low-latitude box. The sea surface temperature $\hat{T}_l$, $\hat{T}_h$ is restored to the air temperature $\hat{T}_{Al}$, $\hat{T}_{Ah}$ at rate constant λ, of a form similar to the boundary conditions used in the OGCM. Let h be the depth of the surface box; H be the depth of the deep ocean box; V_l, V_h, and V_d be the volume of each box, (here, for the given latitude range of our chosen surface boxes, we have $V_l = V_h =$

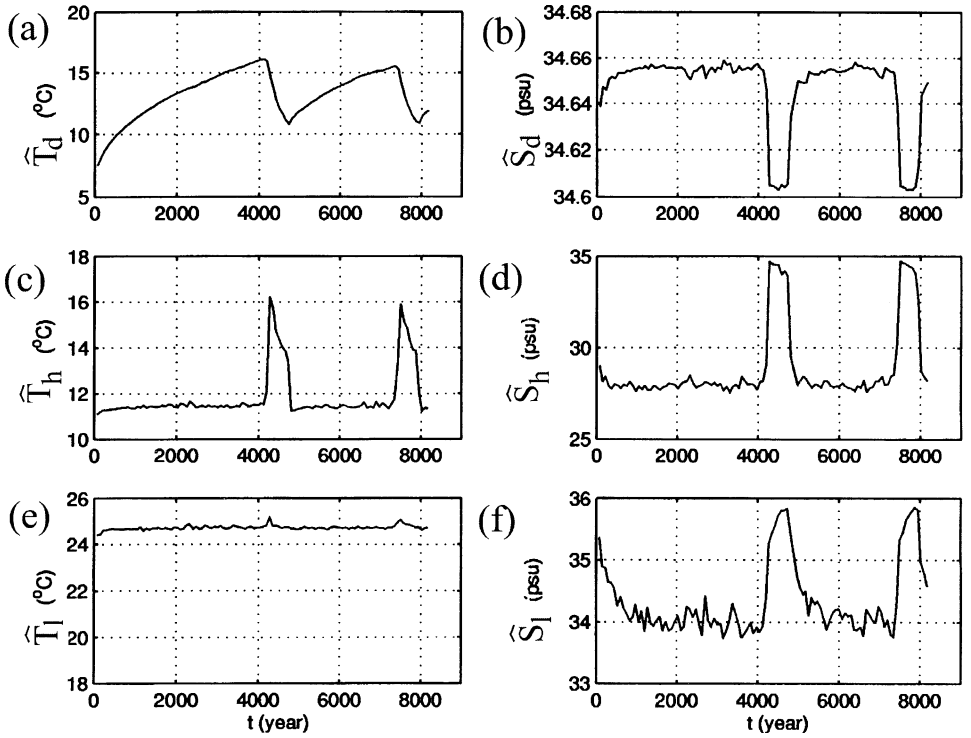

FIG. 3. Time series of mean temperature and salinity in each diagnosed region of the OGCM: (a) $\hat{T}_d$, (b) $\hat{S}_d$, (c) $\hat{T}_h$, (d) $\hat{S}_h$, (e) $\hat{T}_l$, and (f) $\hat{S}_l$. The total salinity of the chosen regions is not conserved. To compare with the three-box model (described in section 3) in which the total salinity is conserved, the mean salinity anomaly of the three regions is subtracted from the salinity of each region.

V); and L be the horizontal distance between the center of the surface boxes (Fig. 5c). Notice that F is the net mean freshwater flux in each box, so it is about half of the peak of the zonal mean latitude-dependent freshwater flux $E - P$. For example, the modern zonal mean $E - P$ profile that has a maximum of about 0.6 m yr^{-1}, corresponds to $F \approx 0.3$ m yr^{-1}.

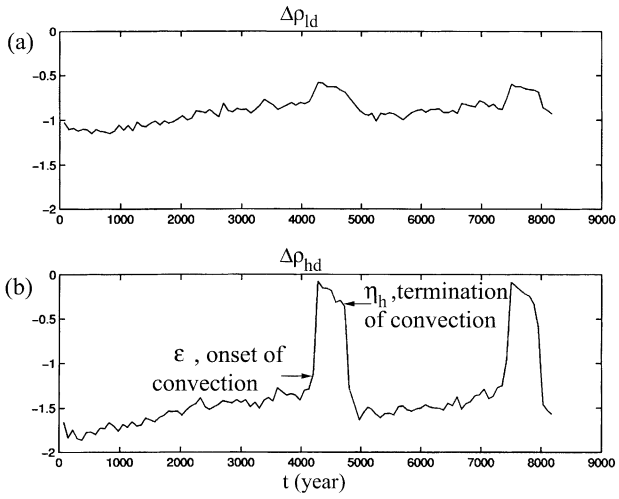

FIG. 4. Time series of mean surface and deep ocean nondimensional density difference of the OGCM: (a) $\Delta\rho_{ld}$ and (b) $\Delta\rho_{hd}$.

The overturning streamfunction q (unit: Sv $\equiv 10^6$ m^3 s^{-1}) is assumed to be linearly proportional to the surface density gradient, as in Stommel's model

$$q = \mu_q(\hat{\rho}_h - \hat{\rho}_l), \tag{1}$$

where $q > 0$ indicates a TM circulation (polar sinking), $q < 0$ indicates an HM circulation (subtropical sinking), and μ_q is the constant of proportionality.

We use a linear equation of state for seawater

$$\hat{\rho}_i = \hat{\rho}_0[1 - \alpha(\hat{T}_i - \hat{T}_r) + \beta(\hat{S}_i - \hat{S}_r)], \tag{2}$$

where $\hat{\rho}_0$, $\hat{T}_r$, $\hat{S}_r$ are the reference density, temperature, and salinity; and α, β are the thermal and saline expansion coefficients, respectively.

The dimensional dynamic equations for $\hat{T}$, $\hat{S}$ in each box are based on an upstream differencing scheme. For example, the high-latitude surface temperature $\hat{T}_h$ evolves due to air–sea heat flux, advection by overturning circulation, and mixing by horizontal diffusion and vertical diffusion.

For the TM ($q > 0$), we have

$$\frac{d\hat{T}_h}{dt} = \lambda(\hat{T}_{Ah} - \hat{T}_h) + \frac{q}{V_h}(\hat{T}_l - \hat{T}_h) + \frac{\hat{K}}{L^2}(\hat{T}_l - \hat{T}_h) + \frac{\hat{M}_h}{\frac{h + H}{2}h}(\hat{T}_d - \hat{T}_h). \tag{3}$$

For the HM ($q < 0$), we have

$$\frac{d\hat{T}_h}{dt} = \lambda(\hat{T}_{Ah} - \hat{T}_h) - \frac{q}{V_h}(\hat{T}_d - \hat{T}_h) + \frac{\hat{K}}{L^2}(\hat{T}_l - \hat{T}_h) + \frac{\hat{M}_h}{\frac{h+H}{2}h}(\hat{T}_d - \hat{T}_h). \quad (4)$$

We can write the above two equations in a form suitable for both the TM ($q > 0$) and HM ($q < 0$), thus,

$$\frac{d\hat{T}_h}{dt} = \lambda(\hat{T}_{Ah} - \hat{T}_h) + \frac{q}{2V_h}(\hat{T}_l - \hat{T}_d) + \frac{|q|}{2V_h}(\hat{T}_l + \hat{T}_d - 2\hat{T}_h) + \frac{\hat{K}}{L^2}(\hat{T}_l - \hat{T}_h) + \frac{\hat{M}_h}{\frac{h+H}{2}h}(\hat{T}_d - \hat{T}_h). \quad (5)$$

The complete dimensional equations for $\hat{T}$, $\hat{S}$ in each box are written out in detail in appendix B.

1) NONDIMENSIONAL EQUATIONS

To nondimensionalize, let $\Delta\hat{T}_A = \hat{T}_{Al} - \hat{T}_{Ah}$ (polar–equator surface air temperature difference) and $\overline{\hat{T}_A} = (\hat{T}_{Al} + \hat{T}_{Ah})/2$ (mean surface air temperature). The controlling nondimensional parameters can then be identified as

$$\gamma = \frac{\Delta\hat{T}_A}{\overline{\hat{T}_A}} \quad \text{(pole–equator air temperature difference)},$$

$$c = \frac{2F}{\lambda h} \quad \text{(freshwater flux)},$$

$$K = \frac{2\hat{K}}{L^2\lambda} \quad \text{(horizontal diffusivity)},$$

$$M_l = \frac{\hat{M}_l}{\frac{h+H}{2}h\lambda} \quad \text{(vertical diffusivity in low latitudes)},$$

$$M_h = \frac{\hat{M}}{\frac{h+H}{2}h\lambda} \quad \text{(vertical diffusivity in high latitudes)},$$

and

$$\mu_f = \frac{\mu_q\hat{\rho}_0\alpha\overline{\hat{T}_A}}{\lambda V}, \qquad \delta = \frac{h}{H}, \qquad R = \frac{\beta\hat{S}_0}{\alpha\Delta\hat{T}_A}.$$

