

Effect of Atmospheric Feedbacks on the Stability of the Atlantic Meridional Overturning Circulation

MATTHIJS DEN TOOM AND HENK A. DIJKSTRA

Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands

ANDREA A. CIMATORIBUS AND SYBREN S. DRIJFHOUT

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

(Manuscript received 22 August 2011, in final form 5 December 2011)

ABSTRACT

The impact of atmospheric feedbacks on the multiple equilibria (ME) regime of the Atlantic meridional overturning circulation (MOC) is investigated using a fully implicit hybrid coupled model (HCM). The HCM consists of a global ocean model coupled to an empirical atmosphere model that is based on linear regressions of the heat, net evaporative, and momentum fluxes generated by a fully coupled climate model onto local as well as Northern Hemisphere averaged sea surface temperatures. Using numerical continuation techniques, bifurcation diagrams are constructed for the HCM with the strength of an anomalous freshwater flux as the bifurcation parameter, which allows for an efficient first-order estimation of the effect of interactive surface fluxes on the MOC stability. The different components of the atmospheric fluxes are first considered individually and then combined. Heat feedbacks act to destabilize the present-day state of the MOC and to stabilize the collapsed state, thus leaving the size of the ME regime almost unaffected. In contrast, interactive freshwater fluxes cause a destabilization of both the present-day and collapsed states of the MOC. Wind feedbacks are found to have a minor impact. The joint effect of the three interactive fluxes is to narrow the range of ME. The shift of the saddle-node bifurcation that terminates the present-day state of the ocean is further investigated by adjoint sensitivity analysis of the overturning rate to surface fluxes. It is found that heat feedbacks primarily affect the MOC stability when they change the heat fluxes over the North Atlantic subpolar gyre, whereas interactive freshwater fluxes have an effect everywhere in the Atlantic basin.

1. Introduction

The circulation of the North Atlantic Ocean is thought to be a particularly sensitive component of the climate system. According to a long-standing hypothesis, the Atlantic meridional overturning circulation (MOC), which controls the meridional heat transport, is vulnerable to freshwater anomalies because of the so-called salt-advection feedback (Stommel 1961; Bryan 1986). This feedback arises because of the fact that the distribution of salinity is affected by the advection of salt, while at the same time determining the strength and

pattern of the circulation through its effect on the density. Although there is a similar nonlinear coupling between the temperature and the circulation, it is far less effective because changes in sea surface temperature (SST) are strongly restricted by variable surface heat fluxes. As a result of the salt-advection feedback there may be another stable equilibrium of the MOC, additional to the present-day temperature driven overturning, in which a self-sustained salinity distribution drives a reversed (and usually much weaker) circulation. Model results show that a collapse of the present-day MOC would strongly affect the Northern Hemisphere climate (Vellinga and Wood 2002) as a result of the reduction of the meridional heat transport.

In a hierarchy of ocean-only models, such as box models (Thual and McWilliams 1992), two- (Cessi and Young 1992) and three-dimensional models (Weijer and Dijkstra 2001), and Earth System Models of Intermediate

Corresponding author address: Matthijs den Toom, Institute for Marine and Atmospheric Research Utrecht (IMAU), Dept. of Physics and Astronomy, Utrecht University, Princetonplein 5, 3584 CC Utrecht, Netherlands.
E-mail: m.dentoom@uu.nl

Complexity (EMICs; Rahmstorf et al. 2005), multiple equilibria of the MOC exist. In these models a relatively rapid (on a decadal-to-century time scale) transition from one equilibrium state to the other can be induced by changing the freshwater boundary conditions or by perturbing the salinity field. MOC-induced abrupt transitions between different equilibria are also an important component of current theories regarding the Dansgaard–Oeschger oscillations that occurred during the last glacial period (Clark et al. 2002; Ganopolski and Rahmstorf 2001).

In state-of-the-art global climate models (GCMs), such as those used in the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC; Meehl et al. 2007), however, multiple equilibria have not been found. Computational constraints currently prevent performing the extensive parameter sweep experiments that are required to detect a bistable regime of the MOC in these models. The fact that no permanent collapse of the MOC is observed in the Special Report on Emissions Scenarios (SRES) A1B scenario runs, as discussed by Schmittner et al. (2005), or in experiments in which the high-latitude freshwater flux is temporarily perturbed, as analyzed by Stouffer et al. (2006), does not necessarily provide conclusive evidence about the existence of multiple stable equilibria in state-of-the-art GCMs. Different conclusions have been reached based on coupled ocean–atmosphere models that use a coarser resolution than the state-of-the-art models. The results of Yin et al. (2006) suggest that the presence of a dynamically active atmosphere causes the bistable regime of the MOC to disappear. In that case the presence of multiple equilibria should be considered as an artifact of ocean-only models and EMICs that arises because of the poor representation of the interaction between atmosphere and ocean. In contrast, a recent extensive investigation by Hawkins et al. (2011) demonstrated the presence of multiple equilibria in Fast Met Office/U.K. Universities Simulator (FAMOUS), which is a low-resolution version of a fully coupled GCM. Although differences in spatial resolution and in the representation of unresolved processes are also likely to affect the MOC stability in different models, this study will focus on the effect of atmospheric feedbacks.

