Comments on "Instability of the Thermohaline Circulation with Respect to Mixed Boundary Conditions: Is It Really a Problem for Realistic Models?"

STEFAN RAHMSTORF

Institut für Meereskunde, Universität Kiel, Kiel, Germany

19 September 1994 and 10 March 1995

ABSTRACT

This comment discusses two issues raised by a recent study of the stability of the thermohaline circulation (Tziperman et al.). A numerical problem is pointed out that may have affected the results of the reported calculations. In addition, it is argued that the state transition of the model circulation found in the above study is triggered by a *convective* instability mechanism, not the *advective* mechanism discussed by its authors. Should this be the case, then the conclusion of Tziperman et al. that the ocean is close to a stability transition point must be questioned.

1. Introduction

Tziperman et al. (1994, hereafter TTFB) have recently published an analysis of the stability of the thermohaline circulation in an OGCM when it is switched from restoring boundary conditions to mixed boundary conditions. They find that the stability of their model during this switch depends on the salinity restoring time used during spinup. If the model is spun up to equilibrium with a salinity restoring time of 30 days, the thermohaline circulation collapses upon transition to mixed boundary conditions. On the other hand, if it is spun up with a restoring time of 120 days the circulation remains stable. The difference is attributed to the different freshwater fluxes diagnosed from the two spinup experiments and used subsequently after switching to mixed boundary conditions.

2. Numerical artifact related to low Baltic salinity

The diagnosed freshwater flux for both experiments of TTFB is shown in their Figs. 1 and 2. It can be seen that for the 30-day restoring time (their Fig. 1) a numerical artifact is present in form of a $2\Delta x$ wave across the Atlantic near 55°N.

I encountered the same time-independent wave pattern when spinning up the GFDL model in a very similar configuration to that of TTFB. A close-up view of this problem is shown in Fig. 1a. Large freshwater fluxes at single grid points, alternating in sign, extend across the Atlantic from an embayment representing the North Sea and part of the Baltic in the model. The origin of the numerical noise is the very low surface salinity in the Baltic. When Levitus (1982) data for the top 50 m are interpolated to the model grid, salinity values in this open embayment are as low as 7.5 psu at the inside end.

The problem disappears when the salinity to which the model is restored at the two innermost grid points in the embayment is adjusted to a more moderate value (32.5 psu). This adjustment seems an ad hoc fix, but it also makes physical sense. The freshwater outflow from the Baltic into the Atlantic is very small as the Baltic is separated off by narrow, shallow straits. A coarse model can only represent the Baltic as an open embayment, giving Baltic waters an unrealistic influence on the Atlantic. By adjusting the model salinity in the embayment toward open ocean values this influence is reduced. The diagnosed freshwater flux from a spinup using this salinity adjustment is shown in Fig. 1b. In all other respects the model run was identical to the one shown in Fig. 1a; in both cases a salinity restoring time of 30 days was used.

Before we discuss the significance of this adjustment for the stability problem considered by TTFB, a related problem in a second region of the model should be mentioned. Figure 2a shows the diagnosed freshwater flux in the Drake Passage region; here a maximum where net precipitation exceeds 6 m yr^{-1} appears just off South America. This is also clearly visible in Figs. 1 and 2 of TTFB and is due to a strong minimum (<30 psu) in the Levitus near-surface salinity in this area. When the Levitus values are superimposed on a map of the South American coastline, it is revealed that these low salinity values originate from inside the fjords of Tierra del Fuego. Due to the coarse resolution

Corresponding author address: Dr. Stefan Rahmstorf, Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, D-14412 Potsdam, Germany.
E-mail: rahmstorf@pik.potsdam.de

