Ocean Dynamics (2001) 52: 26–35

Stefan Rahmstorf

A simple model of seasonal open ocean convection

Part I: Theory

Received: 25 April 2001 / Accepted: 3 July 2001

Abstract A simple tutorial model of open-ocean deep convection is presented, based on the seminal box model of Welander. The model is extended to include prognostic variables of the deep box. An approximate analytical solution is found and discussed for a surface forcing which includes a seasonal cycle. The model helps to illustrate, e.g., the basic balances at work in determining the duration of the winter convection season or the rate dependence of the response of convection to surface warming.

Keywords Seasonal convection · Stability · Box Model · Labrador Sea

1 Introduction

Oceanic convection is a key process in the climate system. In essence, it is a vertical mixing process (Send and Marshall 1995) driven by static instability of the water column. In thermally driven convection, heat is taken from throughout the convecting water column and released to the atmosphere. In the global thermohaline circulation system, convection provides the mechanism of upward heat transport which balances the downward heat diffusion by turbulent diabatic mixing in the stably stratified bulk of the ocean (the latter being the ultimate driving force of the circulation, Sandström 1908). Currents provide the large-scale horizontal heat transport which connects the regions of upward and downward heat flux. In other words, convection, together with wind mixing, is the mechanism by which the heat transported by ocean currents is brought up to the surface to be passed on to the atmosphere. Since mixing heat upwards requires an unstable thermal stratification, wind or other sources of turbulent energy are only needed for heat release where convection is hampered by a stable salinity stratification. Convection can thus be called the chief heat extraction process from the oceans. In the Atlantic, about 1 PW of heat are transported across 24°N (Roemmich and Wunsch 1985) and vented to the atmosphere north of this latitude. The main convection regions, where much of this heat is extracted, are in the Greenland-Norwegian Sea, the Labrador Sea, and possibly the Irminger Sea.

There are many studies investigating the convection process in detail from observations, in laboratory experiments, or with the help of numerical models (for a review see, e.g., Marshall and Schott 1999). Nevertheless, it can be instructive to use the most simple conceptual or "tutorial" models to illustrate some fundamental qualitative aspects of convection. A main benefit of simple models is that analytical solutions can often be found. The seminal convection model of Welander (1982) illustrates fundamental feedbacks of the convection process resulting from the different time scales of heat and freshwater exchange at the surface. They lead to oscillatory solutions for salinity-driven convection, and bistable solutions for thermally driven convection.

One important application of conceptual models is to interpret properties of more complex circulation models. In ocean circulation and climate models, oceanic convection is parameterized as a mixing of vertically adjacent grid boxes when static instability occurs (known as convective adjustment parameterization). Individual grid cells of such models resemble in many respects Welander's conceptual model. Indeed, both the oscillatory and the bistable flip-flop behavior of the conceptual model are found in general circulation models and can play an important role for the climate of such models (Lenderink and Haarsma 1994; Rahmstorf 1995).

In this paper, we extend Welander's model in two respects: a seasonal cycle is introduced to the forcing, and the deep-ocean reservoir is not infinite but has a finite storage capacity and response time scale. In this

Responsible Editor: Jörg-Olaf Wolff

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

way, a simple representation of the heat flow through the system as discussed above is achieved, i.e., the delivery of heat to the deep box by currents and its extraction by seasonal convection. An approximate analytical solution for this system is presented. A number of fundamental questions can be asked of the conceptual model, e.g., what determines the duration of convection each winter? What is the role of the "memory" in the deep water column? How does the system respond to changing surface conditions, e.g., mimicking global warming, on different time scales? Does the bistable behavior of Welander's model still hold for seasonal convection, or is it an artifact of the permanent convection in his model? A companion paper (Kuhlbrodt et al. 2001, part II of this paper) investigates the effects of stochastic variability in the conceptual model.

2 The model

A schematic view of the model is presented in Fig. 1. The model domain is a convection region, surrounded by stratified water. The convective column is modeled with two layers: a surface layer of thickness h, corresponding to the mixed-layer thickness in the absence of convection, and a deep layer of thickness H, representing the water involved in convective overturning. Mixing between the two layers is considered intense during times of convective activity, and absent at other times. Convection is thus grossly simplified as a rapid vertical mixing process which is either "on" or "off"; all subtleties of the process are ignored. A major simplification is that mixing to variable depths is not considered. This is a feature of convection which could be accounted for by including more vertical layers, as is done in one-dimensional column models (e.g., Rahmstorf 1992). To allow analytical treatment, only the simple case with two layers is considered here. The model is driven at the surface by a restoring condition on temperature and a prescribed freshwater flux.

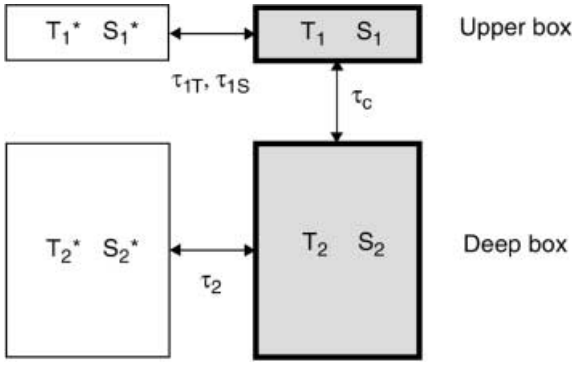

Fig. 1 Schematic of the simple convection box model. The prognostic variables are the temperatures and salinities of the *upper* and (in contrast to Welander's seminal box model) the *deep box*

In these respects, the model is identical to Welander's (1982) flip-flop model. However, we relax Welander's condition that temperature and salinity of the deep box are fixed, allowing the deep water column to adjust in response to convection. Introducing a "memory" in the deep box in this way allows us to investigate a range of phenomena which cannot be modeled by Welander's model, e.g., the transient response to changes in surface forcing. The two model layers are coupled to the surrounding surface and deep waters by simple restoring conditions, which represent the heat and salt exchange due to horizontal mixing and currents.