A reduction of the MOC due to freshwater anomalies leads to changes in sea surface temperatures, which, in turn, induce a change in wind fields, near-surface air temperatures, and evaporation and precipitation patterns. These atmospheric processes constitute changes in the ocean forcing and hence may feed back on the response of the MOC to freshwater anomalies. The southward shift of the intertropical convergence zone (ITCZ) in response to a reduction of the MOC is a well-documented

example of how the atmosphere interacts with the ocean, both as a mechanism for decadal variability (Vellinga and Wu 2004) and as a process that destabilizes the collapsed state of the MOC (Vellinga et al. 2002; Yin et al. 2006).

The main question of this paper is, how do atmospheric processes affect the bistable regime of the Atlantic MOC? We approach this problem by starting from the ocean-only case for which the stability properties have been systematically analyzed over the last decade. It is known that multiple equilibria in an ocean-climate model are connected to the existence of two saddle-node bifurcations (Dijkstra 2005). The distance between the two saddle-node bifurcations in terms of the magnitude of a freshwater perturbation determines the width of the hysteresis that is found in quasi-equilibrium simulations (Rahmstorf et al. 2005). Using techniques from numerical bifurcation theory, the saddle-node bifurcations have been determined for a global ocean-only model (Dijkstra and Weijer 2005) and an indicator was suggested to monitor the presence of the multiple equilibria regime (Huisman et al. 2010). In this study, we will use a fully implicit Hybrid Coupled Model (HCM) and techniques of numerical bifurcation theory to determine the changes (and the physical processes causing these changes) in the position of the saddle-node bifurcations when atmospheric feedbacks are represented. The HCM consists of a global ocean general circulation model coupled to an empirical atmosphere model. The latter is based on linear regressions of the momentum, heat, and freshwater fluxes generated by a fully coupled climate model onto local as well as Northern Hemisphere averaged sea surface temperatures (Cimatoribus et al. 2012).

The outline of this paper is as follows: the model and methodology used are described in section 2. Section 3 presents the results on the changes in the bifurcation diagrams, of which the physics is further analyzed in section 4. A summary and discussion follows in section 5.

2. Model

The behavior of the coupled ocean–atmosphere system is investigated using numerical steady-state continuation and adjoint sensitivity analysis. These methods require that our model admits true steady-state solutions. So, it must exhibit less rich dynamics than state-of-the-art GCMs, but it should include the relevant long-term-mean ocean–atmosphere interaction. This is achieved by adopting a hybrid approach in which we combine a fully implicit, coarse-resolution ocean model with an empirical description of the fluxes at the ocean–atmosphere interface.

The core of the hybrid model is the Thermohaline Circulation Model (THCM; Weijer et al. 2003), which is a fully implicit ocean model designed specifically to perform numerical bifurcation analysis. The THCM is based on the rigid-lid primitive equations in spherical coordinates, formulated in longitude–latitude–depth coordinates on an Arakawa C-grid. The model domain covers a near-global region extending from 85.5°S to 85.5°N, and incorporates realistic bottom topography. The zonal (meridional) grid spacing is 3.75° (4.5°) and there are 12 levels, ranging in thickness from 50 m for the top layer to 950 m for the bottom layer. The implicit approach can be used effectively only when the eddy viscosity is relatively high (Wubs et al. 2006) and when no convective adjustment scheme is applied (Vellinga 1998; den Toom et al. 2011). Despite these limitations, many of the characteristics of the large-scale circulation, and, specifically, its equilibrium sensitivity to an anomalous freshwater forcing, are consistent between THCM and traditional ocean models with the same resolution (Weijer et al. 2003; Dijkstra and Weijer 2005).

The new element presented in this paper is the treatment of the surface forcing. In addition to boundary conditions corresponding to fixed atmospheric conditions, we use the empirical model developed by Cimatoribus et al. (2012) to calculate momentum, heat, and freshwater fluxes as functions of changes in sea surface temperature. We perform a number of computations in which we track the state of the global ocean as a function of the strength of a persistent anomalous freshwater flux. Each of these computations results in a bifurcation diagram that provides exact information about the parameter regime for which the present-day and “collapsed” states of the MOC exist. The effect of different atmospheric feedbacks is evaluated by comparing the bifurcation diagram of a model in which the boundary conditions are (partially) interactive to the diagram of a model in which atmospheric feedbacks are absent.

We start with a description of the boundary conditions for the ocean-only (indicated by superscript oo) version of the model. The momentum fluxes, denoted by F_u^{oo} and F_v^{oo} for the zonal and meridional directions, respectively, are given by the annual mean wind stress climatology of Trenberth et al. (1989). The buoyancy fluxes are defined with respect to a reference (superscript R) solution for which the anomalous freshwater flux is zero.¹ The reference freshwater flux F_E^R (positive for evaporation; Fig. 1a) is equal to the flux implied by

¹ Although the boundary condition is implemented as a flux of salt, we adhere to the “freshwater” terminology throughout the paper.