© 1996 American Meteorological Society

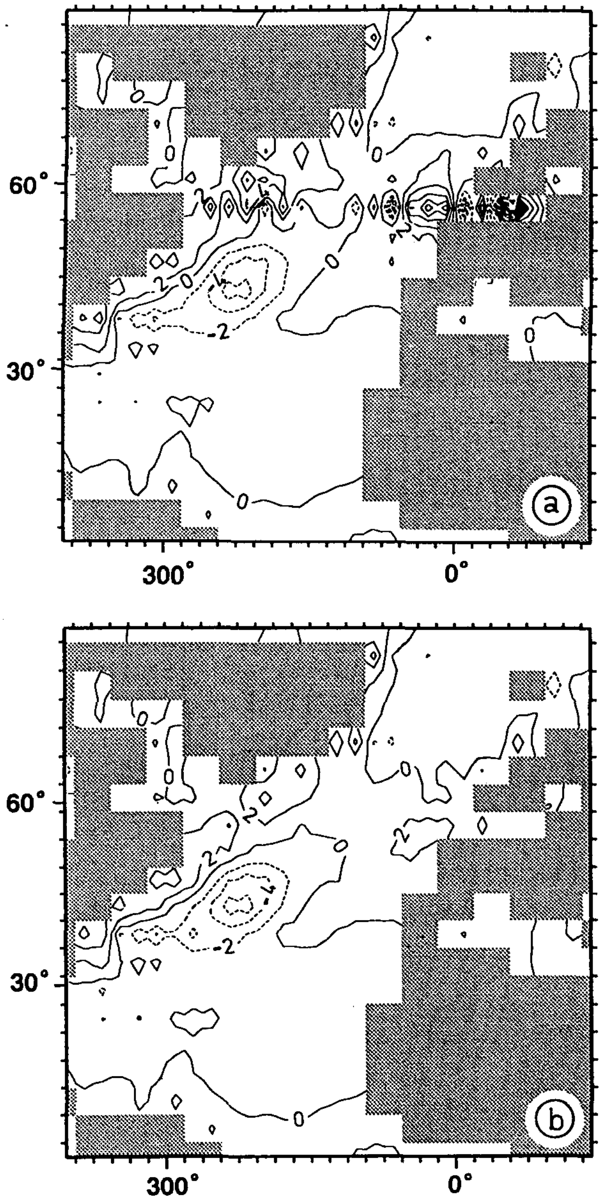

FIG. 1. (a) Freshwater flux (in meters per year) in the North Atlantic diagnosed from a spinup experiment with restoring boundary conditions. Tick marks show the model grid. Note the numerical noise near 55°N. (b) Same as (a) except that the very low salinity values used during restoring at two grid points inside the Baltic were replaced by higher values.

(and the wide Drake Passage) of the present model configuration, the salinity of some ocean grid points is derived mainly from these low values. This is clearly unrealistic and leads to the excess net precipitation visible in Fig. 2a. It can be corrected by setting a lower limit for the restoring salinity in this region; 33.4 psu as recorded in the hydrographic atlas (Olbers 1992) would seem to be realistic. Figure 2b shows the diagnosed freshwater flux from a spinup with this correction; the strong precipitation maximum has disappeared.

We will now turn to the question of whether the numerical wave in the North Atlantic may have affected the experiments of TTFB. In their experiment with a 30-day salinity restoring timescale, the overturning cell in the North Atlantic collapsed after the switch to mixed boundary conditions. The mechanism for such a collapse is the so-called "polar halocline catastrophe," where convection in the northern North Atlantic