The nondimensional dynamical variables are

$$\Delta T = \frac{\hat{T}_l - \hat{T}_h}{\Delta\hat{T}_A}, \qquad \overline{T} = \frac{\hat{T}_l + \hat{T}_h}{2\overline{\hat{T}_A}}, \qquad T_d = \frac{\hat{T}_d}{\overline{\hat{T}_A}},$$

$$\Delta S = \frac{S_l - \hat{S}_h}{\hat{S}_0}, \qquad \overline{S} = \frac{\hat{S}_l + \hat{S}_h}{2\hat{S}_0}, \qquad S_d = \frac{\hat{S}_d}{\hat{S}_0}.$$

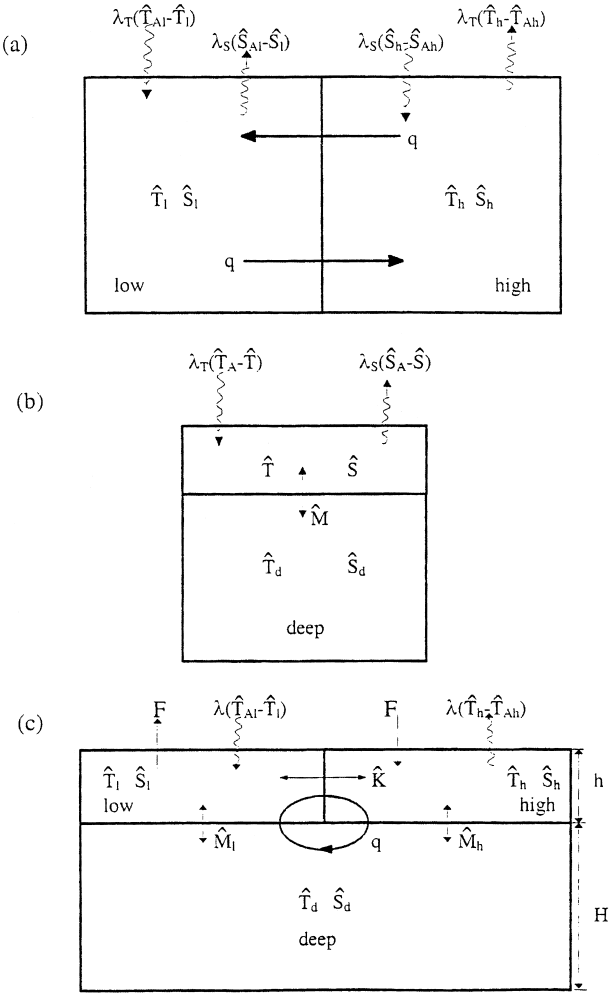

FIG. 5. (a) Schematic diagram of the Stommel-type box model. (b) Schematic diagram of Welander-type convection model. (c) Schematic diagram of the three-box model. The horizontal diffusivity between surface boxes is K; M_l, M_h are the vertical diffusivities between surface and deep ocean in low and high latitude, respectively. Net mean freshwater flux F entering the high-latitude surface box is balanced by outflux from the low-latitude surface box. The SST $\hat{T}_l$, $\hat{T}_h$ is restored to the air temperature $\hat{T}_{Al}$, $\hat{T}_{Ah}$ with restoration rate λ. The overturning strength is q. The surface box has depth h; the deep box has depth H.

The nondimensional overturning streamfunction is $f = q/(\gamma\lambda V)$, and Eq. (1) becomes

$$f = \mu_f(\Delta T - R\Delta S). \quad (6)$$

Based on the dimensional equations written out in full in appendix B, we write the nondimensional equations in terms of the nondimensional variables ΔT, $\overline{T}$, T_d, ΔS, $\overline{S}$, S_d for both TM ($f > 0$) and HM ($f < 0$) circulation. They are (where $t' = \lambda t$ is nondimensional time)

$$\frac{d\Delta T}{dt'} = 1 - \Delta T + f(T_d - \overline{T}) - \frac{3}{2}|f|\gamma\Delta T - K\Delta T + M_l\left[\frac{1}{\gamma}(T_d - \overline{T}) - \frac{\Delta T}{2}\right] - M_h\left[\frac{1}{\gamma}(T_d - \overline{T}) + \frac{\Delta T}{2}\right], \quad (7)$$

$$\frac{d\overline{T}}{dt'} = 1 - \overline{T} + \frac{\gamma}{2}|f|(T_d - \overline{T}) + \frac{\gamma^2}{4}f\Delta T + \frac{M_l}{2}\left(T_d - \overline{T} - \frac{\gamma\Delta T}{2}\right) + \frac{M_h}{2}\left(T_d - \overline{T} + \frac{\gamma\Delta T}{2}\right), \quad (8)$$

$$\frac{dT_d}{dt'} = -\delta\left[\frac{\gamma}{2}|f|(T_d - \overline{T}) + \frac{\gamma^2}{4}f\Delta T + \frac{M_l}{2}\left(T_d - \overline{T} - \frac{\gamma\Delta T}{2}\right) + \frac{M_h}{2}\left(T_d - \overline{T} + \frac{\gamma\Delta T}{2}\right)\right], \quad (9)$$

$$\frac{d\Delta S}{dt'} = c + f\gamma(S_d - \overline{S}) - \frac{3}{2}|f|\gamma\Delta S - K\Delta S + M_l\left(S_d - \overline{S} - \frac{\Delta S}{2}\right) - M_h\left(S_d - \overline{S} + \frac{\Delta S}{2}\right), \quad (10)$$

$$\frac{d\overline{S}}{dt'} = \frac{\gamma}{2}|f|(S_d - \overline{S}) + \frac{\gamma}{4}f\Delta S + \frac{M_l}{2}\left(S_d - \overline{S} - \frac{\Delta S}{2}\right) + \frac{M_h}{2}\left(S_d - \overline{S} + \frac{\Delta S}{2}\right), \quad (11)$$

$$\frac{dS_d}{dt'} = -\delta\frac{d\overline{S}}{dt'}. \quad (12)$$

We will examine both steady and time-dependent solutions of these equations, respectively.

2) Representation of convection

The nondimensional vertical diffusivities M_l, M_h in the above equations depend on the ocean state due to convective process. We make the values of M_l and M_h functions of the nondimensional mean vertical density difference between the surface box and deep ocean box $\Delta\rho_i (i = ld, hd)$ as discussed in section 2. If no convection occurs at low or high latitudes, then $M_l = M$ or $M_h = M$, where $M = 2\hat{M}/[(h + H)h\lambda]$ is the nondimensional background vertical diffusivity ($\hat{M}$ is the dimensional background vertical diffusivity). When the density difference is conducive to convection at low (high) latitude, $M_l(M_h)$ becomes much larger than M.

Based on the diagnosis of the OGCM in section 2, we represent the effect of local polar convection in our three-box model by introducing a threshold, ε, for the onset of polar convection. Here the physical meaning of ε is the same as found in the OGCM (see section 2 and Fig. 4b): during the unsteady HM [$d(\Delta\rho_{hd})/dt > 0$], when the mean nondimensional density difference between the polar surface and deep ocean reaches this threshold, that is, $\Delta\rho_{hd} = \varepsilon$, polar convection starts. If it does not reach this threshold, that is, $\Delta\rho_{hd} < \varepsilon$, polar convection can never occur. Here $\varepsilon < 0$ represents the effect of localized polar convection discussed in section 2. Similarly, when the system is in its polar convection phase, we introduce a threshold, η_h, for the termination of polar convection (see section 2 and Fig. 4b): when $\Delta\rho_{hd} < \eta_h$ during the transient TM [$d(\Delta\rho_{hd})/dt < 0$], polar convection terminates. During the polar convective phase, the high-latitude surface density is spatially more homogeneous than that in the nonconvective phase, so $\eta_h > \varepsilon$ and η_h is set to be close to zero in the box model.

To summarize, we can write the following simplified rules for convective adjustment that include the effect of local convection in the three-box model.

At high latitudes,

if $\Delta\rho_{hd} < \varepsilon$, then $M_h = M$ (no convection);
if $\Delta\rho_{hd} \geq \varepsilon$, and $d(\Delta\rho_{hd})/dt > 0$, then $M_h = M_{sc}$ (strong convection);
if $\Delta\rho_{hd} \geq \varepsilon$, and $d(\Delta\rho_{hd})/dt \leq 0$, then
  if $\Delta\rho_{hd} < \eta_h$, then $M_h = M$ (no convection);
  if $\Delta\rho_{hd} \geq \eta_h$, then $M_h = M_{sc}$ (strong convection).

Here $M_{sc} \gg M$, indicating that vertical mixing of cold fresh surface water in high latitudes is strong and deep.

At low latitudes,

if $\Delta\rho_{ld} < \eta_l$, then $M_l = M$ (no convection);
if $\Delta\rho_{ld} \geq \eta_l$, then $M_l = M_{wc}$ (weak convection).

Here $M_{sc} > M_{wc} \gg M$, indicating that vertical mixing of warm salty surface water in low latitudes is weak and shallow compared to strong, deep convection in high latitudes. Because the density distribution in low latitudes is horizontally more homogeneous than in high latitudes, we use the same threshold, η_l, close to zero, for the onset and termination of convection in low latitudes. Note that η_l and η_h are chosen to be slightly nonzero to avoid false numerical oscillations caused by the discontinuity in M_l (as discussed in Welander 1982).