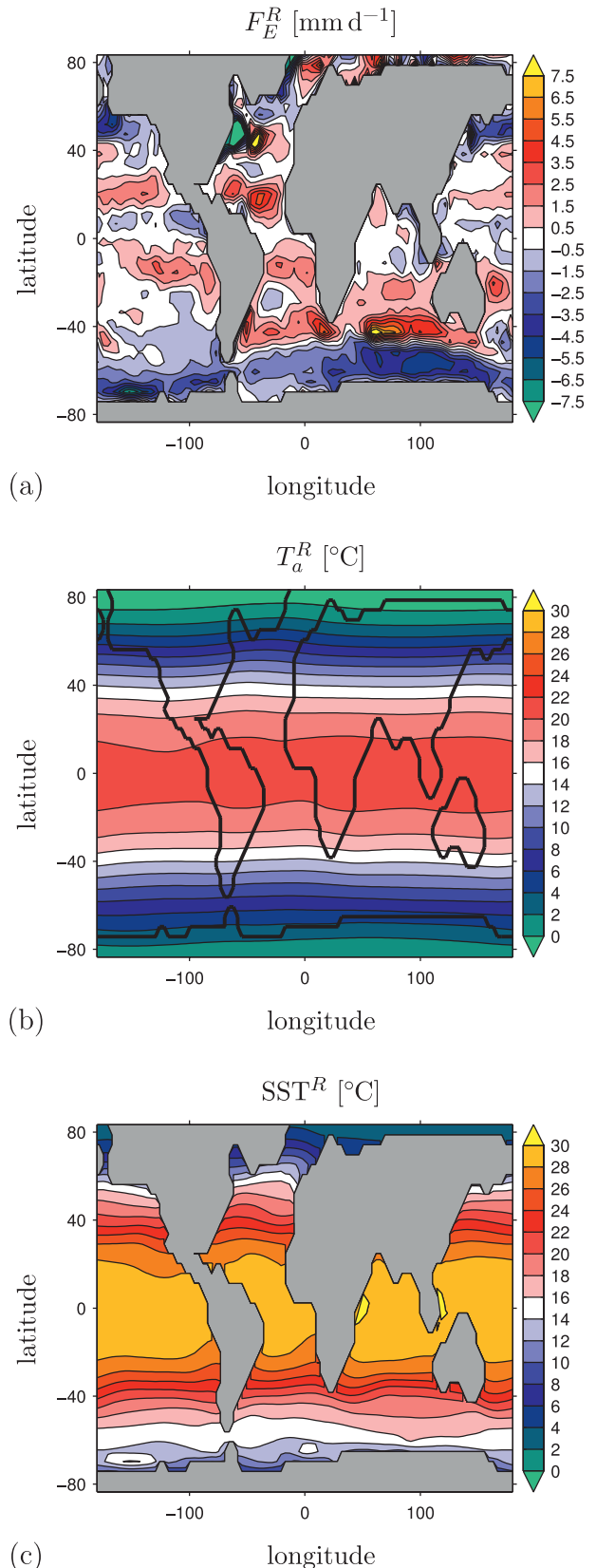


FIG. 1. Reference fields of (a) evaporation minus precipitation, (b) atmospheric temperature, and (c) SST.

a restoring boundary condition on the reference sea surface salinity using the Levitus dataset (Levitus et al. 1994) as the target field. The total freshwater flux for grid point (i, j) is now given by

$$F_E^{\text{oo}}(i, j) = F_E^R(i, j) - \gamma^p F_E^p(i, j) - \Delta F_E. \quad (1)$$

Here, γ^p is the control or bifurcation parameter in all computations. It represents the amplitude of a freshwater perturbation that is applied over a region south of Greenland, indicated by F_E^p , and defined as the area confined by longitudes 60° and 24°W and latitudes 54° and 66°N. The term ΔF_E is a spatially constant compensation that is defined such that the flux integrates to zero. In the formulation of Weijer et al. (2003), the boundary condition for sea surface temperature follows from a one-layer energy balance model, in which the atmospheric temperature T_a is computed. To avoid interference with the empirical description of the atmosphere adopted in this study, the energy balance model is disabled. This is done by fixing the atmospheric temperature to the reference field T_a^R (Fig. 1b), while retaining the coupling to the ocean. The heat flux into the ocean is then given by

$$F_Q^{\text{oo}}(i, j) = \frac{I_0}{4} S(\phi)(1 - \alpha)C_0 - \mu[SST(i, j) - T_a^R(i, j)]. \quad (2)$$

Here, $I_0 = 1360 \text{ W m}^{-2}$ is the solar constant,

$$S(\phi) = 1 - 0.241(3 \sin^2 \phi - 1)$$

the latitudinal (ϕ) dependence of the shortwave radiative heat flux, $\alpha = 0.3$ the albedo, $C_0 = 0.43$ the transmission coefficient, and $\mu = 13.0 \text{ W m}^{-2} \text{ }^\circ\text{C}^{-1}$ a coupling coefficient. The formulation of the heat flux in Eq. (2) is mathematically equivalent to a restoring condition, but the relaxation time scale (about 200 days) is much longer than the often-used 30-day damping time scale. Furthermore, note that the target field would not be T_a^R , but rather $T_a^R + (I_0/4\mu)S(\phi)(1 - \alpha)C_0$. So, since $S(\phi)$ is positive everywhere, sea surface temperatures must exceed T_a^R to realize a balance between incoming solar radiation and heat loss to the atmosphere, as illustrated in Fig. 1c for the reference sea surface temperature.