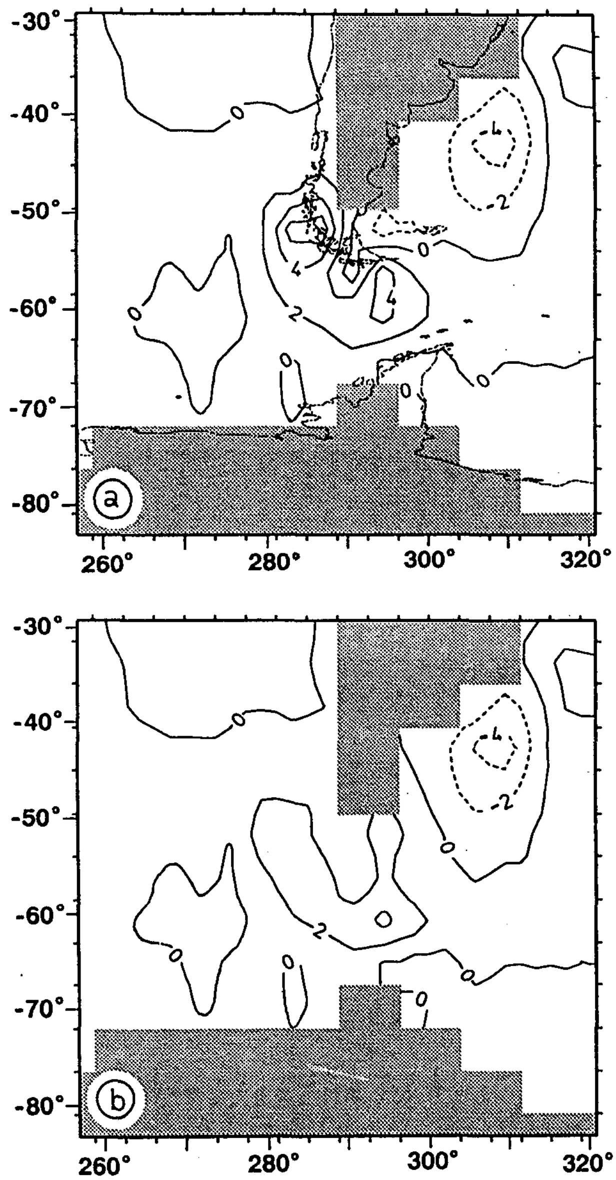

FIG. 2. (a) Freshwater flux (in meters per year) near South America diagnosed from a spinup experiment with restoring boundary conditions. The real coastline is superimposed on the model grid. (b) Same as (a) except that the low salinity values used during restoring near the fjords of Tierra del Fuego were replaced by higher values.

is interrupted and a freshwater lens caps the former deep-water formation region (see section 3). Whether the collapse is triggered thus depends on how easily convection can be interrupted. This can in turn depend not only on the freshwater forcing but also on the details of the convection pattern in the model equilibrium (see Rahmstorf 1995b; Rahmstorf and Willebrand 1995). In the TTFB model, the convection region south of Greenland where North Atlantic Deep Water (NADW) forms is strongly affected by the numerical artifact. Figure 3 shows the convection depth in this region for the two model equilibria shown in Fig. 1, that is, with and without the numerical wave originating from the Baltic. It is clear that the numerical noise in the freshwater forcing generates spurious deep convection in Fig. 1a. This might affect the stability behavior of this model as well as its water mass properties and tracer uptake (e.g., Toggweiler 1989; Toggweiler et al. 1989; England 1993).