We note that the above convective adjustment at high latitudes also depends upon the time rate of change of the stratification, that is, $d(\Delta\rho_{hd})/dt$. This is because we

wish the box model to represent essential physics; the zonal mean stratification thresholds for the onset (ε) and the termination (η_h) of polar convection are *different.* During the quasi-steady HM, the polar surface density is highly inhomogeneous and the onset of the polar convection is triggered by localized convection even though the zonal mean stratification is still quite stable; during the transient TM, the zonal mean stratification is significantly smaller and the polar surface density is more homogeneous (Fig. 4b). Thus we make $\eta_h > \varepsilon$. The box model has to judge whether it is facing the onset or the termination of polar convection and uses a mean stratification threshold dependent on the current state of the system as reflected in the sign of $d(\Delta\rho_{hd})/dt$. When the system evolves from the quasi-steady HM toward the transient TM, $d(\Delta\rho_{hd})/dt > 0$ just before and during the onset of polar convection; when the system evolves from the transient TM toward the quasi-steady HM, $d(\Delta\rho_{hd})/dt < 0$ just before and during the termination of polar convection. Thus given the sign of $d(\Delta\rho_{hd})/dt$, the box model can judge which stratification threshold—whether it be ε or η_h—to use.

b. Steady solutions of the three-box model

Let us first look at the steady solutions of Eqs. (7)–(12). They can be found by setting the rhs to zero, expressing ΔT, $\overline{T}$, T_d, ΔS, $\overline{S}$, S_d in terms of f, and substituting the expressions for ΔT, ΔS into Eq. (6) to yield a fifth-order equation in f:

$$f = \frac{2\mu_f(M_l + M_h + |f|\gamma)}{(M_l + M_h + 2|f|\gamma)^2 - (M_h - M_l)^2 + 2(K+1)(M_l + M_h + |f|\gamma)} - \frac{2\mu_f(M_l + M_h + |f|\gamma)Rc}{(M_l + M_h + 2|f|\gamma)^2 - (M_h - M_l)^2 + 2K(M_l + M_h + |f|\gamma)}. \tag{13}$$

Only the real roots of this equation are physically possible solutions. Substituting the appropriate vertical diffusivities M_l, M_h for the corresponding convective/nonconvective states (discussed in appendix C) into Eq. (13), we can obtain the steady TM ($f > 0$) and HM ($f < 0$) for given parameters.

1) Regions on the bifurcation diagram

Figure 6 shows the bifurcation diagram on the f–c (nondimensional overturning-freshwater flux) plane. The steady TM and HM overturning streamfunction f, as a function of the freshwater forcing c, are obtained by solving the real roots of Eq. (13) with other parameters fixed. The constants of the three-box model are summarized in Table 1. The fixed parameters for the bifurcation diagram are listed in Table 2 (here μ_f is chosen in accord with modern ocean overturning strength and surface density gradient). As in Stommel's box model (1961), there are three branches of steady solution on the f–c plane (Fig. 6). The upper branch (thin solid line) is the stronger stable steady TM ($f > 0$); the middle branch (dashed line) is the weaker unstable steady TM ($f > 0$), both obtained with $M_l = M$, $M_h = M_{sc}$ (polar convection); the lower branch is the steady HM ($f < 0$) obtained with $M_l = M_{wc}$, $M_h = M$ (subtropical convection).

In the case of convective adjustment, when $\varepsilon = 0$, the steady HM will always be stable as in Stommel's box model (1961); but if $\varepsilon < 0$ (allowing localized polar convection), the steady HM can be unstable and there are four different regions on the bifurcation diagram (Fig. 6), separated by three critical values of freshwater flux ($c_{r1} < c_{r2} < c_{r3}$).

Region I ($c \le c_{r1}$): Only the globally stable steady TM exists.

Region II ($c_{r1} < c < c_{r2}$): Both the locally stable steady TM and the locally stable limit cycle exists, depending on initial conditions.

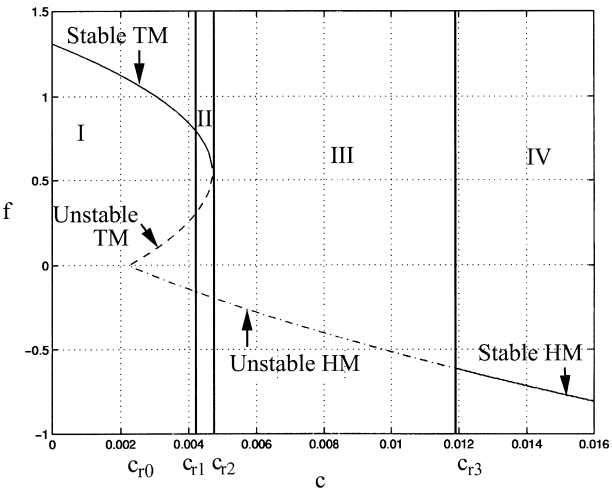

FIG. 6. Bifurcation diagram on the f–c plane of the three-box model: stable TM (thin solid line), unstable TM (dashed line), unstable HM (dot–dashed line), stable HM (thin solid line). Region I ($c \le c_{r1}$): a globally stable steady TM exists. Region II ($c_{r1} < c < c_{r2}$): a local stable steady TM and a locally stable limit cycle exist, depending on initial conditions. Region III ($c_{r2} \le c \le c_{r3}$): a globally stable limit cycle exists. Region IV ($c_{r3} < c$): a globally stable HM exists.

TABLE 1. The three-box model constants.

Constant	Value
α (K^{-1})	2×10^{-4}
β (psu^{-1})	7×10^{-4}
L (km)	3190
V (m^3)	3.265×10^{15}
λ (day^{-1})	1/90
$\hat{S}_0$ (psu^{-1})	35
$\hat{K}$ (m^2 s)	1×10^4
M_{sc}	0.2
M_{wc}	0.1
h (m)	50
H (m)	4000
δ	1/80

Table 2. Parameters for the bifurcation diagram on f–c plane.

Parameter	Value
$\Delta\hat{T}_A$(K)	14
$\overline{\hat{T}_A}$(K)	291
$\gamma = \dfrac{\Delta\hat{T}_A}{\overline{\hat{T}_A}}$	0.0481
$R = \dfrac{\beta\hat{S}_0}{\alpha\Delta\hat{T}_A}$	8.75
$\mu_f = \dfrac{\mu_g\hat{\rho}_0\alpha\overline{\hat{T}_A}}{\lambda V}$	1.5
$M = \hat{M}\Big/\left(\dfrac{h+H}{2}h\lambda\right)$	0.0025

Region III ($c_{r2} \le c \le c_{r3}$): Only the globally stable limit cycle exists.

Region IV ($c_{r3} < c$): Only the globally stable steady HM exists.

The stability of the steady solutions were studied using linear analysis. The stability of the limit cycle was addressed numerically. We now discuss the regions in details.

On the f–c plane in Fig. 6, c_{r0}, c_{r2} are the two bifurcation points. When $c_{r0} < c < c_{r2}$, all three steady solutions (two TM, one HM) exist; when $c < c_{r0}$ only a steady TM exists; when $c > c_{r2}$ only a steady HM exists. Here $c_{r2} \approx 0.0047$ for the given parameters in Table 2 and conversion from c to dimensional freshwater fluxes F is outlined in Table 3.

For the steady HM ($f < 0$), the nondimensional mean density difference between high-latitude surface and deep ocean boxes, $\Delta\rho_{hd}$, can be obtained from Eqs. (7)–(12) and (6):

$$\Delta\rho_{hd} = \frac{f}{\mu_f}\frac{M_l - \gamma f}{M_l + M_h - \gamma f}. \tag{14}$$

There is a critical value of freshwater forcing c_{r3} determined by setting $\Delta\rho_{hd} = \varepsilon$ (threshold for the onset of polar convection) and $M_l = M_{wc}$, $M_h = M$ (subtropical convection) in Eq. (14):

$$\Delta\rho_{hd} = \frac{f(c_{r3})}{\mu_f}\frac{M_{wc} - \gamma f(c_{r3})}{M_{wc} + M - \gamma f(c_{r3})} = \varepsilon. \tag{15}$$

When $c > c_{r3}$, we have $\Delta\rho_{hd} < \varepsilon$ and the steady HM is stable (thin solid line). When $c \le c_{r3}$, we have $\Delta\rho_{hd} \ge \varepsilon$, and the steady HM is unstable (dot–dashed line) in the sense that small perturbations lead to $d(\Delta\rho_{hd})/dt > 0$ triggering polar convection ($M_h = M_{sc}$), making the system suddenly jump away from the steady HM to a transient TM. It never returns to the steady HM.

In region III ($c_{r2} \le c \le c_{r3}$), no stable steady solution exists and the system must oscillate. Because the steady HM is the only stable solution without convective adjustment, the system will evolve toward it, that is, $d(\Delta\rho_{hd})/dt > 0$ in the absence of polar convection if $\Delta\rho_{hd} < \varepsilon$ initially. But when the condition $\Delta\rho_{hd} = \varepsilon$ is satisfied, polar convection begins. The system jumps away before reaching the steady HM, strong convection increases the polar surface density significantly and induces the circulation to switch to the transient nonsteady TM [$d(\Delta\rho_{hd})/dt > 0$, $f > 0$]. Since a steady TM does not exist, freshwater forcing gradually becomes dominant again and $\Delta\rho_{hd}$ continues to decrease until it becomes less than η_h. Polar convection terminates, the system switches back to the quasi-steady HM ($f < 0$) and once again evolves toward the steady HM, completing the limit cycle. This is a globally stable limit cycle; small perturbations will not destroy it and it can be reached from any initial condition.