The prognostic variables of the model are the temperatures and salinities of both layers, with index 1 for the surface and 2 for the deep layer. The model equations are then

$$h\frac{\mathrm{d}T_1}{\mathrm{d}t} = \gamma(T_a - T_1) + c(T_2 - T_1) + u(T_s - T_1) \tag{1a}$$

$$h\frac{\mathrm{d}S_1}{\mathrm{d}t} = -F + c(S_2 - S_1) + u(S_s - S_1) \tag{1b}$$

$$H\frac{\mathrm{d}T_2}{\mathrm{d}t} = c(T_1 - T_2) + q(T_2^* - T_2) \tag{1c}$$

$$H\frac{\mathrm{d}S_2}{\mathrm{d}t} = c(S_1 - S_2) + q(S_2^* - S_2)\ . \tag{1d}$$

The surface heat exchange coefficient is γ and T_a is an apparent atmospheric temperature (including radiation effects), the exchange coefficient with the surrounding surface waters (which are at temperature T_s and salinity S_s) is u, and with the surrounding deep waters (at T_2^* and S_2^*) is q. In Eqs. (1a) and (1b), the surface and horizontal forcing terms can be combined to make them analogous to the deep layer equations (Eqs. 1c and 1d):

$$h\frac{\mathrm{d}T_1}{\mathrm{d}t} = c(T_2 - T_1) + r(T_1^* - T_1) \tag{2a}$$

$$h\frac{\mathrm{d}S_1}{\mathrm{d}t} = c(S_2 - S_1) + u(S_1^* - S_1)\ , \tag{2b}$$

with

$$\begin{aligned} r &= \gamma + u, \\ T_1^* &= \frac{\gamma T_a + u T_s}{\gamma + u}, \\ S_1^* &= S_s - F/u\ . \end{aligned} \tag{3}$$

The convective exchange coefficient is c; this is taken as either very large or zero, depending on whether convection is active:

$$\begin{aligned} &c \gg r, u, q \quad \text{for } \Delta\rho > 0 \\ &c = 0 \quad \text{for } \Delta\rho < 0 \\ &\Delta\rho/\rho_0 := \alpha(T_2 - T_1) - \beta(S_2 - S_1)\ , \end{aligned} \tag{4}$$

where $\Delta\rho$ is the density difference between upper and lower box, and α (unit 1/K) and β (unit 1/psu) are thermal and haline expansion coefficients. Note that we use a linear equation of state for simplicity. With a

nonlinear equation of state, the model can show self-sustained oscillations in a certain parameter regime driven by surface cooling (Cai and Chu 1997). With a linear equation of state, such oscillations are only possible for warm, salty water overlying cooler, fresher water, i.e., for tropical and not for high-latitude convection.

The variables can be made nondimensional by multiplying temperatures with α and salinities with β; we can further subtract T_2^* and S_2^*, since only temperature and salinity *differences* enter the equations; and finally we can divide all nondimensional temperatures and salinities by the negative nondimensional restoring salinity for the surface layer (effectively using this as a density unit). This leads to

$$\begin{aligned}\hat{T}_1 &= \alpha(T_1 - T_2^*)/\beta(S_2^* - S_1^*),\\ T^* &= \alpha(T_1^* - T_2^*)/\beta(S_2^* - S_1^*)\end{aligned} \tag{5}$$

etc., so that in the absence of convection, the nondimensional bottom-layer salinity and temperature are both restored to zero, the surface-layer salinity is restored to -1 and the surface-layer temperature is restored to T^*. Note that the intuitive sign of temperature and salinity changes is preserved in this way (a more negative nondimensional salinity corresponds to a lower salinity), and that a unit increase in temperature cancels a unit increase in salinity in its effect on density; T^* thus measures the strength of the thermal buoyancy forcing relative to the haline buoyancy forcing. We further introduce the thickness ratio $h^* = h/H$, scale the time with the annual period t_a (which we will need later to introduce a seasonal cycle) and convert the coupling constants (which have the dimension m s^{-1}) to time scales (nondimensional, i.e., measured in years): $\tau_1^T = h/rt_a$ is the surface temperature-restoring time scale, $\tau_1^S = h/ut_a$ is the surface salinity-restoring time scale, $\tau_2 = H/qt_a$ is the deep-layer time scale and $\tau_c = H/ct_a$ is a convective mixing time. Omitting the hats, the equations are then simply:

$$\begin{aligned}\dot{T}_1 &= \frac{T_2 - T_1}{h^*\tau_c} + \frac{T^* - T_1}{\tau_1^T},\quad \dot{S}_1 = \frac{S_2 - S_1}{h^*\tau_c} + \frac{-1 - S_1}{\tau_1^S}\\ \dot{T}_2 &= \frac{T_1 - T_2}{\tau_c} - \frac{T_2}{\tau_2},\quad \dot{S}_2 = \frac{S_1 - S_2}{\tau_c} - \frac{S_2}{\tau_2}\ ,\end{aligned} \tag{6}$$

and the condition for convection to occur is

$$T_1 - S_1 < T_2 - S_2\ . \tag{7}$$

The independent parameters of the problem are the strength of the thermal surface forcing T^*, the thickness ratio h^*, and the time scales of the model.

3 Solutions for steady forcing and transitions

The steady solution for constant forcing parameters is, in the absence of convection ($\tau_c \to \infty$), simply

$$T_1 = T^*;\ \ S_1 = -1;\ \ T_2 = 0;\ \ S_2 = 0\ . \tag{8}$$

These values are approached exponentially at the various time scales of upper and deep layer, respectively.

We still need to confirm for which parameter values this solution can exist, i.e., insert Eq. (8) in the convection condition (Eq. 7), leading to the necessary condition for static stability

$$T^* > -1\ . \tag{9}$$

When T^* is less than -1, no stably stratified steady solution is possible, because surface cooling then more than outweighs the effect of surface freshening, making the upper layer denser than the lower layer.

The case with active convection ($\tau_c \to 0$) has the solution

$$T = T_e + (T_i - T_e)\mathrm{e}^{-t/\tau_e^T};\ \ S = S_e + (S_i - S_e)\mathrm{e}^{-t/\tau_e^S} \tag{10}$$

approaching equilibrium values T_e and S_e:

$$T_e = \frac{T^*}{1 + \frac{\tau_1^T}{\tau_2 h^*}};\quad S_e = \frac{-1}{1 + \frac{\tau_1^S}{\tau_2 h^*}};\ \text{with}\ \tau_e^{T,S} = \frac{(1+h^*)\tau_1^{T,S}\tau_2}{\tau_1^{T,S} + h^*\tau_2} \tag{11}$$

T and S are temperature and salinity of the convectively mixed water column (i.e., $T_1 = T_2 = T$). For the steady state, the necessary condition for static instability is

$$\frac{T^*}{\tau_1^T + \tau_2 h^*} + \frac{1}{\tau_1^S + \tau_2 h^*} < 0\ . \tag{12}$$

The left-hand side can be interpreted as a buoyancy flux through the system; depending on how the forcing parameters are set, this can be positive or negative. Note that this condition is *not* the inverse of condition (9) except if the restoring times for temperature and salinity are the same ($\tau_1^T = \tau_1^S$). Depending on the specified forcing, *both* conditions may be satisfied so that there are two steady states (with and without convection), or *neither* may be satisfied, leading to oscillatory convection. This behavior is familiar from Welander's box model.