The effect of atmospheric feedbacks is represented by two additional contributions to the boundary conditions that depend linearly on SST (Cimadoribus et al. 2012). Both are assumed to vanish at the reference solution, for which $SST = SST^R$. The first contribution is proportional to local SST anomalies, while the second depends on SST averaged over the Northern Hemisphere

and is assumed to account for the large-scale changes that arise when the strength of the Atlantic MOC decreases. For heat ($C = Q$), net evaporation ($C = E$), and momentum ($C = u$ or $C = v$), the hybrid model fluxes can be written as

$$F_C(i, j) = F_C^{\text{oo}}(i, j) + \xi_C p_1^{F_C}(i, j)[SST(i, j) - SST^R(i, j)] + \xi_C p_2^{F_C}(i, j)[\langle SST \rangle_{\text{NH}} - \langle SST^R \rangle_{\text{NH}}], \quad (3)$$

where the $\langle \cdot \rangle_{\text{NH}}$ indicate the average over the Northern Hemisphere. The homotopy parameters ξ_C serve to control which of the individual feedbacks are active. The fields $p_1^{F_C}$ and $p_2^{F_C}$ in Eq. (3) are derived from a linear regression analysis on the output of two simulations with the fully coupled Speedy-Ocean (SPEEDO) model (Severijns and Hazeleger 2009) as detailed by Cimadoribus et al. (2012). The local regression parameters $p_1^{F_C}$ are calculated from a yearly averaged time series of the unperturbed SPEEDO simulation, while the large-scale parameters $p_2^{F_C}$ derive from a simulation with the same model in which the Atlantic MOC is forced to collapse.

Cimadoribus et al. (2012) have evaluated the performance of the empirical atmosphere model by coupling it to SPEEDO's ocean component, the Coupled Large-Scale Ice–Ocean model (CLIO; Goosse and Fichefet 1999). It was shown that the relevant long-term mean behavior of the MOC is essentially the same in the hybrid CLIO model and in the fully coupled SPEEDO model, thus validating the approach of regressing atmospheric fluxes onto SST.

It should be kept in mind, however, that the regression coefficients were derived with reference to the unperturbed present-day state. Predictions of the changes in MOC stability are therefore probably more reliable for the present-day MOC than for the collapsed MOC.

In addition, there are several differences between the hybrid CLIO model of Cimadoribus et al. (2012) and the implicit hybrid model described here that should be considered in the interpretation of the results. First, as discussed above, the implicit methodology places constraints on the complexity of the parameterizations of subgrid-scale mixing in THCM, while CLIO includes sophisticated schemes to represent the unresolved physics. As a consequence, both models need to use their own set of ocean-only boundary conditions. Otherwise a very unrealistic circulation would result because of the strong sensitivity of the overturning to the patterns of heat and freshwater fluxes. The hybrid CLIO model uses the climatological mean fluxes of SPEEDO, while the implicit hybrid model uses the fluxes F_C^{oo} defined above. Hence, the fundamental assumption behind Eq. (3) is that its expression of the atmospheric response to changes in

MOC strength is sufficiently generic as to give reliable results in different ocean models. A second difference is that the climatological mean fluxes and the local regression parameters vary seasonally in the hybrid CLIO, while the forcing in the implicit hybrid model is necessarily constant in time. As a result it does not capture the possible effects due to the rectification of interactive atmospheric fluxes. Several tests with the hybrid CLIO in which constant local regression parameters were used, however, indicate that this rectification effect is small. Finally, the implicit hybrid model does not include a sea ice component, whereas the hybrid CLIO does. The regression parameters p_1^{Fc} and p_2^{Fc} used here therefore include the effect of changing sea ice cover, so that they express the “effective” feedback of the combined atmosphere–sea ice system. This issue is further discussed in the paper by Cimatoribus et al. (2012, section IVB).

3. Bifurcation diagrams

a. Computation of steady states

Here we briefly sketch the techniques to find steady-state solutions of an ocean model without using conventional time stepping methods. A more elaborate description of the steady-state continuation method implemented in THCM may be found in the work of Weijer et al. (2003).