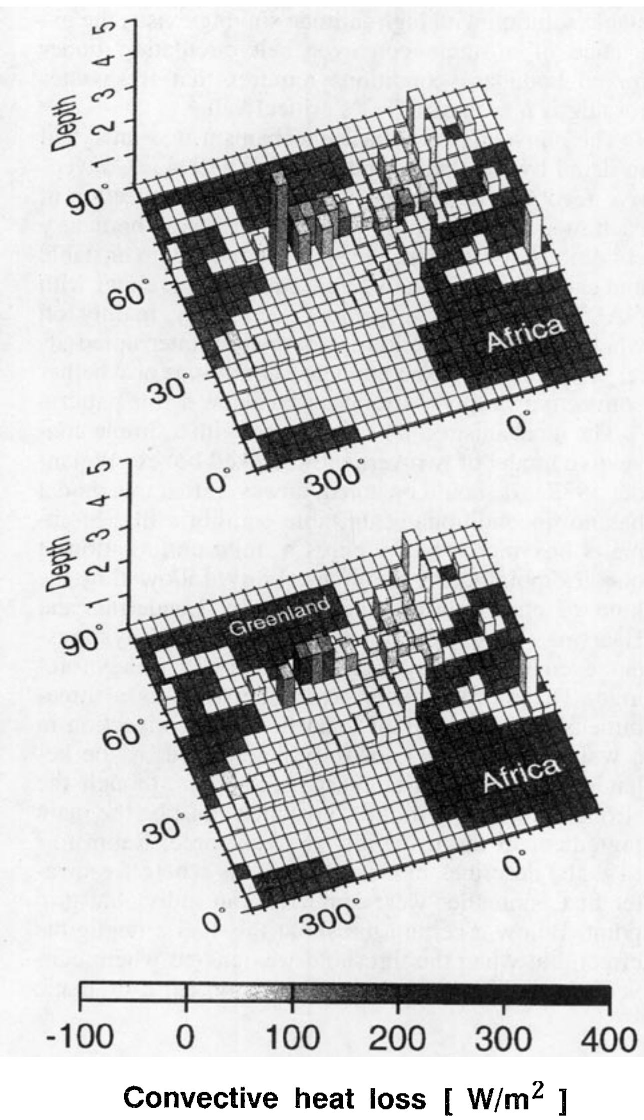

FIG. 3. Convection patterns for the final state of the two spinup experiments shown in Figs. 1a and 1b. Vertical columns show convection depth in kilometers, shading is convective heat loss in watts per square meter. Note the spurious deep convection near 55°N in (a), caused by the numerical noise shown in Fig. 1a.

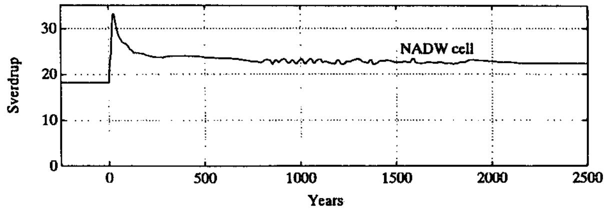

FIG. 4. Time series of NADW overturning in the North Atlantic after switching from restoring to mixed boundary conditions at $t = 0$.

In order to test whether the collapse of the thermohaline circulation discussed by TTFB also occurs when the numerical problem in the North Atlantic is removed, I have repeated their experiment[1] with Baltic salinity corrected as discussed above. The model equilibrium shown in Figs. 1b and 3b, obtained with a 30-day restoring time on both salinity and temperature, was switched to mixed boundary conditions and integrated further for 2500 years. The time series of the NADW overturning cell in the North Atlantic for this experiment is shown in Fig. 4. There is no collapse of the thermohaline circulation; instead it increases from 18 to 22 Sv (Sv $\equiv 10^6$ m^3 s^{-1}) under mixed boundary conditions. This agrees with further experiments that I performed with a salinity restoring time of 50 days, where the thermohaline cell in the Atlantic also increased after switching to mixed boundary conditions (Rahmstorf 1995a). Power and Kleeman (1994) report similar experiments where the spinup was performed with even shorter restoring time (20 days), also resulting in a stable state with 22 Sv overturning after switching to mixed boundary conditions. They also used the GFDL model but with a different global topography that does not include the Baltic. These results, which are contrary to those of TTFB, suggest that the collapse of the thermohaline cell observed by TTFB may be dependent on the numerical $2\Delta x$ wave in their freshwater forcing.

3. Mechanisms of thermohaline state transitions

In the remainder of their paper, TTFB use a box model analysis to explain the varying stability of the thermohaline circulation in their two GCM experi-

[1] Some model details may differ, most importantly diffusivity values (not reported by TTFB). Also I have used my own interpolation of Levitus data to the model grid (based on the mean value of the Levitus grid cells overlapping a model grid box, weighted by the area of overlap) that differs somewhat from the interpolation used at GFDL.

ments. In this section we will discuss the validity of this box model analysis.