In contrast to Welander's model, however, the present model displays a subtle property which has also been found in circulation models: a breakdown of convection can be triggered by a sudden *change* in forcing conditions, even if the forcing after the change (as before) would still allow a stable convecting equilibrium. The concept is easily understood in the absence of salinity forcing, for convection driven purely by surface cooling ($T^* < 0$) and a temperature T between T^* and 0. If T^* is abruptly increased beyond T, the buoyancy flux may change sign (the water column being heated from above) even if T^* remains less than zero; convection is then interrupted. In the purely thermal case, convection would eventually start again as the deeper layer warms towards its nonconvecting equilibrium ($T = 0$), but in the presence of surface freshening, a sudden *change* in T^* can trigger a permanent transition from a convecting to a nonconvecting equilibrium.

The conditions for such a shutdown of convection are easily derived; here we include S^* explicitly, rather than setting it to -1, in order to represent also changes in freshwater forcing in an explicit manner. From Eq. (6), differential equations for $(T_1 - T_2)$ and $(S_1 - S_2)$ can be derived, which show that these terms relax to their equilibrium values at the fast convective time scale τ_c, i.e., much faster than any changes in T, so that we can assume them to be in equilibrium at any time. Their equilibrium values are

$$\begin{aligned} \frac{T_1 - T_2}{\tau_c}\left(1 + \frac{1}{h^*}\right) &= \frac{T^*}{\tau_1^T} + T\left(\frac{1}{\tau_2} - \frac{1}{\tau_1^T}\right) \\ \frac{S_1 - S_2}{\tau_c}\left(1 + \frac{1}{h^*}\right) &= \frac{S^*}{\tau_1^S} + S\left(\frac{1}{\tau_2} - \frac{1}{\tau_1^S}\right) \quad , \end{aligned} \tag{13}$$

so that the static instability condition (Eq. 7) for any instantaneous value of T, S becomes:

$$\frac{T^*}{\tau_1^T} + T\left(\frac{1}{\tau_2} - \frac{1}{\tau_1^T}\right) - \frac{S^*}{\tau_1^S} - S\left(\frac{1}{\tau_2} - \frac{1}{\tau_1^S}\right) < 0 \quad . \tag{14}$$

Equation 14 (together with Eq. 11) allows us to calculate, e.g., the critical value up to which T^* can be increased instantaneously without violating the condition for convection, if we start from an equilibrium that corresponds to some previous (or initial) forcing conditions T_p^*, S_p^*:

$$T^*_{\text{crit}} = T_p^* h^* \frac{\tau_2 - \tau_1^T}{\tau_1^T + h^* \tau_2} + S_p^* \frac{(1 + h^*)\tau_1^T}{\tau_1^S + h^* \tau_2} \tag{15}$$

As a simple example, let us consider a case where $T_p^* = -1.5$ and $S_p^* = -1$ and the time scales are $\tau_1^T = 0.42$ (i.e., 5 months), $\tau_1^S = 8$ and $\tau_2 = 20$, with $h^* = 1/36$. These parameters are as in the Labrador Sea example (Kuhlbrodt et al. 2001), except that for illustration the initial T_p^* is chosen so low as to be in the monostable, always convecting regime. According to conditions (9) and (12), the bistable regime, where both equilibria (with/without convection) are possible, is $-1 < T^* < -0.11$. Equation 15 shows that if T^* is *suddenly* increased from -1.5 to -0.89 or warmer, convection will be interrupted.

Note that a *slow* increase of T^* is possible up to a value of -0.11 without interrupting convection. We can compute a critical linear rate of increase of T^* that must not be exceeded to maintain convection; this is a consideration relevant to the global warming problem. Equation 6 can be solved to yield $T(t)$ if the constant T^* is replaced by $T^* + at$, i.e., increasing with time at rate a. In this case,

$$T = T_p + a\frac{T_e}{T^*}\left(t - \tau_e^T + \tau_e^T \mathrm{e}^{-t/\tau_e^T}\right) \quad , \tag{16}$$

as illustrated in Fig. 2. When this is inserted in the stability condition (14) and solved for $a(t)$, we obtain the critical rate of warming a that can be sustained for a time t until convection breaks down. This is easily done; the resulting formula is lengthy and not very illuminating so is omitted here. However, the leading order effect

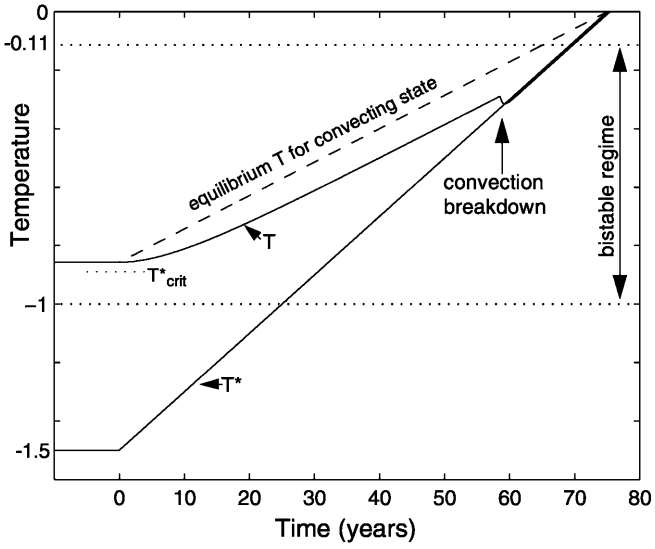

Fig. 2 A "global warming" solution for the box model without seasonal cycle. The diagram illustrates that at faster warming rates convection already breaks down at lower temperature. At year 0, the surface relaxation temperature T^* starts a linear rise at a rate of 2 per century. Convection breaks down at year 58 at a nondimensional temperature of -0.34. Note the small cooling at the time of convection breakdown. For a step-function increase in T^* above T^*_{crit}, convection would also break down. For an infinitely slow rise of T^*, convection is maintained until the limit of the bistable regime at $T^* = -0.11$ is reached. In the range $-1 < T^* < -0.11$ both convecting and nonconvecting solutions are stable; this is a consequence of the salinity forcing, even though this is only weak