In region IV ($c_{r3} < c$), the steady HM always satisfies $\Delta\rho_{hd} < \varepsilon$, and so it is globally stable and no limit cycles exist.

In region I ($c \le c_{r1}$), when $c < c_{r0}$, only one globally stable steady TM exists; when $c_{r0} \le c \le c_{r1}$, the basin of attraction of the stable steady TM is sufficiently large that when the unstable HM switches to the transient TM due to polar convection, it will be ultimately attracted by the globally stable steady TM and stay there forever. There are thus no limit cycles; only one globally stable steady TM exists, no matter what the initial conditions.

In region II ($c_{r1} < c < c_{r2}$), the basin of attraction of the stable steady TM is small enough that the whole limit cycle is outside this basin of attraction. Oscillations can still exist depending on the initial conditions: if the initial state in phase space is close to the stable steady TM, then it will evolve until it reaches the locally stable

Table 3. Conversion between dimensional and nondimensional variables of the bifurcation diagrams using constants in Tables 1 and 2.

Nondimensional variable	Dimensional variable
c (freshwater forcing)	F (m yr^{-1}) $= \dfrac{\lambda h}{2}c \approx 100c$
M (background vertical diffusivity)	$\hat{M}$ (m^2 s^{-1}) $= \dfrac{h+H}{2}h\lambda M \approx 0.013M$
f (overturning stream function)	$\hat{q}$ (Sv) $= \gamma\lambda V f \approx 21f$

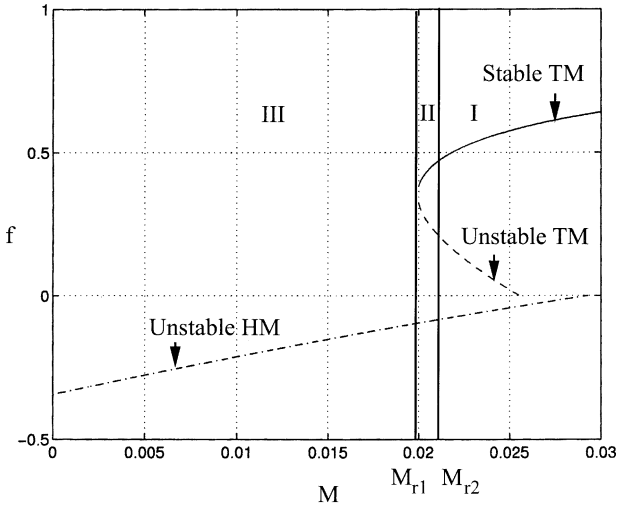

FIG. 7. Bifurcation diagram on the f–M plane of the three-box model: Stable TM (thin solid line), unstable TM (dashed line), unstable HM (dot–dashed line). Region I ($M_{r2} \leq M$): a globally stable steady TM exists. Region II ($M_{r1} < M < M_{r2}$): a locally stable steady TM and a locally stable limit cycle exist, depending on initial conditions. Region III ($M \leq M_{r1}$): a globally stable limit cycle exists.

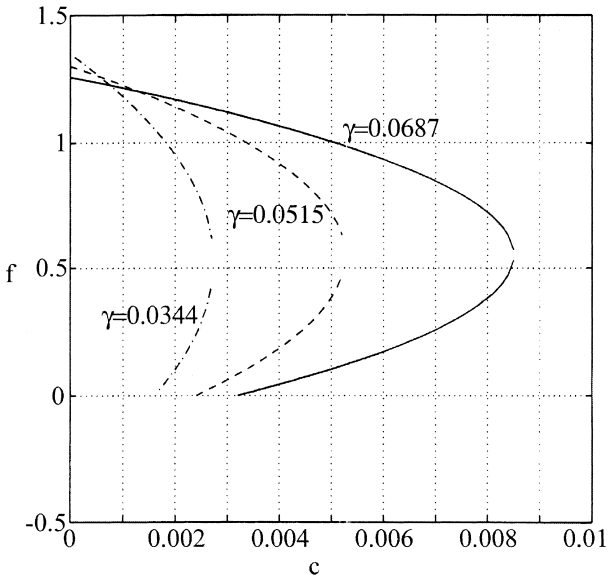

FIG. 8. Bifurcation diagram on the f–c plane for different values of the Pole–equator temperature gradient, $\gamma(\gamma = \Delta\hat{T}_A/\overline{\hat{T}_A}$, $\overline{\hat{T}_A}$ = 291 K), other parameters are as in Fig. 6.

steady TM; if the initial state is close to the steady HM, then it will evolve until it reaches the locally stable limit cycle and keep oscillating. Since the system's phase space is five-dimensional it is very difficult to find the value of c_{r1} analytically—we can only obtain its value by numerical methods (discussed in section 3c). We find that $c_{r1} \approx 0.0043$ for the given parameters (Table 2).

Thermohaline oscillations are only possible in the window of the freshwater forcing in regions II and III. This is consistent with the 2D OGCM results found by Winton and Sarachik (1993). They found that such oscillations exist for certain ranges of freshwater forcing, but when the freshwater forcing is very large, only a stable steady HM exists.

Similarly we can plot the bifurcation diagram on the f–M (nondimensional overturning-vertical diffusivity) plane (Fig. 7), with other parameters (except M) fixed as in Table 2, with $c = 0.0065$ ($F \approx 0.65$ m yr^{-1}, see Table 3). We see two critical values $M_{r1} \approx 0.0199$ (bifurcation point), $M_{r2} \approx 0.0209$, and three regions of physically possible solutions on the f–M plane.

Region I ($M_{r2} \leq M$): Only the globally stable steady TM exists.

Region II ($M_{r1} < M < M_{r2}$): Both the locally stable steady TM and the locally stable limit cycle exist, depending on initial conditions.

Region III ($M \leq M_{r1}$): Only the globally stable limit cycle exists.

Thermohaline oscillations are possible only in the window of the vertical diffusivity in regions II and III. Here we do not find a fourth region in which only the stable steady HM exists. However, for much larger values of c, a region IV on the f–M plane is also possible.

As both bifurcation diagrams are shown with nondimensional variables, f, c, and M, we summarize the numerical conversion between dimensional and nondimensional variables of the bifurcation diagrams in Table 3 to estimate the physical quantities.

The region of the bifurcation diagram corresponding to the present climate is confined by $c < c_{r2}$; we observe a stable steady TM circulation and it is difficult to envisage moving from our present rather cold climate into region III. However, c_{r2}, the critical freshwater forcing at the bifurcation point beyond which there is no steady TM solutions, is also sensitive to the parameter γ (polar–equator surface air temperature difference) and decreases with it (Fig. 8 and Table 4, where $\overline{\hat{T}_A}$ is fixed as 291 K. For warm equable climates, γ is smaller due to smaller $\Delta\hat{T}_A$. Thus c_{r2} will be smaller and it should be easier to reach the oscillatory solutions in region III during warm equable climates.

TABLE 4. Variations of critical freshwater forcing at the bifurcation point with polar–equator surface air temperature difference.

$\Delta\hat{T}_A$ (air temperature difference, K)	20	15	10
γ (nondimensional)	0.0687	0.0515	0.0344
F_{r2} (critical freshwater forcing, m yr^{-1})	0.85	0.53	0.27
c_{r2} (nondimensional)	0.0085	0.0053	0.0027

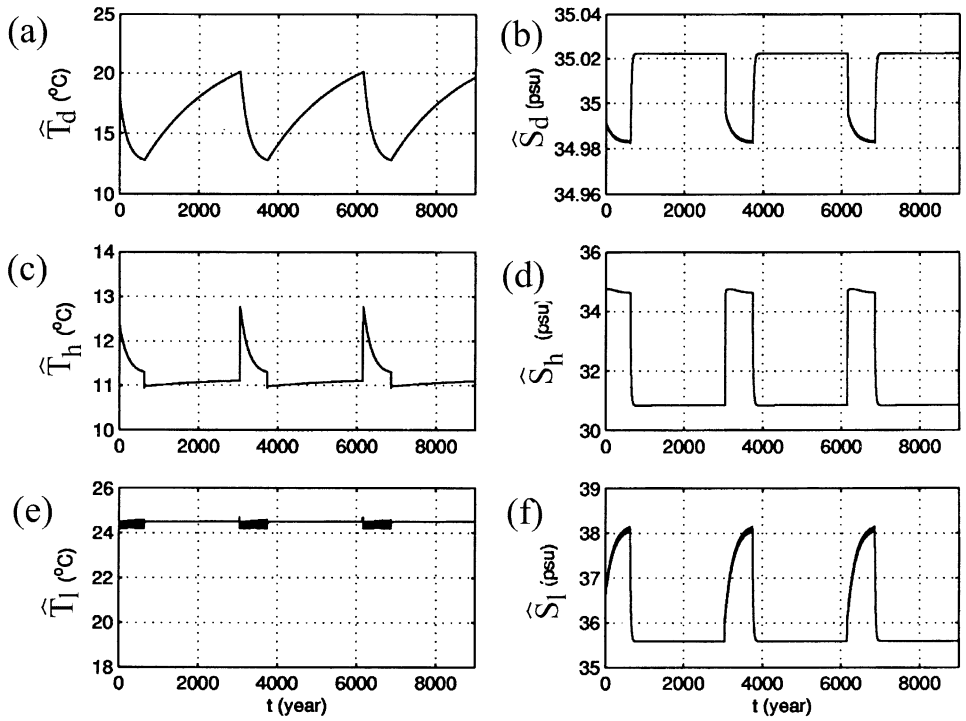

FIG. 9. Time series of temperature and salinity in each box of the three-box model: (a) $\hat{T}_d$, (b) $\hat{S}_d$, (c) $\hat{T}_h$, (d) $\hat{S}_h$, (e) $\hat{T}_l$, and (f) $\hat{S}_l$.