The analysis of TTFB is based on an advective box model modified from Stommel's (1961) classical box model. This type of model has either one or three equilibrium solutions for a given freshwater forcing. With strong net precipitation in high latitudes, only "inverse flow" (i.e., deep water forming in low latitudes) is possible; with net evaporation only "normal flow" (deep water forming in high latitudes) is possible; but for weak net precipitation the effects of cooling and freshening compete in high latitudes and allow three equilibria: a strong normal flow, a sluggish normal flow, and an inverse flow. In Stommel's box model, the sluggish normal branch is unconditionally unstable, and it is generally accepted that the strong normal flow is the box model's equivalent of today's overturning circulation in the North Atlantic.

TTFB split the two boxes of Stommel in the vertical to obtain four boxes and then show that in this case part of the "strong, normal" branch is also unstable. The point that separates the stable and unstable parts of this branch is termed "stability transition point" by TTFB; it is a critical value of the net freshwater forcing in high latitudes beyond which no stable solution exists under mixed boundary conditions. TTFB's argument rests on the idea that the two spinups of their GCM produced two different freshwater forcing fields that are on either side of the critical value. The experiment with less net precipitation therefore remained stable, and the one with more net precipitation collapsed after switching to mixed boundary conditions. Because the two forcing fields (obtained with two salinity restoring time scales) do not differ very much in their average net precipitation over the northern North Atlantic, TTFB conclude that both GCM experiments must be close to the "stability transition point," that is, the critical freshwater forcing value.

There is an alternative interpretation of the collapse of the circulation in TTFB's experiment, namely, that it was triggered by the interruption of convection. This convective instability mechanism is often called a "polar halocline catastrophe." This mechanism has not been considered by TTFB and cannot be modeled by Stommel's or TTFB's box model. Through this mechanism it is possible for the circulation to collapse after switching to mixed boundary conditions, even if freshwater forcing does not exceed the critical value identified by TTFB for their box model. An example of this behavior is found in the experiments of Hovine and Fichefet (1994), where a stable conveyor circulation under mixed boundary conditions was perturbed by a step function decrease in high-latitude precipitation. As expected, this intensified the overturning circulation. However, when the perturbation was turned off again after 100 years and the fluxes returned to those that previously had driven the stable conveyor, a polar halocline catastrophe occurred and the circulation collapsed.

The concept of convective instability can also explain the transition to *stronger* overturning as found in the experiment shown in Fig. 4. Experiments both with idealized models (Hughes and Weaver 1994; Rahmstorf 1994; Rahmstorf 1995b) and with realistic global models (Rahmstorf 1995a) show that not only one but many equilibrium states with a conveyor belt circulation exist in three-dimensional models due to the existence of multiple stable convection patterns. Transitions between these are easily triggered by a change in surface feedback (Rahmstorf 1995a). The transition to a new state with a different NADW formation rate cannot be explained by TTFB's instability model. In this model, when forcing is beyond the critical value, *no* stable solution with high-latitude sinking exists; the existence of a stable conveyor belt circulation under mixed boundary conditions requires that freshwater forcing is *less* than TTFB's critical value.

The convective feedback mechanism was analyzed in detail by Lenderink and Haarsma (1994). Convective feedback determines whether deep convection at each model grid point is stable under mixed boundary conditions; it can make the spinup equilibrium unstable and cause a transition to a new equilibrium, either with NADW formation or without, depending mainly on whether deep convection is completely interrupted after switching to mixed boundary conditions or whether convection simply rearranges into a new stable pattern.