We define the state of the model by an N -dimensional vector \mathbf{x} , containing the three velocity components, the pressure, the temperature, and the salinity at each grid point. For the configuration described above, $N = 284\,544$. The space-discretized model can be expressed as a set of N coupled differential equations of the form

$$\begin{aligned} \mathbf{M} \frac{d\mathbf{x}}{dt} &= \tilde{\mathcal{F}}(\mathbf{x}, \gamma^p, \xi_C) \\ &= \mathcal{F}(\mathbf{x}, \gamma^p) + \sum_{C=Q,S,u,v} \xi_C \mathbf{L}_C (\mathbf{x} - \bar{\mathbf{x}}^R), \end{aligned} \quad (4)$$

where \mathbf{M} is a diagonal matrix, and $\tilde{\mathcal{F}}$ a nonlinear operator, which we split into two parts. The operator \mathcal{F} on \mathbf{x} is the expression of all the processes included in the ocean component of the model: the pressure gradient and Coriolis forces, advection, diffusion, etc. It depends on many parameters, but here we only highlight its dependence on the control parameter γ^p . The second part of $\tilde{\mathcal{F}}$ is the expression of the empirical atmosphere model and corresponds to the terms added to F_C^{oo} in Eq. (3). The vector $\bar{\mathbf{x}}^R$ is the reference steady state at $\gamma^p = 0$. The matrices \mathbf{L}_C define the action of the coefficients $p_1^{Fc}(i, j)$ and $p_2^{Fc}(i, j)$ on the part of \mathbf{x} that corresponds to the SSTs.

Steady-state solutions $\bar{\mathbf{x}}$ of Eq. (4) are found by Newton’s method, which can be used to find steady

states irrespective of their linear stability. Starting from an initial guess \mathbf{x}^0 , the Newton updates $\mathbf{x}^{k+1} - \mathbf{x}^k$ of each iteration (with counter k) are found from

$$\tilde{\Phi}^k (\mathbf{x}^{k+1} - \mathbf{x}^k) = -\tilde{\mathcal{F}}^k, \quad (5)$$

where $\tilde{\Phi}^k$ is the Jacobian matrix of $\tilde{\mathcal{F}}^k = \tilde{\mathcal{F}}(\mathbf{x}^k, \gamma^p, \xi_C)$. The method to solve the linear system in Eq. (5) is described in more detail in the appendix. The Newton iterations stop when the largest entry of $\tilde{\mathcal{F}}^{k+1}$ is smaller than some user-defined tolerance level. Provided that the initial estimate \mathbf{x}^0 is sufficiently close to $\bar{\mathbf{x}}$ and that $\bar{\mathbf{x}}$ is a regular point, the convergence of Newton’s method is quadratic.

b. Results

The bifurcation structure of the ocean-only model (Fig. 2), which is computed by setting all $\xi_C = 0$, includes two saddle-node bifurcations, labeled L_1 and L_2 . These bifurcations define three branches of solutions, whose characteristics are discussed in detail by Dijkstra and Weijer (2005). Figure 2a shows the bifurcation diagram with the Atlantic meridional overturning strength, Ψ_A , as a measure of the solution. We define Ψ_A as the maximum of the meridional overturning streamfunction, measured below a depth of 500 m to exclude the contribution from shallow wind-driven cells. The solutions on the branch labeled ON are linearly stable and correspond to present-day climate states, characterized by a strong overturning in the Atlantic with a typical pattern as shown in Fig. 3a. The ON state can only be maintained for $\gamma^p < \gamma^p(L_1) = 209.8 \text{ mSv}$ ($= \text{Sv} \times 10^{-3}$, where $1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$). For $\gamma^p < \gamma^p(L_2) = 48.0 \text{ mSv}$ it coexists with an unstable (UNS; Fig. 3b) and a linearly stable collapsed (OFF) state of the Atlantic MOC (Fig. 3c). In Fig. 2b we plot the diagram using the Northern Hemisphere averaged SST ($\langle \text{SST} \rangle_{\text{NH}}$) as a measure of the solution. The reduction in the Atlantic MOC strength leads to a reduction in $\langle \text{SST} \rangle_{\text{NH}}$, as is required for our representation of the large-scale atmospheric feedbacks [see discussion surrounding Eq. (3)] to be valid. The difference in temperature between ON and OFF states is approximately 0.2°C , which is comparable to the response in the ocean-only version of the hybrid CLIO model of Cimatoribus et al. (2012).

The next step is to allow the operation of air–sea interactions by setting all $\xi_C = 1$. Figure 4a shows that, although the size of the multiple equilibria regime is reduced compared to the ocean-only case, the bistability of the Atlantic overturning circulation is maintained in the hybrid coupled model. Both the ON and OFF states are destabilized by atmospheric feedbacks, in the sense that bifurcation point L_1 occurs at a smaller and L_2 at