2) WHAT SETS THE CONVECTION SWITCH, ε, AND THE FRESHWATER BOUNDARY SEPARATING STABLE/UNSTABLE STEADY HM?

In our OGCM, we estimate that the threshold for polar convection is $\varepsilon \approx -1.1$ (Fig. 4b). Given the bathymetry we used for the late Permian ocean, salt is transported from the surface of the narrow Tethys Sea into the deep ocean, because the Tethys Sea is isolated from the open ocean and surface evaporation there is strong. This makes the mean deep ocean salinity much higher than normal. To conserve the total salinity, the mean surface polar salinity $\hat{S}_{h\_\mathrm{gcm}}$ has to be very low, around 28 psu during the quasi-steady HM. In our three-box model the typical value for $\hat{S}_{h\_\mathrm{box}}$ during the quasi-steady HM with similar parameters as the OGCM (i.e., $F = 0.65$ m yr^{-1}, $\hat{M} = 3.33 \times 10^{-5}$ m^2 s^{-1}) is higher, about 30.8 psu. As discussed in section 2, $\varepsilon \approx (\hat{\rho}_h - \hat{\rho}_{\mathrm{eddy}})/(\hat{\rho}_0 \alpha \Delta \hat{T}_A)$, and assuming $\hat{\rho}_{\mathrm{eddy}}$ and $\hat{T}_h$ in the quasi-steady HM for both the OGCM and the three-box model are similar, we estimate

$$\frac{(\hat{\rho}_{h\_\mathrm{gcm}} - \hat{\rho}_{h\_\mathrm{box}})}{(\hat{\rho}_0 \alpha \Delta \hat{T}_A)} \approx \frac{\beta(\hat{S}_{h\_\mathrm{gcm}} - \hat{S}_{h\_\mathrm{box}})}{(\alpha \Delta \hat{T}_A)} \approx -0.7. \quad (16)$$

This gives us a rough guide to the difference between ε in the OGCM and ε in the box model due to the salinity difference and the effect of the Tethys sea. Choosing $\varepsilon \approx -0.4$ in our three-box model (in reality ε may depend on c, M_l, M_h, etc.) and substituting it into the relation $\Delta\rho_{hd} = \varepsilon$ [Eq. (15)], we obtain the critical value of the freshwater forcing separating stable and unstable steady HM. It is $c_{r3} \approx 0.0119$, which corresponds to a mean freshwater forcing $F \approx 1.19$ m yr^{-1}.

c. Time-dependent solutions: Explicit oscillations

We now discuss the time-dependent solutions of the three-box model obtained numerically. Equations (7)–(12) are integrated forward using a Runge–Kutta method and convective adjustment employed at each time step. By choosing $c = 0.0065$ ($F \approx 0.65$ m yr^{-1}), $M = 0.0025$ ($\hat{M} \approx 3.3 \times 10^{-5}$ m^2 s^{-1}), similar to that used in the OGCM with other parameters fixed at values in Table 2, the system is within the window on the f–c plane in which oscillations are possible. Indeed we obtain oscillatory solutions with a period of about 3000 yr. Figures 9–10 shows the time series of temperature $\hat{T}_i$ and salinity $\hat{S}_i (i = l, h, d)$ in each box and the nondimensional vertical density difference between the surface box and deep ocean box $\Delta\rho_i (i = ld, hd)$, respectively. Comparing the three-box model results (Figs. 9, 10) with the OGCM (Figs. 3, 4), we see that they are very similar. Figure 11 shows the time series of the nondimensional overturning circulation f of the three-box model. The system resides for a long period in the quasi-steady HM ($q \approx -7.4$ Sv), until it suddenly jumps to the transient TM ($q \approx 15.5$ Sv), then quickly returns to the HM. In the three-box model, the strength of the

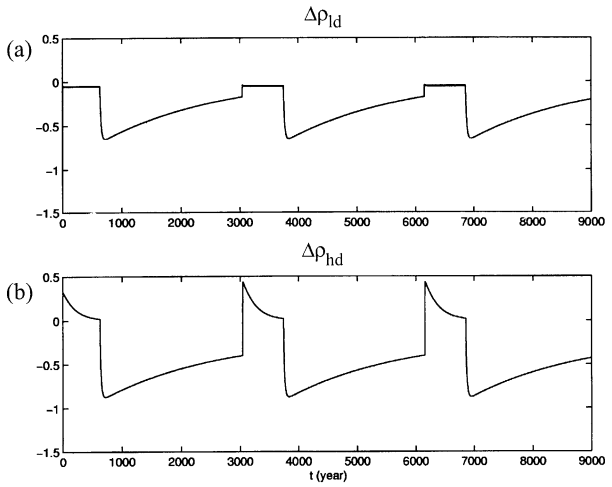

FIG. 10. Time series of surface and deep ocean nondimensional density difference of the three-box model: (a) $\Delta\rho_{ld}$ and (b) $\Delta\rho_{hd}$.

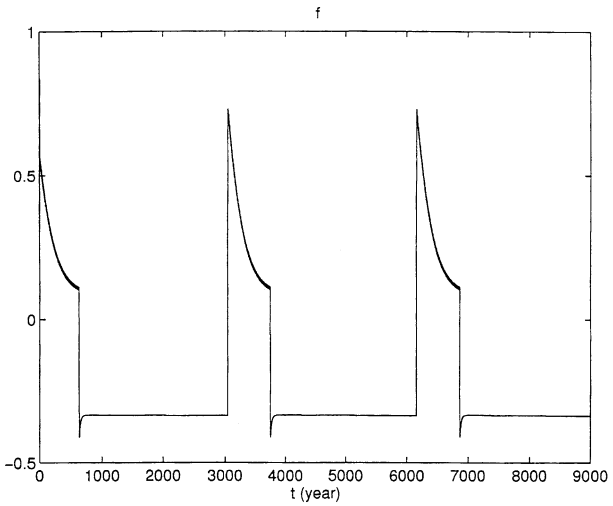

FIG. 11. Time series of the nondimensional overturning circulation of the three-box model.

transient TM is much smaller than that in the OGCM. This may be, for example, because the simple assumption [Eq. (1)] used in the box model does not capture the details of the dynamics of the overturning strength in the ocean. Figure 12 is the projection of the phase portrait of the limit cycle onto the $\hat{T}_d$–$\hat{S}_d$ plane for the OGCM and the three-box model. Both exhibit a loop with two fast branches (the onset and the termination of polar convection) and two slow branches. In the three-box model, there are small noisy oscillations during the transient TM before the termination of polar convection (Figs. 9–11). This is due to convection in low latitudes: when $\Delta\rho_{ld} \geq \eta_l$, convection occurs in low latitudes, and convective mixing $M_l = M_{\rm wc}$ gradually decreases $\Delta\rho_{ld}$ until $\Delta\rho_{ld} < \eta_l$, then convection ceases and we have $M_l = M$. The surface freshwater forcing ensures that $\Delta\rho_{ld} \geq \eta_l$ once more and convection resumes with $M_l = M_{\rm wc}$. This kind of low-amplitude, high-frequency oscillation is similar to the heat–salt oscillator of Welander's convection model (Welander 1982). It does not affect the existence of the low-frequency lower-amplitude oscillation induced by polar convection. Here we have set $\eta_l = -0.05$, $\eta_h = 0.02$: they are not exactly zero so as to avoid spurious numerical oscillations that occur when M_l is discontinuous during the transient TM. In our OGCM (Fig. 4a), we always have $\Delta\rho_{ld} < \eta_l$ due to higher levels of deep ocean salinity, and we do not observe this kind of high-frequency oscillation.

By decreasing c from $c_{r2} = 0.0047$ and experimenting with different initial conditions, we found that when $c \leq 0.0043$, the solution always ends up in the stable steady TM, no matter what the initial conditions. When $c > 0.0043$, both the stable steady TM and oscillatory solution are possible depending on initial conditions, suggesting that $c_{r1} \approx 0.0043$. Similarly we deduce that $M_{r2} \approx 0.0209$ (Fig. 7) by experimenting with different initial conditions. Again, to convert to dimensional parameters, see Table 3.

Dependence of oscillation period on freshwater flux/vertical diffusivities

If we vary c in region II, III on the f–c plane and choose appropriate initial conditions when in region II, we can study the relation of the oscillation period to the amplitude of freshwater forcing. Figure 13 plots the total oscillation period ($\tau_{\rm os}$, line with circles), the duration of the quasi-steady HM ($\tau_{\rm HM}$, line with dots), and the duration of the transient TM ($\tau_{\rm TM}$, line with stars) from one cycle of the oscillation, as a function of c in the oscillatory window of the f–c plane (all other parameters are the same as in Fig. 6). The total oscillation period $\tau_{\rm os}$ is dominated by the duration of the quasi-steady HM $\tau_{\rm HM}$. When freshwater forcing c is small, $\tau_{\rm HM}$ decreases as c increases; when c is large, $\tau_{\rm HM}$ increases as c increases.