The mechanism can be illustrated with a simple convective model of two vertically stacked boxes (Welander 1982). It should be noted, however, that this model has no unconditionally unstable equilibria like Stommel's box model but requires a finite perturbation at one grid point to trigger an instability, followed by the kind of chain reaction described by Lenderink and Haarsma (1994). Such perturbations are always present, even in a fully equilibrated model with steady forcing[2]. The mechanism is best understood as a three-dimensional interplay of local flow and convection in a whole region. It is not fully determined by the behavior of individual convective columns, though the strong nonlinearity in the convection must be the main ingredient in the instability. For example, Rahmstorf (1995b) describes model experiments where freshwater flux anomalies were applied to an individual grid point. Below a certain threshold this had a negligible effect, but when the threshold was passed where convection at this one point was interrupted, a dramatic

[2] One cause of such small perturbations is the leapfrog time-stepping scheme of the GFDL model (Bryan 1969). To prevent the "splitting" of the model into two solutions, every *nmix* time steps a forward step is substituted for the leapfrog step. Time series of convection show that regular convective events are triggered on this forward time step.

state transition was triggered, resembling the one shown in Fig. 4.

4. Why the convective instability mechanism?

There are several reasons why instability of the convection pattern would seem to be the correct explanation for the model transitions that take place after switching to mixed boundary conditions.

The first reason is the timescale of the transition. In our experiments with an idealized circulation model we have seen both types of transition (Rahmstorf et al. 1996): a fast collapse of the conveyor circulation after a brief freshwater pulse that interrupted convection (a "polar halocline catastrophe") and a winding down of the NADW circulation after a permanent increase of precipitation in high latitudes, which did not interrupt convection but made it impossible to sustain a conveyor (this is the mechanism analyzed by TTFB). The former transition typically shows a strong initial response within decades, the latter slowly adjusts over centuries. This indicates that both the collapse within 100 years described in TTFB and the transition shown in Fig. 4 were probably caused by a convective instability; the advective mechanism is too slow to produce such a rapid response. It is interesting that Bryan (1986) already encountered these two timescales for transitions from a symmetrical circulation to a pole to pole circulation. With a negative salinity perturbation, convection was interrupted and a fast collapse (taking 50 years) of the symmetrical state occurred. With a positive salinity perturbation, convection continued, and the symmetrical state decayed on a timescale of hundreds of years. Bryan rightly attributed this difference to the different feedbacks: convective in the fast case, advective in the slow case. TTFB show from the box model that their advective mechanism can have different timescales, but close to the "stability transition point" the timescale of the instability should be very long.

The sequence of events in the transition shown in Fig. 4 is another strong indication of convective instability operating at least in this case. In the time series, changes in convection generally precede changes in overturning. For example, ten years before the overturning peak visible in Fig. 4, convective heat loss in the northern North Atlantic peaked. Figure 5 shows time series of the five years before and after the switch to mixed boundary conditions in this experiment. Figure 5a demonstrates that rearrangements in the North Atlantic convection pattern (and likewise in the Southern Ocean, not shown here) start within the first week after switching to mixed boundary conditions, with some previously convecting points shutting down and some previously stratified points starting to convect. Figure 5b gives a more large-scale view of convection. It shows that about a year after the switch there is a sudden increase in convective heat loss in the northern North Atlantic, which precedes any change in overturning. These results strongly suggest that convective changes trigger adjustments in large-scale flow rather than the other way round.