is simple for the case with weak salinity forcing $(\tau_1^S \gg \tau_1^T)$ and for $\tau_2 \gg \tau_1^T$. When the exponential adjustment term in (16) has decayed, the actual temperature lags behind the equilibrium temperature for the current forcing (i.e., the first term in the bracket) by the time τ_e^T. It is this lag which makes the buoyancy forcing positive already before the stability threshold would otherwise be reached (namely at $T = -0.11$). It is interesting that τ_e^T (defined in Eq. 11) is the crucial time scale for this problem; this is the response time scale of the whole water column (boxes 1 and 2 combined) under the influence of surface forcing and interior heat transport. Convection breaks down already at a threshold which is lower than the limit of the bistable regime, namely by

$$\Delta T = -a\tau_e^T \frac{\tau_2 h^*}{\tau_1^T} \quad . \tag{17}$$

Thus, the lowering of the temperature threshold is proportional to the *rate* of warming a, and the reason is the thermal inertia τ_e^T of the convectively mixed water column. (For the parameter values in Fig. 2, $\Delta T = -0.23$, resulting in a convection breakdown in year 58 at $T = -0.34$.)

This is the theoretical underpinning of an important fact suggested first by Rahmstorf et al. (1996) and confirmed subsequently in model experiments by Stocker and Schmittner (1997), namely that more global warming can be sustained without interrupting convection if

the warming proceeds at a slower rate. This analysis (and the model of Stocker and Schmittner) does not include the effect of stochastic variability, which would lead to a probabilistic result (warming reducing the probability of convection) rather than a deterministic threshold for a shutdown (Kuhlbrodt et al. 2001). Nevertheless, we expect the rate dependence to apply also in this case. Greenhouse warming has a time scale similar to the adjustment time scale of the deep-water column in convection regions, and coupled global warming scenario runs show greatly reduced winter convection (Manabe and Stouffer 1994). In one general circulation model (Wood et al. 1999), convection in the Labrador Sea shuts down early in the 21st century as a result of global warming. In all of these models, convection is parameterized in a simple way similar to our analytical model, albeit for more vertical levels. To what extent the feedbacks leading to the rate dependence are also important in the far more complex real ocean requires further study.

4 Seasonal convection

4.1 General considerations

We now add a seasonal cycle to the thermal forcing of the surface layer,

$$T^*(t) = T_0^* + A \sin 2\pi t \ , \tag{18}$$

where A is the nondimensional amplitude of the seasonal forcing cycle (note that time is nondimensionalized with the annual period). If the parameters are chosen appropriately, the model will now go through a period of winter convection, interrupted by a time of stratification. A typical seasonal cycle is shown in Fig. 3; the analytical solution describing this cycle is discussed in the following.

First we consider the *convective phase* of the cycle, again assuming vigorous convection ($\tau_c \to 0$), with $T_1 \approx T_2 =: T$. The temperature evolution equation is

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{(T_e - T)}{\tau_e^T} + \frac{h^*}{(1+h^*)\tau_1^T} A \sin 2\pi t \ , \tag{19}$$

with $T_e := T_0^*/(1 + \tau_1^T/h^*\tau_2)$ and τ_e^T as in Eq. (11). The solution is

$$T = T_e + A_c \sin 2\pi(t-\varphi) + C\mathrm{e}^{-t/\tau_e^T} \ . \tag{20}$$

As expected, this describes a forced oscillation of temperature, with an amplitude

$$A_c = \frac{A}{\sqrt{[2\pi\tau_1^T(1+h^*)/h^*]^2 + (1+\tau_1^T/h^*\tau_2)^2}} \tag{21}$$

and a phase lag $\varphi_c = \arctan(2\pi\tau_e^T)/2\pi$. The amplitude is reduced if the surface coupling is weak (large τ_1^T) or if the damping through the deep layer is strong (small τ_2); the scaling factor h^* enters only because time scales rather than dimensional coupling constants (as in Eq. 1) are used. In addition, there is an exponential term in (20) describing the decay of an initial deviation from the steady cycle; the constant C (like C_1 and C_2 below) is determined by the initial condition.

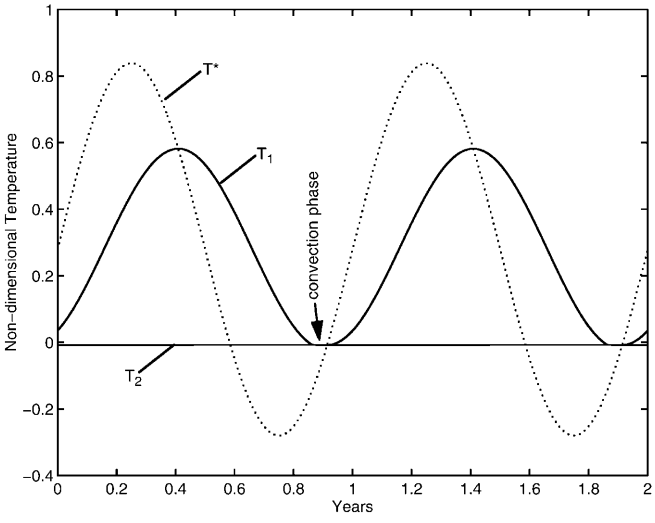

Fig. 3 A typical seasonal temperature cycle for realistic parameter magnitudes in the absence of salinity forcing. The relaxation temperature T^* is given by Eq. (18) (and arbitrarily sets the phase of the annual cycle – the time axis does not relate to particular calendar months). Surface temperature T_1 follows Eq. (22) and deep temperature T_2 follows Eq. (24) except during the brief convective phase, where both follow Eq. (20). Due to the much greater thickness and thermal inertia of the deep layer (in this case, 36 times that of the surface layer), the seasonal variation of T_2 is very small. A zoom into the winter months of this cycle is presented in Fig. 4

In the *nonconvective phase* ($\tau_c \to \infty$), the surface layer is decoupled from the deep layer, and the solution is

$$T_1 = T_0^* + A_n \sin 2\pi(t-\varphi) + C_1\mathrm{e}^{-t/\tau_1^T} \ , \tag{22}$$

with amplitude

$$A_n = \frac{A}{\sqrt{1+(2\pi\tau_1^T)^2}} \ , \tag{23}$$

and a phase lag $\varphi_n = \arctan(2\pi\tau_1^T)/2\pi$. The minimum temperature would be $T_1^{\min} = T_0^* - A_n$ if there was no convection throughout the year, but if convection does take place then $T_1^{\min}$ is never quite reached. The deep layer simply approaches 0 during the decoupled phase according to

$$T_2 = C_2\mathrm{e}^{-t/\tau_2} \ . \tag{24}$$

Salinities, in both cases, approach their respective equilibrium values (see previous section) following similar exponential laws.