If we vary M (background mixing rate) in region II, III on the f–M plane and choose appropriate initial conditions when in region II, we can also illustrate the relation of oscillation period to background vertical diffusivity. Figure 14 shows $\tau_{\rm os}$ (line with circles), $\tau_{\rm HM}$ (line with dots), and $\tau_{\rm TM}$ (line with star) as a function of M in the oscillatory window of the f–M plane, other parameters are the same as in Fig. 7. The duration of the HM, $\tau_{\rm HM}$, always decreases as background mixing rate M increases because the larger M becomes, the faster the deep ocean temperature increases due to vertical diffusion, and the earlier polar convection starts.

4. Summary and Discussion

a. Haline–thermal mode switching mechanism

The underlying mechanism of haline–thermal mode switching was studied in an OGCM of the late Permian ocean circulation and a box model was constructed to study the stability properties of the steady states and

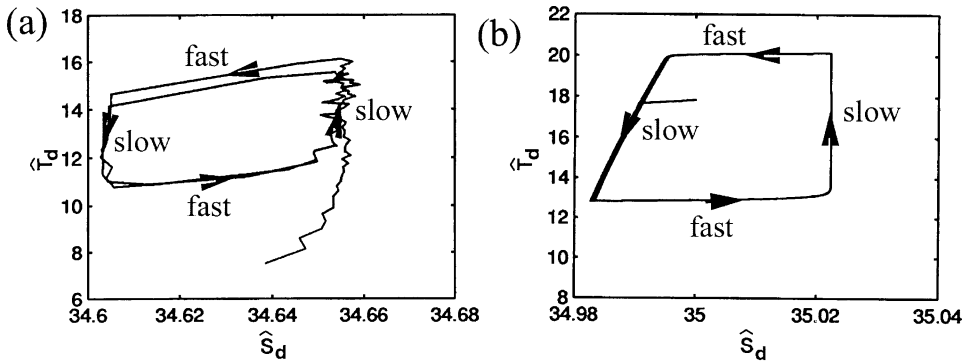

FIG. 12. Projection of the phase portrait of the limit cycle on $\hat{T}_d$–$\hat{S}_d$ plane: (a) OGCM and (b) three-box model.

transitions between them. Within certain parameter regimes of forcing, mixing, and Pole–equator temperature gradient—which in the context of the box model can be precisely determined (see below)—a quasi-steady HM evolves toward a steady HM. Polar stratification decreases due to the increase in abyssal ocean temperatures of the HM. Eventually an unstable stratification occurs associated with the formation of warm, salty, large-scale eddies in the polar surface region that act as ''preconditioning'' centers for polar deep convective activity. Thus before reaching the steady HM attractor, the system jumps away to a transient, nonsteady TM; strong convection increases the polar surface density significantly and induces a mode switch. Since a steady TM does not exist, freshwater forcing gradually becomes dominant again and the density stratification continues to increase until polar convection terminates and the cycle repeats itself.

The idealized box model was used to elucidate this underlying mechanism which, in certain parameter regimes, is a property of ocean circulation models and perhaps the real climate system. The box model illustrates the inherent instability of the HM and the importance of polar convection in the flushing mechanism. A key property of our box model is the manner in which deep convection is parameterized. Motivated by our OGCM results, we employ a stratification threshold ε, so that polar convection can occur even though the large-scale, mean density structure is statically stable—see Marshall and Schott (1999) for a discussion of the observed patchiness of deep convective process in the current ocean.

b. Stability analysis of the box model

Using steady solutions, and time-dependent integrations of the box model, we are able to fully explore the circulation and stability properties of the system over a wide range of parameter space. On the bifurcation di-

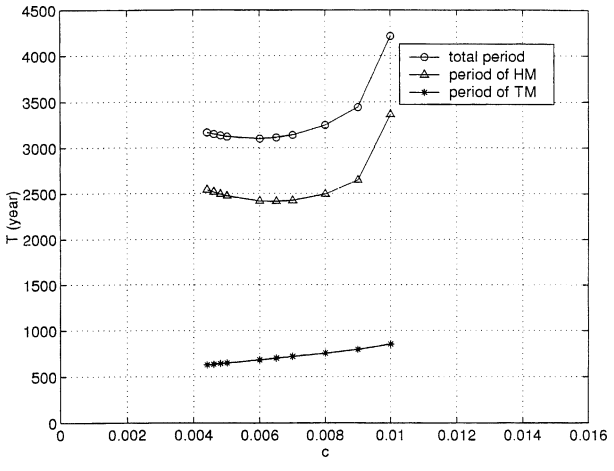

FIG. 13. Oscillation periods as a function of freshwater forcing *c*. The total oscillation period (line with circles), the period of HM during one cycle of the oscillation (line with triangles), and the period of TM during one cycle of the oscillation (line with stars).

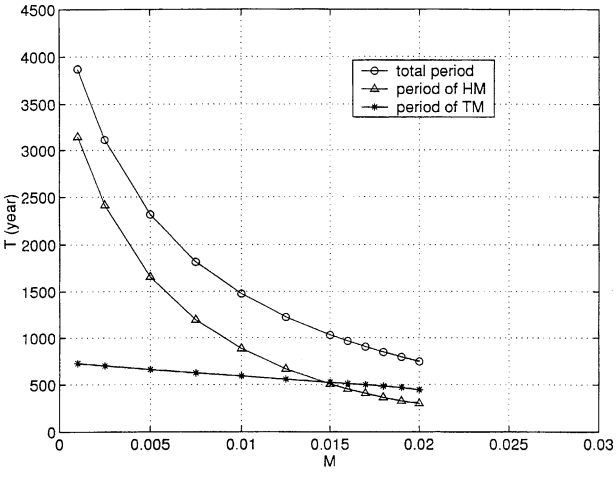

FIG. 14. Oscillation periods as a function of background vertical diffusivity *M*. The total oscillation period (line with circles), the period of HM during one cycle of the oscillation (line with triangles), and the period of TM during one cycle of the oscillation (line with star).

agram of overturning strength and freshwater forcing, there is a window in which the steady HM is unstable and thermohaline oscillations are possible. When the freshwater forcing exceeds an upper limit, only stable HMs exist. For the freshwater forcing below a lower limit, only stable TMs exist. But within a broad window of freshwater forcing the HM oscillates and the oscillation period exhibits a minimum.

Similarly, on the bifurcation diagram of overturning strength and background vertical diffusivity, there is a window where the steady HM is unstable, and thermohaline oscillations are possible. When the background vertical diffusivity exceeds an upper limit, only stable TMs exist. Below that limit thermohaline oscillations are possible, and the oscillation period decreases monotonically as the vertical diffusivity increases.

The box model also shows that, in a warm equable climate with, presumably, a smaller Pole–equator surface temperature gradient, a smaller critical intensity of freshwater flux $E - P$ is required to induce an HM.

c. Implications for paleoclimate and biogeochemical cycles

Global mode switching of the thermohaline circulation is a hypothesis which, on the basis of the work presented here and elsewhere—see, especially Marotzke (1989); Wright and Stocker (1991); Weaver and Sarachik (1991a,b); Weaver et al. (1993); Winton and Sarachik (1993); Huang (1994)—should be taken seriously. Our box model suggests that it may have been much easier to switch into an unstable HM with thermohaline oscillations, during warm equable paleoclimates such as the mid-Cretaceous and late Permian. During glacial periods, the meridional temperature gradient is stronger and the amplitude of freshwater forcing beyond which no steady TM exists would be larger. On the other hand, the freshwater forcing can be significantly enhanced in cold climates by ice melt and a related oscillation may be a possibility.

In our OGCM of the late Permian ocean and the time-dependent solutions of three-box models, a thermohaline oscillation is obtained when the amplitude of the freshwater flux, $E - P$, increased to about 1.3 m yr^{-1}, twice the value of the present climate. Is this a possibility in warm climates? Manabe and Stouffer (1994) showed in a CO_2 quadrupling experiment that the intensity of $E - P$ increased to about 1.5 that of present levels. In the late Permian, the P_{CO_2} level could have been very much higher than that of today (Budyko and Ronov 1979). Moreover, increased dust and sulfate aerosol due to stronger volcanic activity during the late Permian could also induce stronger freshwater flux (Kozur 1998). Here in our OGCM study, we assumed that the spatial distribution of $E - P$ is similar to that of the modern climate. However, little is known about the actual distribution of $E - P$ in the late Permian. Even if each component—evaporation and precipitation—were not very different from today, a change in the spatial distribution of either could have induced significant change in ocean circulation.

In our OGCM of the late Permian ocean and the time-dependent solutions of the three-box model, the thermohaline oscillation is obtained with a vertical diffusivity of $\hat{M} \sim 3 \times 10^{-5}$ m^2 s^{-1}. The physics that control the level of diapycnal mixing in the ocean remain uncertain even for the modern ocean. Measurements of the vertical spread of deliberate tracer releases in the main thermocline (Ledwell et al. 1993) yield 1.1×10^{-5} m^2 s^{-1}, in the lower range of that which is assumed in large-scale ocean circulation models and adopted in our model. The somewhat reduced vertical diffusivity employed in our simulations of the HM is nevertheless within the observed range of mixing in the modern ocean and does not imply that we think the late Permian ocean mixing rate was necessarily weaker than today's.