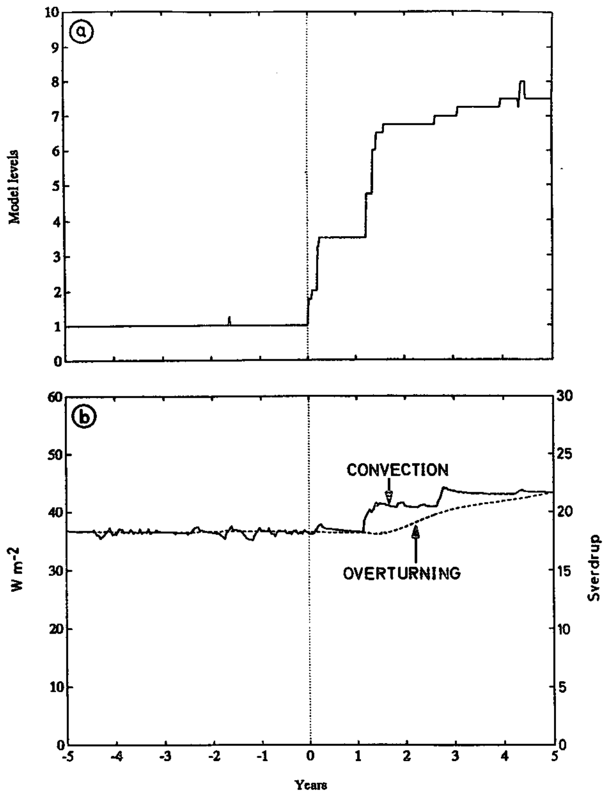

FIG. 5. High-resolution time series of the experiment shown in Fig. 4, showing the five years before and after switching to mixed boundary conditions at $t = 0$. Sampling interval is 7 days. (a) Average convection depth of four grid cells just off the southern tip of Greenland. At one of these points convection already starts within the first week after the switch. At other grid cells (not shown) convection shuts down immediately after the switch. (b) Atlantic overturning (i.e., maximum of the NADW cell in 10^6 m^3 s^{-1}) and convective heat loss averaged over the North Atlantic north of 49°N (in watts per square meter). Note that convection changes precede overturning changes.

An indication of convective instability in the TTFB model is the zonal-average freshwater flux for both their experiments, shown in their Fig. 5. The two curves are very similar except at the latitudes of deep convection. This is a characteristic feature of restoring boundary conditions. When the restoring timescale is decreased, the model is forced closer to the observations; in deep convection areas this requires much larger fluxes than in regions where only the surface level of the model is affected. This means that in high-latitude convection areas large net precipitation will be found in a restoring run with a short restoring timescale. After switching to mixed boundary conditions, this precipitation makes convection unstable, and the model makes a transition to a new convection pattern. It is possible that all deep convection in the North Atlantic is interrupted or that convection merely shifts to grid points with less precipitation. TTFB point out that the isolated peaks in the $E - P$ pattern make little difference to the large-scale freshwater forcing and thus to their proposed large-scale advective instability mechanism; this leads them to conclude that the model must be close to the ''stability transition point.'' However, the doubling of net precipitation near 60°N makes a big difference to the stability of convection in this area and therefore quite naturally explains the difference between their two runs if convective instability is the relevant mechanism.

A final argument is that practically all other published studies with global ocean models under mixed boundary conditions or coupled to an atmosphere model are in a parameter regime where both types of equilibria are possible, with or without conveyor belt circulation (e.g., Manabe and Stouffer 1988; Maier-Reimer et al. 1989; England 1993; Moore and Reason 1993; Power and Kleeman 1993). These models are clearly not beyond TTFB's stability transition point, as beyond this point there is no stable conveyor equilibrium (see TTFB's Fig. 4). I am aware of only one exception: in the coupled scenario calculation of Manabe and Stouffer (1994) for a quadrupling of CO_2, the conveyor circulation slowly winds down, apparently not through a ''halocline catastrophe'' but in a way consistent with TTFB's advective mechanism.

5. Discussion

A major conclusion of TTFB is that their ocean model under mixed boundary conditions is close to a ''stability transition point.'' This conclusion is based on two GCM experiments and on the assumption that the collapse of the conveyor in one of these GCM experiments is due to an advective instability mechanism described by a simple box model. TTFB further conclude, ''this also implies that the ocean itself is near this transition point.''

In this comment I have argued that a collapse or other state transitions of the thermohaline circulation after switching to mixed boundary conditions are probably caused by an instability of *convection*, not by a large-scale *advective* instability. Arguments were presented as to why this convective mechanism (together with the numerical problem in TTFB's spinup run) is likely to have caused the thermohaline collapse in TTFB's model rather than the mechanism investigated by the authors.

If the proposed convective mechanism is responsible for the thermohaline state transitions, then there is no basis for the conclusion that the ocean is close to the stability transition point proposed by TTFB. Rather, the transitions are entirely a result of changing a model from restoring to mixed boundary conditions. They are an artifact of the spinup procedure and tell us nothing about how easily state transitions in the real ocean might be triggered.

