Inherently unstable climate behaviour due to weak thermohaline ocean circulation

Eli Tziperman

Department of Environmental Sciences, Weizmann Institute of Science, Rehovot, 76100, Israel and Atmospheric and Oceanic Sciences Program, Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey 08544, USA

The oceanic thermohaline circulation (THC) carries light, warm surface water polewards and dense, cold deep water equatorwards, thereby transporting a large amount of heat towards the poles and significantly affecting high latitude climate. The THC has been remarkably stable, and its variability quite low, over the Holocene period (the past 10,000 years). The much greater climate instability and high-frequency variability recorded in ice[1] and deep-sea[31] cores throughout the preceding 150,000 years has been linked to greater THC variability[2,3]. Here we argue, using a global coupled ocean–atmosphere–ice general circulation model with realistic geography, that there is a wide range of weak mean states of the THC that cannot be stably sustained by the climate system. When the model THC is forced into a state in the unstable range, the THC may rapidly strengthen, collapse or display strong oscillations. The existence of this unstable regime may account for the greater variability of the THC and climate before the Holocene period.

The THC is driven by the north–south gradient in water density created by atmospheric thermal forcing and weakened by the north–south salinity gradient maintained by air–sea exchange of fresh water (FW) due to subtropical evaporation and high-latitude precipitation. The present-day stable THC is thermally dominant, that is, driven by cooling and sinking of dense water masses at high latitudes. Numerous ocean-only model studies, from simple box models[4] to three-dimensional models[5], have shown that a weaker THC than that of today, obtained under larger freshwater forcing (that is, a less thermally dominant THC) may become unstable. Once unstable, such a weak THC cannot be sustained with small amplitude variability such as exists today. Instead, the THC may reverse[4,6–8] (as it may have done during the Cretaceous period[9]), undergo a significant decrease[10], collapse[10,11], display strong oscillations[12–14] or increase to a stable value outside the unstable range[10]. On the basis of such ocean-only models, it was further speculated[6,15] that a weaker and less thermally dominant North Atlantic THC than today's, obtained under stronger FW forcing (for example, a larger FW flux into the northern North Atlantic Ocean), might be unstable.

Ocean-only models heretofore used for investigating the stability of the THC are known to be prone to various artefacts[8,16,17] which can affect the model's stability behaviour. Here I use a coupled ocean–atmosphere–ice general circulation model (GCM), in which the air–sea interaction is more fully represented, to examine the spontaneous stability of a range of strong (thermally dominant) and weak (less thermally dominant) THC states. We consider a given climate state to be stable if, when used to initialize the coupled model, it results in a small-amplitude variability around this initial state, similar to the weak THC variability during the Holocene. On the other hand, a climate state is considered unstable if, when used to initialize the coupled model, it results in a sizeable drift of the THC away from the initial state, or in large-amplitude THC oscillations. A THC instability may thus result in an increase as