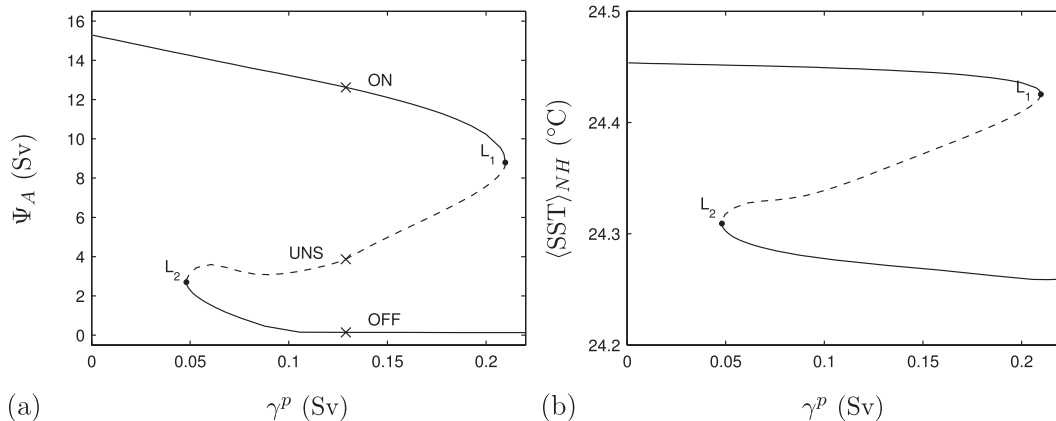


FIG. 2. Bifurcation diagram for the ocean-only model showing (a) the maximum of the Atlantic meridional overturning streamfunction, measured below 500 m, and (b) the NH averaged SST. Solid lines indicate that the branch is linearly stable, while dashed lines mean it is linearly unstable.

a larger value of γ^p than in the ocean-only model. Quantitatively, however, the shift of L_1 is bigger than the shift of L_2 (Table 1).

The relative importance of the heat ($C = Q$), freshwater ($C = E$), and momentum ($C = u$ and $C = v$) flux feedbacks in the hybrid coupled model can be assessed by setting the respective homotopy parameter $\xi_C = 1$, and leaving the others at zero. If only heat is interactive ($\xi_Q = 1$), the bifurcation diagram shifts toward smaller values of the freshwater anomaly γ^p (Fig. 4b). The destabilization of the ON state, as measured by the change in the value of γ^p corresponding to L_1 , is about as large as the stabilization of the OFF state (Table 1). In contrast, switching on only the freshwater feedback ($\xi_E = 1$) results in a significant decrease in the size of the multiple equilibria regime, because both the ON and OFF states are destabilized (Fig. 4c). When only the momentum feedback is switched on ($\xi_u = \xi_v = 1$), both the ON and OFF states can be maintained across a range of γ^p that is only slightly larger than for the ocean-only model (Fig. 4d).

The shifts of the bifurcation points L_1 and L_2 in the full hybrid model (Fig. 4a) appear to be consistent with a simple combination of the shifts in each of the partially interactive models (Figs. 4b–d). This is only true in a qualitative sense, however. As listed in Table 1, the value of γ^p at L_1 is 18.0 mSv smaller in the full hybrid model than in the ocean-only model, while the shifts of the individual feedbacks added together amount to a shift of only 13.2 mSv. The actual shift of L_2 in the full hybrid model, on the other hand, is much smaller than simple addition would predict: 3.1 rather than 18.9 mSv. Although a complete analysis of this issue is beyond the scope of this paper, two aspects of it are worth discussing in more detail.

First, we emphasize the role of interactive heat fluxes in determining the SST field. For solutions that are relatively close to the reference state, atmospheric feedbacks do not significantly alter the pattern of the circulation compared to the ocean-only case. This explains why the value of Ψ_A at saddle-node bifurcation L_1 is approximately the same in all of the cases shown in Fig. 4. Despite the similarity of the circulation, the SST field differs depending on whether the heat feedback is enabled or not. As can be deduced from Eq. (3), the sea surface temperature contrast $\text{SST}(i, j) - \text{SST}^R(i, j)$ at a particular location (i, j) is either amplified [if $F_Q(i, j) - F_Q^0(i, j)$ is of the same sign as $\text{SST}(i, j) - \text{SST}^R(i, j)$] or damped [if the two quantities are of opposite sign] by the interactive heat flux. Figure 5 shows that the Northern Hemisphere averaged SST is less sensitive to changes in Ψ_A when the heat feedback is enabled ($\xi_Q = 1$). As a consequence, the large-scale interactive freshwater and momentum fluxes, which are proportional to $\langle \text{SST} \rangle_{\text{NH}} - \langle \text{SST}^R \rangle_{\text{NH}}$, will be smaller when the heat feedback is switched on. Whether the local interactive freshwater and momentum fluxes are enhanced or reduced by the direct feedback of heat on SST depends on the location that is considered, as will be shown in section 4 for a solution close to bifurcation point L_1 . The fact that the shift of L_1 is larger in the full hybrid model than the sum of the shifts in each of the partially interactive models suggests that the heat feedback modifies the SST field in such a way that those fluxes are favored that destabilize the ON state of the MOC.