We can now match the solutions for the convective and stratified periods to obtain a periodic seasonal solution. The conditions for matching the curves are that convection sets in when the boxes become unstably stratified, and convection stops as soon as the (nonconvective) buoyancy gain of the upper box is faster than that of the lower box. Heat and salt content of the

whole water column are continuous at the onset of convection (at time t_1), while temperature and salinity in each box undergo a discontinuity there due to the instantaneous convective mixing; the latter are, however, continuous across the end of convection at time t_0. Finding the start and end times t_1 and t_0 of seasonal convection analytically is not possible in general (leading, e.g., to conditions of the type $\sin t_1 = a\exp(bt_1)$ to find t_1). However, the exact form is hardly interesting, given the idealized nature of the model. What is interesting is only the physically meaningful, leading-order parameter dependence of the convection times for a short convection season. This can be derived by Taylor series expansion of the functions involved. To avoid burdening the reader with algebra, this will be carried out here only for a simple, but realistic and instructive case.

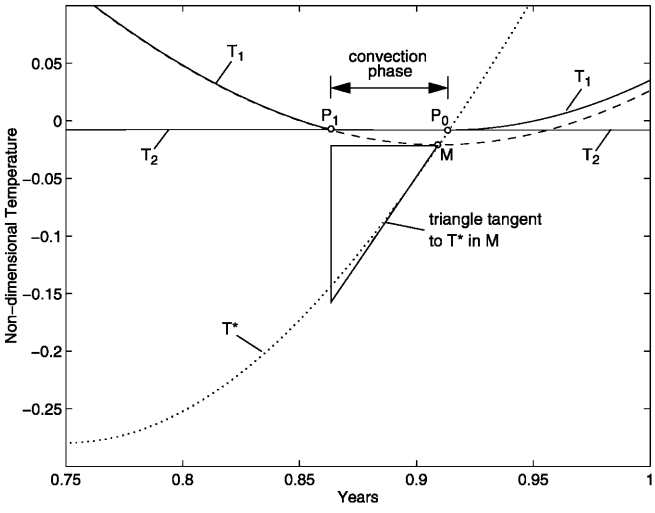

Fig. 4 Closeup view of the convective phase of the seasonal cycle shown in Fig. 3. Convection starts at point P_1 and ends at point P_0, lasting for 19 days. The *dashed line* shows the analytical solution of the seasonal cycle of T_1 in the absence of convection from Eq. (22) (equilibrated, i.e., without the exponential spinup term), reaching its minimum ($T_1^{\min}$) at point M. The actual T_1 never reaches this minimum, as convection sets in beforehand. After convection ends at P_0, T_1 again follows Eq. (22), exponentially approaching the dashed equilibrium solution. The time variation of T_2 is too small (compared to that of T_1) to be seen on the scale of the plot, but is a crucial part of the heat budget; see condition (3) discussed in the text

4.2 Convection duration for purely thermal convection

The case considered in the following is a purely thermal model (without salinity) in the limit of short convection duration ($t_0 - t_1 \ll 1$). The first consequence of this is that the seasonal cycle of the deep temperature T_2 is small compared to that of T_1; during a short convection season not much heat can be released to the atmosphere, and the deep heat reservoir is much larger than that of the thin surface layer. This idea is supported by data from Ocean Weather Station Bravo in the Labrador Sea (Lazier 1980) and simplifies the algebraic conditions considerably. There are three such conditions; Fig. 4 helps to illustrate them.

Condition 1: End of convection. Convection stops during the warming phase, as soon as the surface warms faster than the deep layer:

$$\frac{\mathrm{d}T_1}{\mathrm{d}t} = \frac{\mathrm{d}T_2}{\mathrm{d}t}\Big|_{(\text{nonconvective})} \tag{25}$$

where the nonconvective warming rate of T_2 is $-T_2/\tau_2$ (Eq. 6) and can be neglected ($T_2 \approx$ const). Therefore, the condition is met as soon as the relaxation temperature T^* exceeds T_1 (which is equal to T_2 during convection). This leads to the condition

$$T^*\,(t = t_0) = T_0^* - A + 2\pi^2 A\left(t_0 - \frac{3}{4}\right)^2 = T_2 \ , \tag{26}$$

where the sinusoidal seasonal cycle (Eq. 16) has been expanded quadratically around its minimum at $t = 3/4$. This condition is easily solved for t_0:

$$t_0 = \frac{3}{4} + \sqrt{\frac{A - T_0^* + T_2}{2\pi^2 A}} \ . \tag{27}$$

An even simpler approximate condition is obtained by assuming that

$$t_0 = \frac{3}{4} + \varphi_n \ , \tag{28}$$

i.e., that convection ends at the minimum of the nonconvective surface temperature cycle given by Eq. (20) (point M in Fig. 4). Figure 4 shows that the convection duration $t_0 - t_1$ is in the leading order determined by the time that t_1 precedes point M, not by the time that t_0 lags point M, because the surface temperature cycle has zero time derivative at M.

Condition 2: Start of convection. Convection restarts when the falling surface temperature T_1 intersects T_2 (see Fig. 4). T_1 is given by Eq. (22). Again expanding the sinusoidal and neglecting the exponential term ($\tau_1^T \ll 1$), we obtain the condition

$$T_1 = T_1^{\min} + 2\pi^2 A_n\left(t_1 - \varphi_n - \frac{3}{4}\right)^2 = T_2 \ . \tag{29}$$

This condition can be solved for t_1:

$$t_1 = \frac{3}{4} + \varphi_n - \sqrt{\frac{T_2 - T_1^{\min}}{2\pi^2 A_n}} \ . \tag{30}$$

Using Eq. (28), the duration of winter convection then is

$$\Delta t_c = t_0 - t_1 = \sqrt{\frac{T_2 - T_1^{\min}}{2\pi^2 A_n}} \ . \tag{31}$$

It is thus proportional to the square root of the difference between the deep temperature T_2 and the minimum temperature of the nonconvective surface cycle. In order to determine this temperature difference, we need a third condition.