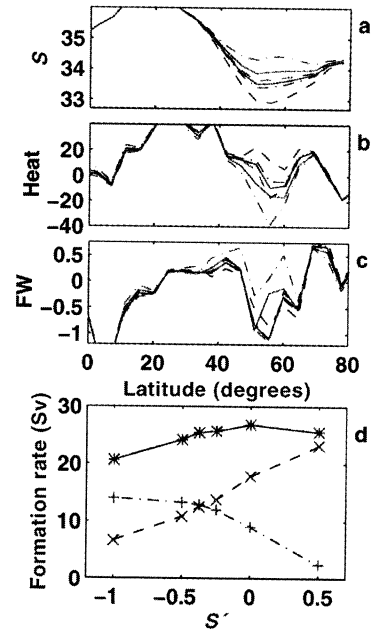

Figure 1 Diagnostics of the steady-state, ocean-only solutions used to initialize the coupled model integrations shown in Fig. 2. **a**, The zonally averaged North Atlantic sea surface salinity fields used to obtain the various steady-state, ocean-only solutions. To the observed North Atlantic salinity (solid grid line), zonally uniform perturbations of the form $\Delta S \exp(-(\theta - 55.5°)^2/4.4^{°2})$ were added, as a function of latitude θ only, where ΔS takes the values (-1, -0.5, -0.375, -0.25, 0, $+0.5$ p.p.t.). **b**, the heat flux and **c**, the freshwater flux adjustments (in watts m^{-2} and $m\,yr^{-1}$, respectively), annually averaged and zonally averaged over the Northern Atlantic ocean, for the control run and for the five modified surface salinity runs. A positive FW adjustment tends to make the ocean saltier, and a positive heat adjustment heats the ocean. Line types for each experiment are as in **a**. **d**, The steady-state, ocean-only rate of formation of North Atlantic Deep Water (dashed line), Southern Ocean Deep Water (dash-dotted line) and their sum (solid line), as function of the restoring salinity perturbation, $S' = \Delta S$, in the North Atlantic.

well as a decrease, a collapse or strong oscillations of the THC[10] which resemble some of the strong climate variability seen during unstable past climates[1].

To explore the stability of weak and strong (less and more thermally dominant) THC states, I initialized several coupled ocean–atmosphere model runs from various steady-state, ocean-only model solutions, following the standard initialization procedure for such coupled models[18]. The different ocean-only steady-state solutions were obtained by forcing (restoring) the upper level temperature and salinity of the ocean model to near an observed monthly sea surface temperature (SST) and a modified observed monthly sea surface salinity (SSS)[12] (termed 'restoring SST and SSS'). The modifications to the SSS to which the surface model salinity is restored makes it fresher or saltier in the North Atlantic area of water mass formation (Fig. 1a), leading to a correspondingly weaker (less thermally dominant) or stronger (more thermally dominant) steady-state ocean-only North Atlantic THC (dashed line in Fig. 1d). The restoring of the model surface temperature and salinity to the specified SST and SSS fields allows various ocean-only steady states to be obtained, some of which may be unstable in the coupled model, where the surface fields are not restored and are free to evolve.

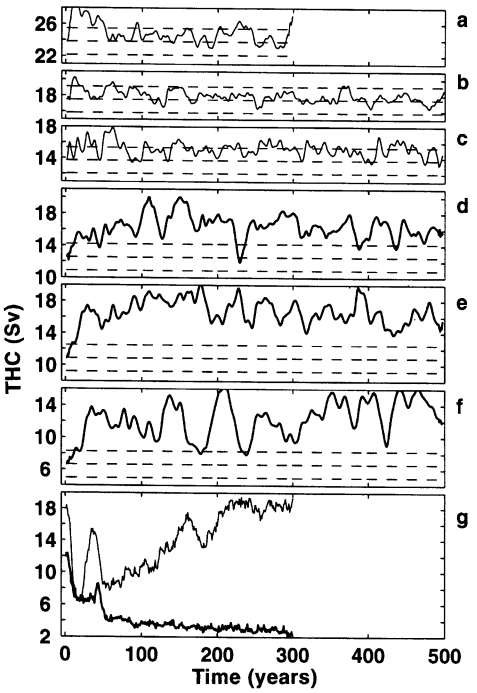

Figure 2 The North Atlantic yearly averaged THC (in Sverdrup) plotted against time (years) from eight coupled model runs. Dashed lines denote the initial state and plus/minus two standard deviation (SD) lines, where the SD is defined from the control run (**b**). Stable initial states are defined as those whose THC remains within two SD during the coupled model run, although the results are not sensitive to this particular choice of two SD. The different panels show **a**, the coupled model run initialized with the stable state [$\Delta S = +0.5$]; **b**, the stable control run[19,20]; **c**, run initialized with the stable state [$\Delta S = -0.25$]; **d**, run initialized with the unstable state [$\Delta S = -0.375$]; **e**, run initialized with the unstable state [$\Delta S = -0.5$]; **f**, run initialized with the unstable state [$\Delta S = -1$]; **g**, thick line: the North Atlantic THC for a coupled model run starting from year 10 of the run shown in **e**, initialized with the unstable initial state [$\Delta S = -0.5$]. In this run, freshwater flux is added to the North Atlantic sinking area for the first 10 years. The THC collapses and does not recover for 300 years. This contrasts with the corresponding experiment of ref. 23, started from the stable control run and shown by the thin line.