Second, we point out that as atmospheric fluxes increase in magnitude, they not only affect the strength of the MOC, but the pattern of the circulation as well. So, in addition to the direct impact of the heat feedback, interactive fluxes also influence the SST field indirectly

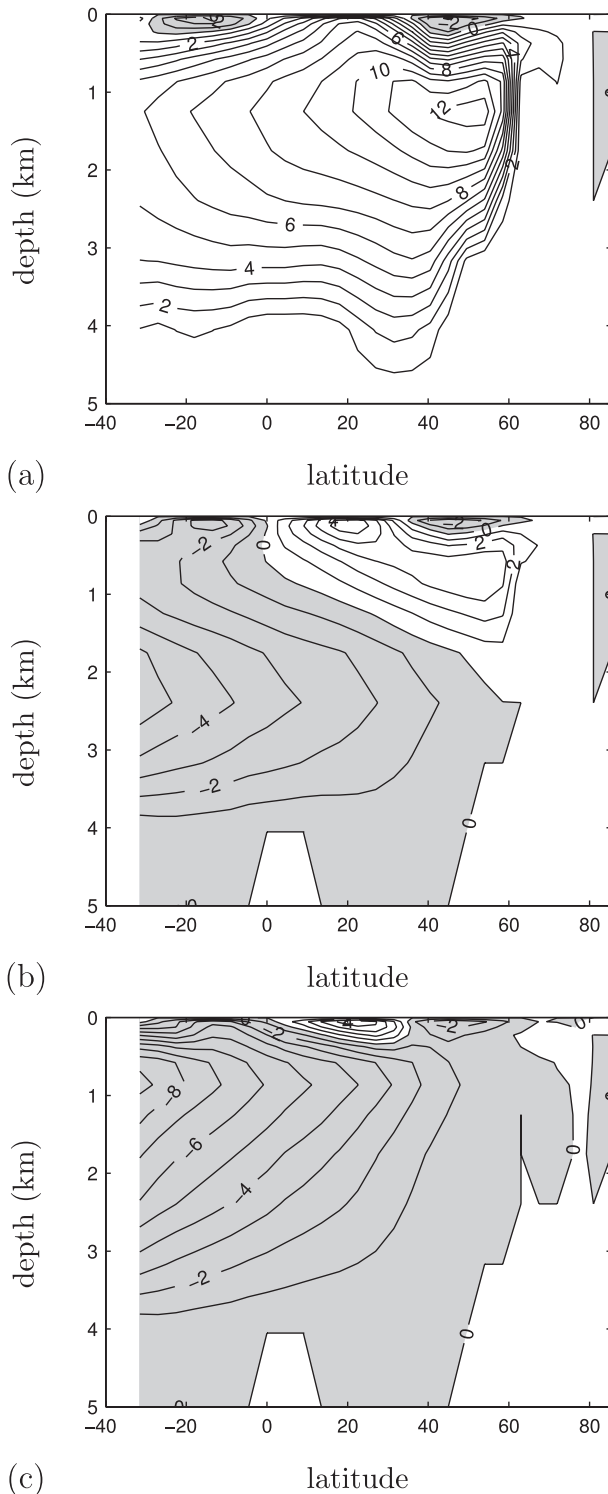


FIG. 3. Contour plots of the Atlantic meridional overturning streamfunction for the (a) ON, (b) UNS, and (c) OFF states at $\gamma^p = 0.13$ Sv (see crosses in Fig. 2a). The contour interval is 1 Sv and shading indicates negative values.

by modifying the surface velocity field. This effect explains the (slight) differences in the plots of Ψ_A versus $\langle \text{SST} \rangle_{\text{NH}}$ (Fig. 5) between the three models without interactive heat fluxes (ocean only, freshwater, and momentum), as well as the difference between the two models with interactive heat fluxes (fully hybrid and heat). The effect described here implies that there is a complex coupling between SST and all interactive atmospheric fluxes. Figure 5 suggests, however, that in our model results, indirect effects are of secondary importance compared to the direct influence of the heat feedback on SST.

As reported above, the window of bistability hardly changes when the feedbacks in the momentum fluxes are switched on (Fig. 4d). This can be attributed to the relatively high eddy viscosity that is used in THCM (Dijkstra and Weijer 2005). Using an EMIC, Arzel et al. (2011) showed that wind stress feedbacks do not alter the stability of the ON state. Consistently, coupled GCM results of Dixon et al. (1999) suggest that the ON state overturning strength is hardly affected by momentum fluxes being interactive or not. We are therefore confident in the result that saddle-node bifurcation L_1 does not shift much because of wind feedbacks. The OFF state, however, is expected to have a substantially different surface circulation as a result of altered winds (Vellinga and Wood 2002). In the hybrid CLIO model of Cimadoribus et al. (2012) the difference in $\langle \text{SST} \rangle_{\text{NH}}$ between the ON and OFF states is about 1.2°C . By switching off the wind feedback in the hybrid CLIO model it is found that interactive winds are responsible for about 1°C of this difference. The destabilization of the OFF state due to interactive freshwater fluxes (Fig. 4c) appears to be reduced in the full hybrid model (Fig. 4a) as a result of the direct impact of interactive heat fluxes on SST. This reduction may well be counteracted, however, by the indirect effect of interactive momentum fluxes that is described here. As this effect is not accurately reproduced by THCM, it is not possible to reliably compare the relative amounts by which the ON and OFF states are destabilized. This deficiency will be discussed further in section 5.