Condition 3: Annual heat budget. A balanced annual heat budget requires that the heat accumulated by

the deep layer throughout the year is equal to the heat lost during the brief convection season; this heat budget determines T_2. The heat accumulated in the absence of convection in a year is to good approximation

$$\Delta T_2|_{(\text{nonconvective})} = -T_2/\tau_2 \ . \tag{32}$$

During convection, the deep box is relaxed towards T^* (Eq. 18) with a time constant $\tau_1^T(1+h^*)/h^*$ during time interval Δt_c. The resulting temperature change is proportional to the surface area between the line P_1P_0 and the T^* curve in Fig. 4. We can approximate this by the triangle shown in Fig. 4, which assumes that $T_2 = T_1^{\min}$ and T^* has a constant slope of $4\pi^2 A\varphi_n$ (namely its slope at $t = \varphi_n$ in quadratic approximation to Eq. 18):

$$\Delta T_2|_{(\text{convective})} = -\frac{2\pi^2 A\varphi_n h^*}{\tau_1^T(1+h^*)}(\Delta t_c)^2 \ . \tag{33}$$

Requiring that these two temperature changes are equal but opposite, and using Eq. (29), finally leads to the third condition:

$$T_2 - T_1^{\min} = -\frac{T_1^{\min}}{\dfrac{\tau_2 h^* \varphi_n \sqrt{1+(2\pi\tau_1^T)^2}}{\tau_1^T(1+h^*)} + 1} \ . \tag{34}$$

Given $\tau_1^T \ll 1$, $\varphi_n \approx \tau_1^T$, and $h^* \ll 1$, this simplifies to

$$T_2 - T_1^{\min} = \frac{-T_1^{\min}}{h^*\tau_2 + 1} \ , \tag{35}$$

and gives (with Eq. 31) the convection duration to leading order as

$$\Delta t_c = \sqrt{\frac{-T_1^{\min}}{2\pi^2 A_n(h^*\tau_2 + 1)}} \ . \tag{36}$$

Although it would have been difficult to guess this result in advance, it is not difficult to interpret. For very short deep restoring time scale ($\tau_2 \to 0$) the deep temperature is essentially fixed at $T_2 = 0$ (this limit corresponds to Welander's original box model with prescribed deep temperature). Equation (36) then just reflects the leading-order (i.e., quadratic) expansion of the sinusoidal cycle around its minimum (point M in Fig. 4) – convection simply lasts from the time where the surface temperature dips below zero up until point M. Whatever heat is extracted from the deep layer during convection can be easily replenished due to the fast τ_2. However, when $h^*\tau_2$ becomes comparable to a year or longer, the deep temperature starts to be lowered significantly by the convective cooling – this shortens the convection season. This is the realistic case, when the need to balance the annual heat budget determines the duration of convection and by how much the deep temperature is lowered. For infinitely slow adjustment ($\tau_2 \to \infty$), convection duration vanishes, as in this case the deep box does not accumulate any heat during the year, so none needs to be released to the surface by winter convection.

The response time of the deep water column thus controls not only the behavior of convection during periods of climate change (as already discussed), but also the seasonal behavior. The presence of seasonal convection in fact implies a net heat flow through the system with subsurface warming through advection or mixing of heat throughout the year, as is the case in the North Atlantic due to the thermohaline circulation. The heat brought north by this circulation is then released during the convection season; the larger the heat transport, the longer the convection season will be. This is an important aspect of convection modeled in the box model with variable deep temperature, which is distilled in Eq. (36).

One might have expected that the surface layer relaxation time enters the equation, and that the shortness of the convection season reflects the relation of the fast heat loss during convection to the slow heat gain during the rest of the year. This is not the case: for a faster surface relaxation time, $T_1^{\min}$ becomes lower and thus convection duration actually increases (Fig. 5).

4.3 Influence of salinity on the seasonal cycle

The influence of salinity is responsible for the rich and interesting stability behavior of the box model (i.e., the existence of multiple equilibria) because of the weaker feedback of salinity on freshwater flux compared to the feedback of temperature on heat flux (i.e., because $\tau_1^S > \tau_1^T$). Nevertheless, the salinity influence is

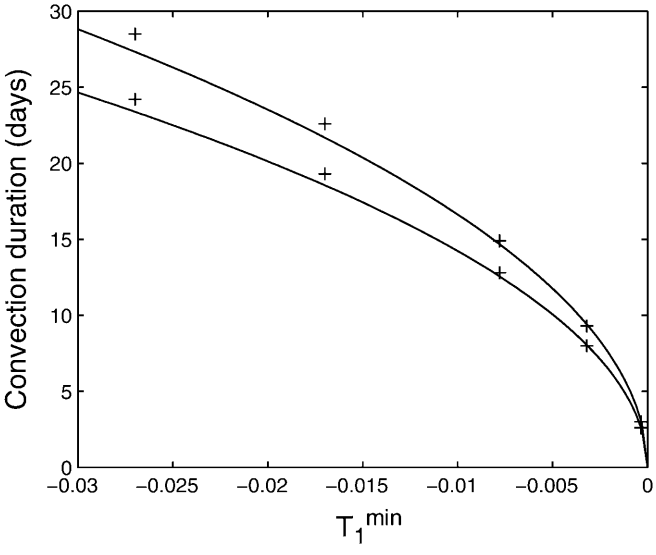

Fig. 5 Convection duration as a function of minimum winter temperature. *Crosses* show numerical results, *solid lines* the approximate analytical solution given in Eq. (36). The *upper curve* is for $\tau_2 = 5$, the *lower curve* for $\tau_2 = 20$. Other parameters correspond to the standard set found for the Labrador Sea (see Kuhlbrodt et al. 2001), except for the absence of salinity forcing. The analytical solution, valid for short convection times, slightly underestimates convection duration, but the error is small (e.g., 1 day for a convection duration of 20 days)