The low-resolution primitive-equation coupled ocean–atmosphere–ice GCM used here, with its realistic geometry, has been used for many studies of greenhouse warming[19], climate variability[20] and multiple climate equilibria[21]. When initialized with ocean-only and atmosphere-only steady states obtained by running the two models separately, the coupled model rapidly drifts from its initial conditions. The air–sea fluxes of heat and fresh water are therefore supplemented at every time step by flux adjustments (FA) which depend on geographical location and month, but have no inter-annual variations and are obtained by averaging the restoring terms from the ocean-only initialization run[18]. FA is used in the present experiments as a parameter that allows the different initial THC states to be obtained (Fig. 1b, c). This is somewhat similar to changing the FW forcing in an ocean-only model[12,14,15].

In the absence of an external perturbation, if the initial THC is stable, we expect the coupled-model initialization procedure to prevent a drift of the THC in the coupled model from the initial ocean-only steady states, as it did in the stable control run[19,20] (although see ref. 22 for a caveat). The strength of the North Atlantic THC as a function of time for the five coupled-model runs initialized using modified surface salinities, and for the control run, is shown in Fig. 2a–f. I denote each initial state of the coupled model by the amplitude ΔS of the restoring salinity perturbation used to obtain it in the ocean-only model runs (Fig. 1a). When ocean-only models in previous studies were initialized by restoring their SST and SSS and were then switched to some simple atmospheric representation, these models invariably showed that weak THC initial states were unstable, leading to a switch to a different THC state. The qualitative behaviour of the realistic-geometry coupled ocean–atmosphere GCM used here is remarkably consistent with that of these simpler ocean-only models. In particular, the more thermally dominant initial states with a strong initial THC (runs [$\Delta S = -0.25$], control [$\Delta S = 0$] and [$\Delta S = +0.5$], Fig 2a–c) are stable. The coupled model runs starting at these initial conditions remain mostly within two standard deviations from the initial state, where the standard deviation is defined from the control run's THC variability which roughly represents a small, Holocene-like, stable variability. On the other hand, the less thermally dominant and thus weaker initial THC states (runs [$\Delta S = -1$], [$\Delta S = -0.5$], [$\Delta S = -0.375$], Fig. 2d–f) are spontaneously unstable within the coupled model, as the THC drift from the initial conditions in these runs clearly exceeds two standard deviations. I call this a spontaneous instability, because no external perturbation is applied to the coupled model once the coupled model integration begins.

The threshold between stable and spontaneously unstable initial states in the coupled model lies between 14 Sv (stable state [$\Delta S = -0.25$], Fig. 2c) and 13 Sv (unstable state [$\Delta S = -0.375$], Fig. 2d). The steady-state ocean-only solution used as the stable initial condition in the control run [$\Delta S = 0$] is different from the initial state [$\Delta S = -0.375$] that is unstable within the coupled model, yet not dissimilar (Fig. 3a–d), indicating that an ocean climate not very far from the present-day climate may be spontaneously unstable within the coupled model. Note that none of the runs displays a stable small-amplitude THC variability, like that of the present day, within the unstable THC regime of below 14 Sv. When forced by a doubling of carbon dioxide[19] and by freshwater input into the North Atlantic ocean[23], the coupled model THC recovers within 150–200 years. Its inability to recover to its original weak initial states within the fairly long integration of 500 years seen in Fig. 2 indicates that the weak THC initial states may indeed be considered to be in an unstable THC regime.