Thermohaline oscillations can have important implications for ocean biogeochemical cycles. During the sustained HM (between flushing events), the overturning circulation is weak and shallow, and deep ocean transport is dominated by weak, small-scale mixing processes. In this situation abyssal oxygen concentrations can gradually become significantly depleted (Zhang et al. 2001). During the transient TM, however, strong deep convection occurs in the polar region and deep ocean oxygen is replenished. Hence, oscillatory overturning circulation could lead to deep ocean anoxic–oxic cycles with periods ranging between a few hundred and a few thousand years. Unfortunately, the resolution of most geological records for such warm paleoclimates cannot resolve centuries–millennial timescales, and the long periods of HM circulation might dominate the sedimentary record. Indeed it would be very difficult to observe such short periods in the paleorecord even if the mode switching had occurred. Thus signatures of the haline–thermal mode switching in the paleorecord must be studied further, particularly at high temporal resolution.

The haline–thermal mode switching mechanism may also be sensitive to the land–sea distribution. The critical parameters separating different regimes on the bifurcation diagrams may be changed by the land–sea distribution. Thus the impact on the mode switching mechanism of various land–sea distributions at different geological periods during earth's history would be of interest to study.

d. Future work

The three-box model explored here represents only one hemisphere since this is the simplest system with which we can illustrate the underlying physical mechanism of mode switching. The transient TM circulation found in OGCM simulations is, on the other hand, highly asymmetric about the equator. However, during the transient TM, deep downwelling only occurs in one hemisphere, and most of the upwelling occurs within

the same hemisphere. Thus in the OGCM deep ocean temperatures covary with polar SST in the same hemisphere with the other hemisphere playing a passive role. Thus a single-hemisphere box model can catch, we believe, the essential physical mechanism of the thermohaline oscillation. Nevertheless studies with interhemispheric box models are called for and may have wider application than the one studied here.

A previous GCM study of high CO_2 climates (Manabe and Bryan 1985), using a nonlinear equation of state, showed that the overturning circulation strength might not be as sensitive to surface polar–equator temperature gradients as found in our study of a box model with a linear equation of state. However, Manabe and Bryan (1985) employed a low spatial resolution OGCM with weak vertical diffusivity—0.31 $cm^2\ s^{-1}$, and found very sluggish ocean overturning circulation and unrealistically small poleward heat transport in the ''normal'' CO_2 experiment. Studies with the same Geophysical Fluid Dynamics Laboratory (GFDL) OGCM, at similar horizontal resolution but at much higher vertical diffusivity—1 $cm^2\ s^{-1}$ (Weaver et al. 1993; Hotinski et al. 2001) showed that the overturning circulation decreased significantly or even switched into the HM, with smaller surface polar–equator temperature gradients. So the sensitivity of overturning circulation strength to surface polar–equator temperature gradients deserves further study with a nonlinear equation of state.

Localized deep convection in a (generally) stably stratified ocean, which is a key component of these oscillations, is consistent with modern observations of the convective process (Marshall and Schott 1999). In today's ocean, deep polar convection is only observed in confined regions. The warm salty eddies that induce the localized convection (described in section 2) are similar to the convecting eddies found in a previous OGCM study (Winton 1993). The eddies transport warm salty water into the polar region and are important for the onset of polar convection. The mechanism of formation of these eddies is worthy of further investigation and is clearly related to rotating baroclinic fluid dynamics with baroclinic instability a likely candidate. The phenomenon is parameterized simply in the box model, which does not explicitly account for rotational effects. Localized convective activity can also be found in 2D OGCMs (Marotzke 1989) when the polar density distribution is not homogeneous for a particular freshwater forcing profile.

Both the OGCM and box model studied in this paper are forced with prescribed distributions of surface atmospheric temperature and freshwater fluxes. Do the circulation regimes obtained exist in a more realistic, dynamic, coupled atmosphere–ocean system? Since the atmosphere provides negative feedback to the air–sea heat flux, we believe that such oscillatory solutions do exist in a coupled model. But this needs to be demonstrated by further study.

Finally, it should be emphasized that our knowledge of past ocean circulation emerges largely from the sedimentary record, through the preservation of isotopic signatures or organic material. To connect our models more closely to the data record, and in order to understand the implications of these significant global climate changes for the biogeochemical system, we are implementing biogeochemical cycles ($\delta^{13}C$, O_2, etc.) into the models reported here. The results will be addressed in a forthcoming paper.

Acknowledgments. We thank Jochem Marotzke for helpful discussion on the box model. We also would like to thank the constructive comments from two reviewers of the paper. This research is supported by NSF Grant OCE-9819488.

APPENDIX A

OGCM Configuration

The late Permian land–sea distribution was quite different from that of today. The supercontinent Pangaea extended from the North Pole to the South Pole, the Panthalassic superocean covered nearly 70% of earth's surface and connected to the small, narrow Tethys sea. A paleobathymetry was kindly provided by D. B. Rowley and used to configure our ocean circulation model for the late Permian. Although epicontinental regions are represented, the deep ocean is assumed to have a flat bottom with a uniform depth of 4000 m due to lack of other evidence. The ocean domain of the model extends from 70°N to 70°S. The surface forcing fields are based on an atmospheric circulation model of the late Permian (Kutzbach et al. 1989, 1990) and we enhanced the amplitude of the freshwater flux $E - P$. All surface forcing patterns were zonally averaged (Fig. A1), adjusted to be symmetric about the equator and then these zonally averaged fields were used to drive the 3D circulation. Mixed surface boundary conditions were employed—the sea surface temperature (SST) is relaxed to a prescribed surface air temperature showed in Fig. A1b (Haney 1971), but, to avoid spurious feedbacks between the freshwater flux and local salinity, the effect of precipitation is represented as a salinity flux, thus, $Q_S = \hat{S}_o(E - P)$, where $\hat{S}_o$ is the mean surface salinity of the ocean. The mean ocean salinity is 34.6 psu. The relaxation timescale for SST in this simulation is about 160 days, and this damping strength is much weaker than that used in previous OGCM simulations with mixed boundary conditions (Weaver and Sarachik 1991b; Marotzke and Willebrand 1991). This corresponds to a sensitivity of surface heat flux to changes in SST of about 15.6 $W\ m^{-2}\ K^{-1}$ for a 50-m-deep mixed layer, which is close to the sensitivity for observed SST anomalies of typical scales (Seager et al. 1995). This weak damping prevents our simulations from excessive variability (Zhang et al. 1993).

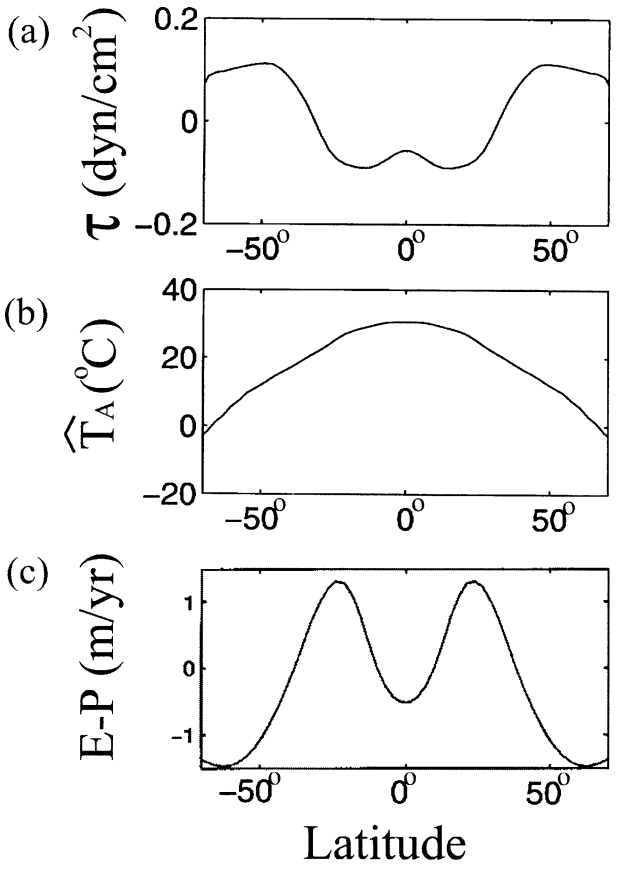

FIG A1. Surface atmospheric forcing fields used to drive the OGCM: (a) surface wind stress, (b) surface air temperature, (c) and freshwater forcing $E - P$.