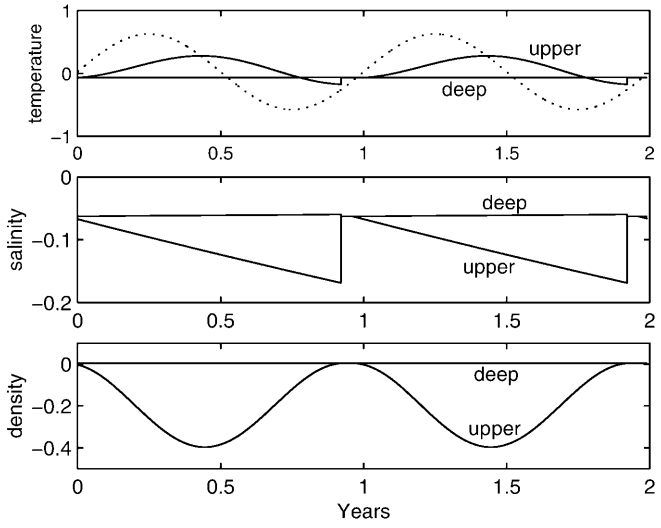

Fig. 6 Example of a seasonal cycle with salinity. This example corresponds to the optimal parameter set found for the Labrador Sea (see Kuhlbrodt et al. 2001), but is shown here for non-dimensional variables and without seasonal cycle in the salinity forcing. Note the freshening of the surface layer throughout most of the year as given by Eq. (37), leading to a salinity difference ΔS at the onset of convection. For the nondimensional variables, one unit of temperature and salinity correspond to one density unit. *Lines* as in Fig. 3

surprisingly easy to include in the seasonal cycle solution discussed above. This is because τ_1^S and τ_2 are typically both much longer than 1 year. A typical seasonal cycle in the presence of salinity is shown in Fig. 6. Salinity decreases in the upper layer and increases in the lower layer during the non-convective part of the year, for the typical thermally driven convection in regions of net surface freshwater input. Convection then simply mixes the two to an average salinity S; given the long time scales of salinity forcing we can assume that S is constant during the brief convection season. During the nonconvective phase, surface and deep salinity drift exponentially away from this mixed-salinity value during 1 year (assuming short convection duration):

$$\begin{aligned} \Delta S_1 &= (S+1)(\mathrm{e}^{-1/\tau_1^S} - 1) \\ \Delta S_2 &= S(\mathrm{e}^{-1/\tau_2} - 1) \end{aligned} \tag{37}$$

The salinity difference ΔS between the layers at the start of convection, i.e., after 1 year, is the difference of the individual changes $(\Delta S_1 - \Delta S_2)$. The value of the mixed salinity S is simply determined by the requirement to balance the freshwater inputs of upper and lower layer over 1 year, $h^* \Delta S_1 + \Delta S_2 = 0$, leading to

$$S = -\frac{h^*(\mathrm{e}^{-1/\tau_1^S} - 1)}{h^*(\mathrm{e}^{-1/\tau_1^S} - 1) + \mathrm{e}^{-1/\tau_2} - 1} \quad . \tag{38}$$

For $\tau_2, \tau_1^S \gg 1$, this salinity is the same as the solution for steady forcing and permanent convection given in Eq. (11).

The salinity cycle is thus independent of the temperature cycle except for the crucial fact that the temperature cycle forces convective mixing once per year. However, the presence of a salinity cycle does alter the temperature cycle in the following three ways (illustrated in Fig. 7), which correspond to changes in the three conditions for convection duration given in Section 4.2.

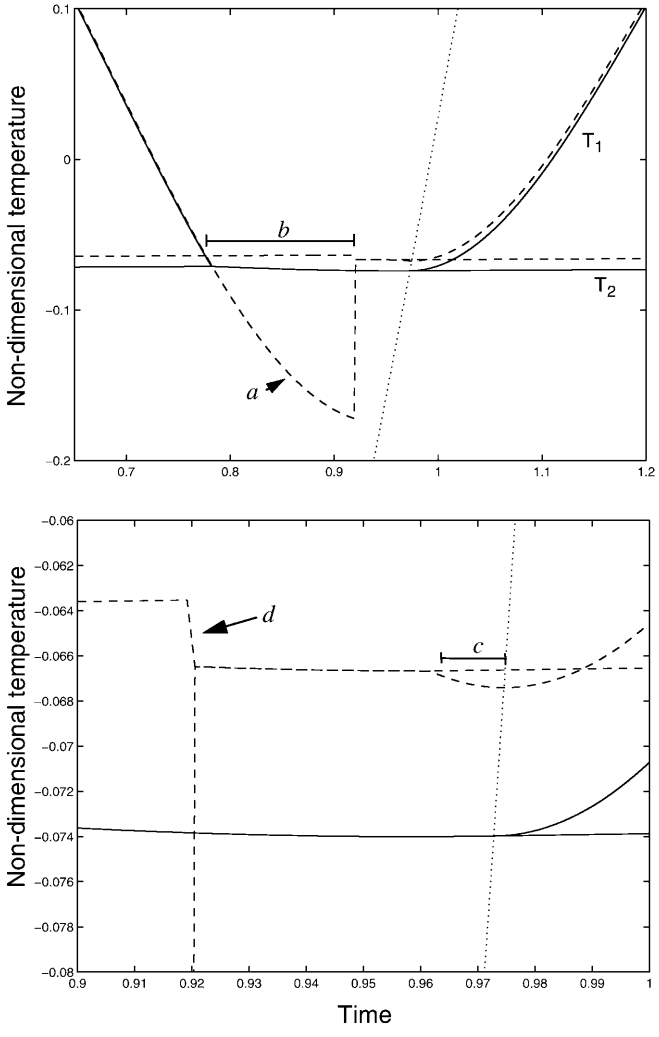

Fig. 7 Illustration of the salinity effect on the temperature cycle. The graphs show the temperature evolution during the convective phase of the seasonal cycle for a temperature-only case (*solid lines*) and with added salinity forcing (*dashed lines*, other parameters remaining the same). The *lower panel* shows a zoom into part of the upper panel. Key changes to the temperature cycle due to the inclusion of salinity, as discussed in the text, are: *a* overshooting of surface temperature below the deep temperature; *b* delayed onset of convection; *c* earlier end of convection (while surface layer still cools); *d* deep layer cooling at onset of convective mixing. The first-order impacts are thus a much colder surface temperature in winter and a greatly reduced convection duration, with only a small (warming) effect on the deep temperature. The *dotted line* is the relaxation temperature as in Figs. 3 and 4

1. Convection ends slightly earlier compared to the no-salt case, due to the freshwater forcing promoting stable stratification. This effect is easily computed, given the mixed salinity of Eq. (38).

2. Convection will now start not when the layer temperatures are equal but rather at a temperature

difference $\Delta T = \Delta S$, i.e., when the surface is cold enough to overcome the stable salinity stratification of the water column. This shortens the convection season and leads to an overshooting of surface temperature below the deep temperature. This shortening of the convection season is much larger than that due to effect (1), as it results from the freshwater accumulated over a whole year.