The instability mechanism leading to the drift away from the unstable initial states in Fig. 2d–f is composed of two stages. First, an advective linear salinity feedback[6,7,15] leads to an increase of salinity and density in the water mass formation area and thus an increase in the formation rate of DW southeast of Greenland. The

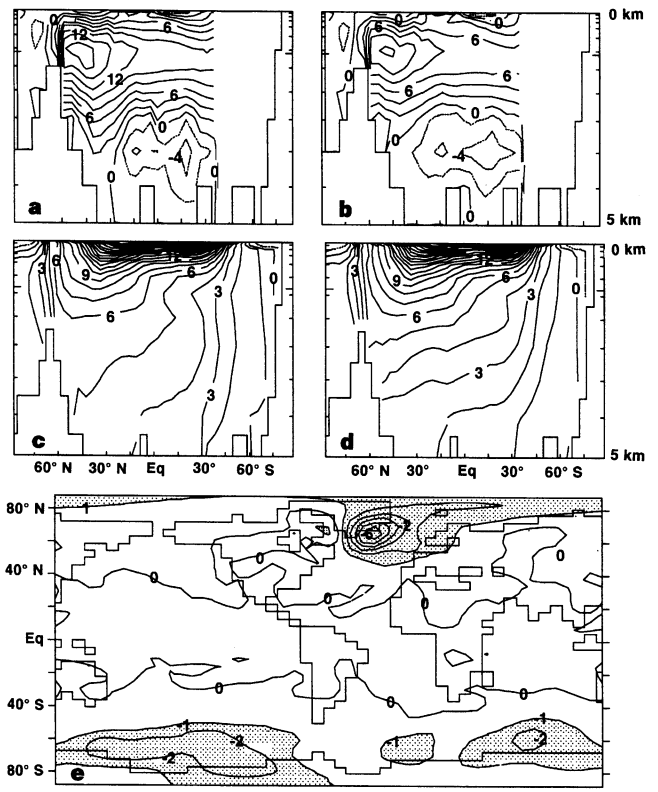

Figure 3 a–d, Upper panel, the steady-state, ocean-only North Atlantic meridional circulation (that is, the meridional stream function); middle panel, the zonally averaged temperature. **a**, **c**, Stable initial state for the control run; **b**, **d**, unstable initial state for run [$\Delta S = -0.375$]. The smaller rate of deep-water formation in the North Atlantic in the unstable solution allows an increased influx of Southern Ocean water at depth, and thus the difference seen in the deep-water properties and in Deep Water formation rate in the Southern Ocean (Fig. 1d). **e**, The difference between the atmospheric surface temperature averaged over years 201–300 for the collapsed THC run of Fig. 2g, and the similarly averaged atmospheric surface temperature for the control run.

runs with the stronger initial THC are stable with respect to this linear instability stage, because the strong THC rapidly advects away any developing salinity perturbations, and because such perturbations are harder to develop because of the reduced meridional salinity gradient in the stable states[6,7,15]. At a second stage, the increased sinking southeast of Greenland modifies the properties of the deep water inside the neighbouring Labrador Sea, activating DW formation and self-sustained convective mixing[24] there as well, and leading to a further increase in the THC strength. Such an increase is not seen in simpler two-dimensional models[4,7,15] which, once unstable, often lead to a collapsed or reversed THC. These models, being two-dimensional, cannot support an inactive DW formation site at the same latitude as the active formation sites.

Past the initial instability, the runs beginning from the unstable initial states are characterized by strong North Atlantic THC oscillations (Fig. 2d–f). This is the first time that such large-amplitude oscillations have been seen in a coupled model, following numerous ocean-only model studies that observed such large-amplitude THC variability[5], and coupled model runs that displayed a smaller variability[20,25]. These large-amplitude oscillations (which, especially in Fig. 2f, seem to result from the inherent instability of the weak THC states) do not involve a complete switch between 'on' and 'off' states of the North Atlantic THC, yet may be able to account for the climate oscillations in the interglacial Eemian period[1,26].

In a final experiment (Fig. 2g), I added a freshwater flux of 1 Sv to the northern North Atlantic over 10 years, and found that the spontaneously unstable initial states were also more prone to a forced THC collapse than stronger THC runs such as the control run[23].