Even should the stability analysis of TTFB be applicable, their conclusion about the stability of the real ocean implies considerable confidence in the OGCM. In particular, it requires that the model's freshwater forcing is close to the actual fluxes. More importantly, it tacitly assumes that the strong thermal restoring boundary condition used by TTFB, with a coupling constant of 80 W m^{-2} K^{-1}, allows a realistic stability behavior of the model. Power and Kleeman (1994) and Zhang et al. (1993) have shown that the stability of the circulation depends strongly on the thermal coupling constant (it being more stable for weaker coupling), and Rahmstorf and Willebrand (1995) show from a simple atmospheric energy balance that the coupling sensitivity depends on the scale of the process considered and should be 9–12 W m^{-2} K^{-1} for large-scale changes in the thermohaline circulation. It is therefore likely that the model of TTFB is too unstable.

The question of how close the North Atlantic is to an instability is of great importance for environmental policy because of the concern as to whether man-made climate changes could cause a transition or even a collapse of the thermohaline conveyor belt. It remains to be established how much it takes to trigger a ''polar halocline catastrophe'' or a more subtle shift in convection sites, which could also have major climatic repercussions (Rahmstorf 1994). Global ocean models are not yet capable of representing NADW formation in the Greenland Sea and its subsequent overflow over the Iceland Sill in a realistic way, and coupled models have the additional problem of large flux corrections in this region. Therefore, present climate models are in my view not likely to give a reliable estimate of the sensitivity of North Atlantic deep circulation.

Acknowledgments. I am grateful to Eli Tziperman, Robbie Toggweiler, Yizhak Feliks, Kirk Bryan, Stephen Griffies, and Bonnie Samuels for constructive criticism of an earlier draft of this comment. Matthew England shared his version of the global GFDL code with me, and Dulcie Smart made many suggestions to clarify the language. The calculations were performed at the German Climate Computer Centre in Hamburg.

REFERENCES

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

Bryan, K., 1969: A numerical method for the study of the circulation of the world ocean. *J. Comput. Phys.,* **4,** 347–376.

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

Hovine, S., and T. Fichefet, 1994: A zonally averaged, three-basin ocean circulation model for climate studies. *Climate Dyn.,* **10,** 313–331.

Hughes, T. M. C., and A. J. Weaver, 1994: Multiple equilibria of an asymmetric two-basin ocean model. *J. Phys. Oceanogr.,* **24,** 619–637.

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

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

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

——, and R. J. Stouffer, 1988: Two stable equilibria of a coupled ocean–atmosphere model. *J. Climate,* **1,** 841–866.

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

Olbers, D., 1992: *Hydrographic Atlas of the Southern Ocean.* Alfred Wegener Institut, 82 pp.

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

——, and R. Kleeman, 1993: Multiple equilibria in a global ocean general circulation model. *J. Phys. Oceanogr.,* **23,** 1670–1681.

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

——, 1995a: Climate drift in an OGCM coupled to a simple, perfectly matched atmosphere. *Climate Dyn.,* **11,** 447–458.

——, 1995b: Multiple convection patterns and thermohaline flow in an idealized OGCM. *J. Climate,* **8,** 3028–3039.

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

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

Toggweiler, J. R., 1989: Simulations of radiocarbon in a coarse-resolution world ocean model. I. Steady state pre-bomb distributions. *J. Geophys. Res.,* **94,** 8217–8242.

——, K. Dixon, and K. Bryan, 1989: Simulations of radiocarbon in a coarse-resolution world ocean model. 2. Distributions of bomb-produced carbon 14. *J. Geophys. Res.,* **94,** 8243–8264.

Tziperman, E., J. R. Toggweiler, Y. Feliks, and K. Bryan, 1994: Instability of the thermohaline circulation with respect to mixed boundary conditions: is it really a problem for realistic models? *J. Phys. Oceanogr.,* **24,** 217–232.

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

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.