3. The temperature change of the deep layer during the convection season has now two terms: a cooling by $h^*\Delta T$ right at the onset of convection, due to mixing with the colder upper layer, and the cooling during the convection season as given in Eq. (33). For a realistic magnitude of freshwater forcing and short convection season (e.g., the parameter set applicable to the Labrador Sea derived in Kuhlbrodt et al. 2001), the former effect strongly dominates. Note that the total cooling of the deep layer still equals the total surface heat loss starting from the time where the surface temperature equals the deep temperature, just as in the no-salt case. The difference is that the surface temperature becomes colder in the absence of convective mixing, so that the heat loss through the relaxation term is less. For this reason, the deep temperature must be warmer than in the no-salt case.

Another simple property of salinity is that a seasonal surface freshwater flux (with zero annual mean) can be added (as is done in Kuhlbrodt et al. 2001) to produce a realistic seasonal surface salinity cycle without affecting other aspects of the solution. Only proviso is that this added freshwater flux is not too strong at the end of the convection season, so that effect (1) does not become too large.

We thus have found an approximate analytical solution for a complete seasonal cycle including salinity forcing, that is valid for a wide range of realistic parameter choices. The full equations are not shown here as they are not very instructive; the major balances at work in determining the convection season and deep water properties have already been demonstrated in the derivation of the no-salt case, and the modifications due to salinity are easily grasped intuitively. Adding salinity effects (assuming this is a net surface freshening) leads to a colder surface temperature in winter, a much shorter convection season, a slightly warmer deep temperature, and a correspondingly reduced annual heat flow through the system.

5 Discussion

We have extended Welander's seminal box model of convection in two ways:

1. We made the deep box temperature and salinity prognostic to study how they are affected by the presence or absence of convection.
2. We introduced a seasonal cycle to investigate the effect of winter (rather than permanent) convection.

An approximate analytical solution is presented for the extended convection model. The key to finding this solution is to determine the beginning and end of the convection season; separate solutions for the time evolution during convective and nonconvective parts of the year are readily calculated.

Convection duration is determined by the requirement to balance the heat gain of the deep box throughout the nonconvective bulk of the year with the heat lost during the brief convection season. This balance is determined by the surface temperature cycle, i.e., for how long and how much surface temperature can drop below the deep temperature, and by the time scale at which the deep box takes up heat in the absence of convection. The slower this time scale (i.e., the smaller the interior heat transport into the convective region), the shorter the convection season.

The effect of introducing salinity (i.e., surface freshening balanced by salt input in the interior) is to greatly reduce the length of the convection season, with only a comparatively small effect on the heat budget and deep temperature. Surface temperature is affected in winter because the onset of convection is delayed by the stable salinity stratification, so that surface temperature drops significantly below the deep temperature before convection sets in.

The simple model highlights the role of the heat and salt budgets of the deep layer and associated feedback effects in determining the duration of convection events and the deep water properties. Although the seasonal cycle studied here is highly idealized, these are basic constraints which also need to be satisfied in more complex situations which cannot be treated analytically, e.g., in the presence of stochastic climate variability. This case is studied in Kuhlbrodt et al. 2001.

The basic stability behavior of Welander's model, i.e., the existence of two equilibria in a certain parameter regime, applies also to the extended model. If convection is suppressed in one winter, the deep temperature starts to drift to warmer conditions (which would tend to make convection easier in subsequent winters), but at the same time a stronger salinity stratification develops with freshwater accumulating in the surface layer. Thus, a brief perturbation can flip the model to another stable state without winter convection. To what extent such a state transition triggered by a short pulse (rather than a lasting anomaly in the forcing) may explain the Great Salinity Anomaly of the early 1970s in the Labrador Sea will also be examined in Kuhlbrodt et al. 2001.

Introducing a "memory" in the deep water column, which responds over a finite time scale to changes in forcing, makes the model's reaction to a gradual surface warming dependent on the rate of this warming. This is a simple candidate mechanism for explaining the rate dependence of the response of the thermohaline circulation and convection to global warming that has been found in more complex models. This is an example of how the simple model can be used as a 'tutorial' tool to help in the interpretation of climate model results or observations of the real climate system.

Acknowledgements I am grateful for the hospitality of and discussions with M. England and H. van den Budenmeyer during a stay at the University of New South Wales in Sydney, where part of this work was done.

References

Cai W, Chu PC (1997) A thermal oscillation under a restorative forcing. Q J R Met Soc 124: 793–809

Kuhlbrodt T, Titz S, Feudel U, Rahmstorf S (2001) A simple model of seasonal open ocean convection. Part II: Labrador Sea stability and stochastic forcing. Ocean Dyn 52: 36–49

Lazier JRN (1980) Oceanographic conditions at Ocean Weather Ship Bravo, 1964–74. Atmos Ocean 18: 227–238

Lenderink G, Haarsma RJ (1994) Variability and multiple equilibria of the thermohaline circulation, associated with deep water formation. J Phys Oceanogr 24: 1480–1493

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

Marshall J, Schott F (1999) Open-ocean convection: observations, theory, and models. Rev Geophys 37: 1–64

Rahmstorf S (1992) Modelling ocean temperatures and mixed-layer depths in the Tasman Sea off the South Island, New Zealand. New Zealand J Mar Freshwater Res 26: 37–51

Rahmstorf S (1995) Multiple convection patterns and thermohaline flow in an idealised OGCM. J Clim 8: 3028–3039

Rahmstorf S, Marotzke J, Willebrand J (1996) Stability of the thermohaline circulation. In: Krauss W (ed) The warm water sphere of the North Atlantic ocean. Borntraeger, Stuttgart, pp 129–158

Roemmich DH, Wunsch C (1985) Two transatlantic sections: meridional circulation and heat flux in the subtropical North Atlantic Ocean. Deep-Sea Res 32: 619–664

Sandström JW (1908) Dynamische Versuche mit Meerwasser. Ann Hydrogr Marit Meteorol 36: 6–23

Send U, Marshall J (1995) Integral effects of deep convection. J Phys Oceanogr 25: 855–872

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

Welander P (1982) A simple heat-salt oscillator. Dyn Atmos Oceans 6: 233–242

Wood RA, Keen AB, Mitchell JFB, Gregory JM (1999) Changing spatial structure of the thermohaline circulation in response to atmospheric CO_2 forcing in a climate model. Nature 399: 572–575