Various palaeo proxy data have shown a link between the North Atlantic THC and the Pacific sector of the Southern Ocean[2]. The present ocean-only steady-state solutions indeed show that as DW formation rates in the Pacific sector of the Southern Ocean and in the North Atlantic Ocean vary between the different initial states by 10–15 Sv at each DW source, their sum remains nearly constant (Fig. 1d). Thus this sum, the total DW formation, is determined through a global teleconnection (Fig. 3). Similarly, the forced instability experiment (Fig. 2g) displays a cooling of the air temperature both over Greenland and in the Pacific sector of the Southern Ocean by about 5 and 2 °C respectively (Fig. 3e; see also ref. 23).

The unknown effects of the flux adjustment procedure introduce an uncertainty as to the applicability of the more quantitative results (the precise location of the instability threshold) to the actual climate system[27,28]. I also note that the phenomenon of a coupling shock[22], at the initialization of coupled model integrations, may lead to an initial model drift which may affect the differentiation of stable and unstable states. However, some previous coupled model results[29], and in particular the remarkable qualitative agreement with the numerous ocean-only studies of THC stability lend additional confidence to the results here. In fact, I feel that a flux-adjusted coupled model with an active atmospheric component may be expected to be more accurate than ocean-only models with fixed freshwater flux and atmospheric temperature, with a highly simplified heat flux parameterization and with no weather noise.

In summary, this study provides confirmation using a fully coupled ocean–atmosphere–ice general circulation model with realistic geometry that a weak, less thermally dominant THC is unstable. The results limit the range of stable equilibria of the THC in this model to above about 14 Sv and below 4 Sv (our collapsed THC run, Fig. 2g). In this model I did not find intermediate values of the THC that are stable, with small variability like that of the present day. This may suggest that the North Atlantic water mass formation and thus THC may have operated in a weak and shallow[3] (Fig. 3b) unstable mode with large variability before the Holocene and during previous interglacial periods[1]. Later, the water mass formation may have switched to a stable mode with small variability, and with stronger and deeper North Atlantic THC, which has lasted throughout the Holocene.

To set this in perspective, I note that the weakening required to make the model THC spontaneously unstable is still significantly larger than the present-day natural variability, which is estimated to be about 5–10% for the coupled model used here[20]. Although the location of the stable THC range found here differs from those found in simpler models[15,21] and is clearly model dependent, I expect the qualitative dependence of the model stability and variability behaviour on the mean THC strength to carry over to the real world. The implications for greenhouse scenarios that predict a weakening of the THC are clearly an important yet open issue. Finally, improved knowledge of past THC magnitude[30], and a more precise determination of the unstable THC regime and its distance from present-day climate, are clearly needed, using better coupled ocean–atmosphere models as they become available. □

Received 10 October 1996; accepted 7 February 1997.

1. GRIP Members Climate instability during the last interglacial period recorded in the GRIP ice core. *Nature* **364**, 203–207 (1993).
2. Broecker, W. S., Peteet, D. M. & Rind, D. Does the ocean–atmosphere system have more than one stable mode of operation? *Nature* **315**, 21–25 (1985).
3. Boyle, E. A. & Keigwin, L. North Atlantic thermohaline circulation during the past 20,000 years linked to high-latitude surface temperature. *Nature* **330**, 35–40 (1987).
4. Stommel, H. M. Thermohaline convection with two stable regimes of flow. *Tellus* **XIII**, 224–230 (1961).