APPENDIX B

Dimensional Dynamic Equations for the Three-Box Model

$$\frac{d\hat{T}_l}{dt} = \lambda(\hat{T}_{Al} - \hat{T}_l) + \frac{q}{2V_l}(\hat{T}_d - \hat{T}_h) + \frac{|q|}{2V_l}(\hat{T}_d + \hat{T}_h - 2\hat{T}_l) + \frac{\hat{K}}{L^2}(\hat{T}_h - \hat{T}_l) + \frac{\hat{M}_l}{\frac{h+H}{2}h}(\hat{T}_d - \hat{T}_l) \tag{B1}$$

$$\frac{d\hat{T}_h}{dt} = \lambda(\hat{T}_{Ah} - \hat{T}_h) + \frac{q}{2V_h}(\hat{T}_l - \hat{T}_d) + \frac{|q|}{2V_h}(\hat{T}_l + \hat{T}_d - 2\hat{T}_h) + \frac{\hat{K}}{L^2}(\hat{T}_l - \hat{T}_h) + \frac{\hat{M}_h}{\frac{h+H}{2}h}(\hat{T}_d - \hat{T}_h) \tag{B2}$$

$$\frac{d\hat{T}_d}{dt} = \frac{q}{2V_d}(\hat{T}_h - \hat{T}_l) + \frac{|q|}{2V_d}(\hat{T}_h + \hat{T}_l - 2\hat{T}_d) + \frac{\hat{M}_l}{\frac{h+H}{2}2H}(\hat{T}_l - \hat{T}_d) + \frac{\hat{M}_h}{\frac{h+H}{2}2H}(\hat{T}_h - \hat{T}_d) \tag{B3}$$

$$\frac{d\hat{S}_l}{dt} = \frac{F\hat{S}_0}{h} + \frac{q}{2V_l}(\hat{S}_d - \hat{S}_h) + \frac{|q|}{2V_l}(\hat{S}_d + \hat{S}_h - 2\hat{S}_l) + \frac{\hat{K}}{L^2}(\hat{S}_h - \hat{S}_l) + \frac{\hat{M}}{\frac{h+H}{2}h}(\hat{S}_d - \hat{S}_l) \tag{B4}$$

$$\frac{d\hat{S}_h}{dt} = -\frac{F\hat{S}_0}{h} + \frac{q}{2V_h}(\hat{S}_l - \hat{S}_d) + \frac{|q|}{2V_h}(\hat{S}_l + \hat{S}_d - 2\hat{S}_h) + \frac{\hat{K}}{L^2}(\hat{S}_l - \hat{S}_h) + \frac{\hat{M}_h}{\frac{h+H}{2}h}(\hat{S}_d - \hat{S}_h) \tag{B5}$$

$$\frac{d\hat{S}_d}{dt} = \frac{q}{2V_d}(\hat{S}_h - \hat{S}_l) + \frac{|q|}{2V_d}(\hat{S}_h + \hat{S}_l - 2\hat{S}_d) + \frac{\hat{M}_l}{\frac{h+H}{2}2H}(\hat{S}_l - \hat{S}_d) + \frac{\hat{M}_h}{\frac{h+H}{2}2H}(\hat{S}_h - \hat{S}_d). \tag{B6}$$

Here $\hat{S}_0$ is mean salinity of the ocean. At the air–sea surface there is no net salt flux so the total salinity of the three boxes is conserved. Hence, there are only five independent dynamic variables.

APPENDIX C

Properties of Steady-State Solutions

For steady TM ($f > 0$), the nondimensional mean density difference between surface and deep ocean boxes can be obtained from Eqs. (7) to (12) and (6) as

$$\Delta\rho_{ld} = -\frac{f}{\mu_f}\frac{M_h + \gamma f}{M_l + M_h + \gamma f} \tag{C1}$$

$$\Delta\rho_{hd} = \frac{f}{\mu_f}\frac{M_l}{M_l + M_h + \gamma f}. \tag{C2}$$

From the above relations, we know that $\Delta\rho_{ld} < 0$ and $\Delta\rho_{hd} > 0$ since $f > 0$, that is, for steady TM ($f > 0$)

the low-latitude surface is always less dense than the deep ocean, and the high surface is always denser than the deep ocean. Combining with the rules of convective adjustment, we can see that, for physically possible steady TM ($f > 0$), no convection occurs in low latitudes ($M_l = M$) while strong convection happens in polar regions ($M_h = M_{sc}$). Substitute $M_l = M$, $M_h = M_{sc}$ into Eq. (13), we can obtain the steady TM solutions for given parameters.

REFERENCES

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

Budyko, M. I., and A. B. Ronov, 1979: Atmospheric evolution in the Phanerozoic. *Geochem. Int.,* **16,** 1–9.

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

Haupt, B., and D. Seidov, 2001: Warm deep-water ocean conveyor during Cretaceous time. *Geology,* **29,** 295–298.

Hotinski, R. M., K. L. Bice, L. R. Lump, R. G. Najjar, and M. A. Arthur, 2001: Ocean stagnation and end-Permian anoxia. *Geology,* **29,** 7–10.

Huang, R. X., 1994: Thermohaline circulation: Energetics and variability in a single-hemisphere basin model. *J. Geophys. Res.,* **99,** 12 471–12 485.

——, J. Luyten, and H. M. Stommel, 1992: Multiple equilibrium states in combined thermal and saline circulation. *J. Phys. Oceanogr.,* **22,** 231–246.

Johnson, S. J., and Coauthors, 1992: Irregular glacial interstadials recorded in a new Greenland ice core. *Nature,* **359,** 311–313.

Kozur, H. W., 1998: Some aspects of the Permian–Triassic boundary (PTB) and of the possible causes for the biotic crisis around this boundary. *Palaeogeogr., Palaeoclimatol., Palaeoecol.,* **143,** 227–272.

Kutzbach, J. E., and R. G. Gallimore, 1989: Pangaean climates: Megamonsoons of the megacontinent. *J. Geophys. Res.,* **94,** 3341–3358.

——, P. J. Guetter, and W. M. Washington, 1990: Simulated circulation of an idealized ocean for Pangaean time. *Paleoceanography,* **5,** 299–317.

Ledwell, J. R., A. J. Watson, and C. S. Law, 1993: Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature,* **364,** 701–703.

Manabe, S., and K. Bryan, 1985: CO_2-induced change in a coupled ocean–atmosphere model and its paleoclimatic implications. *J. Geophys. Res.,* **90,** 11 689–11 707.

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

Marotzke, J., 1989: Instabilities and multiple steady states of the thermohaline circulation. *Oceanic Circulation Models: Combining Data and Dynamics,* D. L. T. Anderson and J. Willebrand, Eds., Kluwer Academic Publishers, 501–511.

——, 1997: Boundary mixing and the dynamics of three-dimensional thermohaline circulations. *J. Phys. Oceanogr.,* **27,** 1713–1728.

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

Marshall, J., and F. Schott, 1999: Open ocean deep convection: Observations, models and theory. *Rev. Geophys.,* **37,** 1–64.

——, C. Hill, L. Perelman, and A. Adcroft, 1997a: Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *J. Geophys. Res.,* **102,** 5733–5752.

——, A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997b: A finite-volume, incompressible Navier–Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.,* **102,** 5753–5766.

Munk, M., and C. Wunsch, 1998: Abyssal recipes II: Energetics of tidal and wind mixing. *Deep-Sea Res.,* **45,** 1977–2010.

Pierce, D. W., T. P. Barnett, and U. Mikolajewicz, 1995: Competing roles of heat and freshwater flux in forcing thermohaline oscillations. *J. Phys. Oceanogr.,* **25,** 2046–2064.

Railsback, L. B., S. C. Ackerly, T. F. Anderson, and J. L. Cisne, 1990: Palaeontological and isotope evidence for warm saline deep waters in Ordovician oceans. *Nature,* **343,** 156–159.

Rooth, C., 1982: Hydrology and ocean circulation. *Progress in Oceanography,* Vol. 11, Pergamon, 131–149.

Ruddick, B., and L. Q. Zhang, 1996: Qualitative behavior and non-oscillation of Stommel's thermohaline box model. *J. Climate,* **9,** 2768–2777.

Seager, R., Y. Kushnir, and M. A. Cane, 1995: On heat flux boundary conditions for ocean models. *J. Phys. Oceanogr.,* **25,** 3219–3230.

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

Taylor, E. L., T. N. Taylor, and N. R. Cuneo, 1992: The present is not the key to the past: A polar forest from the Permian of Antarctica. *Science,* **257,** 1657–1677.

Toole, J. M., K. L. Polzin, and R. Schmitt, 1994: Estimates of diapycnal mixing in the abyssal ocean. *Science,* **264,** 1120–1123.

Weaver, A. J., and E. S. Sarachik, 1991a: The role of mixed boundary conditions in numerical models of the ocean's climate. *J. Phys. Oceanogr.,* **21,** 1470–1493.

——, and ——, 1991b: Evidence for decadal variability in an ocean general circulation model: An advective mechanism. *Atmos.–Ocean,* **29,** 197–231.

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

Welander, P., 1982: A simple heat–salt oscillator. *Dyn. Atmos. Oceans,* **6,** 233–242.

Winton, M., 1993: Deep decoupling oscillations of the oceanic thermohaline circulation. *Ice in the Climate System,* NATO ASI Series, Vol. 112, Springer Verlag, 417–432.

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

Wright, D. G., and T. F. Stocker, 1991: A zonally averaged model for the thermohaline circulation. Part I: Model development and flow dynamics. *J. Phys. Oceanogr.,* **21,** 1713–1724.

Zhang, R., M. Follows, J. P. Grotzinger, and J. Marshall, 2001: Could the late Permian deep ocean have been anoxic? *Paleoceanography,* **16,** 317–329.

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.