5. McWilliams, J. C. Modelling the oceanic general circulation. *Annu. Rev. Fluid. Mech.* **28,** 215–248 (1996).
6. Walin, G. The thermohaline circulation and the control of ice ages. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **50,** 323–332 (1985).
7. Marotzke, J., Welander, P. & Willebrand, J. Instability and multiple steady states in a meridional-plane model of the thermohaline circulation. *Tellus A* **40,** 162–172 (1988).
8. Mikolajewicz, U. & Maier-Reimer, E. Mixed boundary conditions in ocean general circulation models and their influence on the stability of the model's conveyor belt. *J. Geophys. Res.* **99,** 22633–22644 (1994).
9. Brass, G. W., Southam, J. R. & Peterson, W. H. Warm saline bottom water in the ancient ocean. *Nature* **296,** 620–623 (1982).
10. Toggweiler, J. R. *et al.* Reply. *J. Phys. Oceanogr.* **26,** 1106–1110 (1996).
11. Bryan, F. High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature* **323,** 301–304 (1986).
12. Weaver, A. J., Sarachick, E. S. & Marotzke, J. Freshwater flux forcing of decadal and interdecadal oceanic variability. *Nature* **353,** 836–838 (1991).
13. Chen, F. & Ghil, M. Interdecadal variability of the thermohaline circulation and high-latitude surface fluxes. *J. Phys. Oceanogr.* **25,** 2547–2568 (1995).
14. Rahmstorf, S. Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nature* **378,** 145–149 (1995).
15. Tziperman, E., Toggweiler, R., Feliks, Y. & Bryan, K. 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 (1994).
16. Zhang, S., Greatbatch, R. J. & Lin, C. A. A reexamination of the polar halocline catastrophe and implications for coupled ocean–atmosphere modeling. *J. Phys. Oceanogr.* **23,** 287–299 (1993).
17. Rahmstorf, S. & Willebrand, J. The role of temperature feedback in stabilizing the thermohaline circulation. *J. Phys. Oceanogr.* **25,** 787–805 (1995).
18. Manabe, S., Spelman, M. J. & Bryan, K. Transient response of a coupled ocean–atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. *J. Clim.* **4,** 785–818 (1991).
19. Manabe, S. & Stouffer, R. J. Century-scale effects of increased atmospheric CO_2 on the ocean–atmosphere system. *Nature* **364,** 215–218 (1993).
20. Delworth, T., Manabe, S. & Stouffer, R. J. Interdecadal variations of the thermohaline circulation in a coupled ocean–atmosphere model. *J. Clim.* **6,** 1993–2011 (1993).
21. Manabe, S. & Stouffer, R. J. Two stable equilibria of a coupled ocean–atmosphere model. *J. Clim.* **1,** 841–866 (1988).
22. Rahmstorf, S. Climate drift in an ocean model coupled to a simple, perfectly matched atmosphere. *Clim. Dynam.* **11,** 447–458 (1996).
23. Manabe, S. & Stouffer, R. J. Simulation of abrupt climate change induced by freshwater input to the North Atlantic ocean. *Nature* **378,** 165–167 (1995).
24. Lenderink, G. & Haarsma, R. J. Variability and multiple equilibria of the thermohaline circulation associated with deep water formation. *J. Phys. Oceanogr.* **24,** 1480–1493 (1994).
25. Griffies, S. & Tziperman, E. A linear thermohaline oscillator driven by stochastic atmospheric forcing. *J. Clim.* **8,** 2440–2453 (1995).
26. Weaver, A. J. & Hughes, T. M. C. Rapid interglacial climate fluctuations driven by North Atlantic ocean circulation. *Nature* **367,** 447–450 (1994).
27. Neelin, J. D. & Dijkstra, H. A. Ocean–atmosphere interaction and the tropical climatology. Part I: the dangers of flux adjustment. *J. Clim.* **8,** 1325–1342 (1995).
28. Marotzke, J. & Stone, P. H. Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. *J. Phys. Oceanogr.* **25,** 1350–1364 (1995).
29. Schiller, A., Mikolajewicz, U. & Voss, R. The stability of the thermohaline circulation in a coupled ocean–atmosphere general circulation model. *Clim. Dynam.* (in the press).
30. Yu, E.-F., Francois, R. & Bacon, M. P. Similar rates of modern and last-glacial ocean thermohaline circulation inferred from radio-chemical data. *Nature* **379,** 689–694 (1996).
31. McManus, J. F. *et al.* High resolution climate records from the North Atlantic during the last interglacial. *Nature* **371,** 326–329 (1995).

Acknowledgements. I thank S. Manabe and R. Stouffer for allowing the use of their coupled model, and S. Griffies and R. Stouffer without whose help the completion of this work would not have been possible. This work is partially funded by the Israeli Academy of Sciences.

Correspondence should be addressed to E.T. at the Weizmann Institute of Science (e-mail: eli@beach.weizmann.ac.il).