4. Adjoint sensitivity analysis

It was shown in the previous section that the addition of atmospheric feedbacks results in changes in the bifurcation diagram of the persistent freshwater perturbation γ^p . In particular, we described the shifts of the two saddle-node bifurcations, which define the multiple equilibria regime, but noted that the results for saddle-node L_2 should be treated with care if feedbacks acting through the momentum fluxes are enabled. In

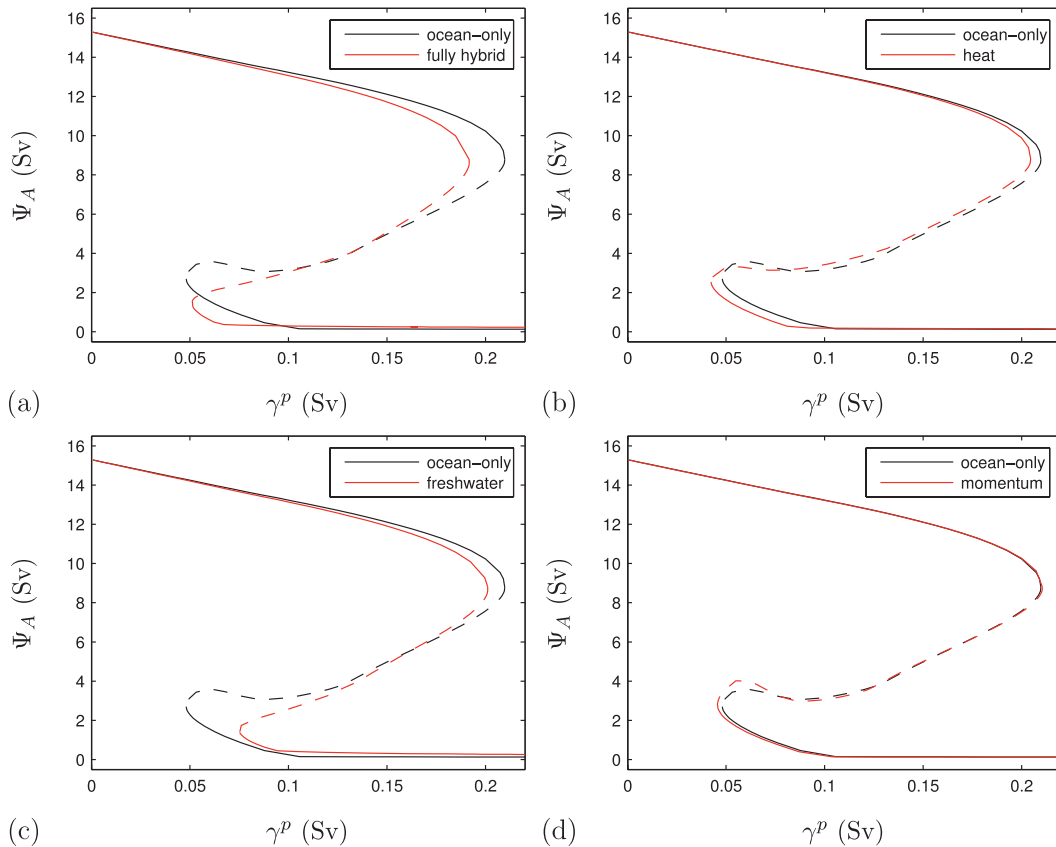


FIG. 4. Changes in the bifurcation diagram due to the surfaces fluxes being interactive. The black line represents the ocean-only case as in Fig. 2a and the red lines correspond to (a) the full HCM, and the cases with only (b) interactive heat fluxes, (c) interactive freshwater fluxes, and (d) interactive momentum fluxes enabled.

addition, we recall that the regression coefficients $p_1^{F_c}$ and $p_2^{F_c}$ [Eq. (3)] were derived with reference to the ON state with $\gamma^p = 0$. The focus of this section is therefore on the shift of saddle-node bifurcation L_1 . Using adjoint sensitivity analysis we aim to link the changes in the stability of the ON state to the physical processes responsible for it.

a. Overturning rate as measure of stability

Figure 4 shows that the overturning strength Ψ_A is roughly the same (about 8.8 Sv) at bifurcation point L_1 , irrespective of which fluxes are interactive. So, those atmospheric fluxes that cause a reduction of the overturning strength push saddle-node bifurcation L_1 to smaller values of γ^p , and vice versa. We therefore argue that it is sufficient to explain how Ψ_A is modified by atmospheric feedbacks to understand the changes in the stability of the present-day branch. In fact, as will be shown below, it is possible to compute a first-order estimate of the shift of L_1 by calculating the change in the freshwater perturbation γ^p that would oppose the

change in Ψ_A caused by enabling the atmospheric feedbacks.

b. Computation of adjoint sensitivities

Numerical steady-state continuation provides the sensitivity of a solution to a single control parameter and bifurcation diagrams can be drawn for any metric of the solution. In contrast, adjoint sensitivity analysis provides the sensitivity of a single metric of the solution to any model parameter. We argued above that a useful metric

TABLE 1. Changes in the value of γ^p (mSv) corresponding to the saddle-node bifurcation points L_1 and L_2 as a result of atmospheric feedbacks. Positive values mean that the saddle-node bifurcation shifts to larger values of γ^p .

| | L_1 | L_2 |
|------------|-------|-------|
| Heat | −5.2 | −5.8 |
| Freshwater | −8.7 | 27.2 |
| Momentum | 0.7 | −2.5 |
| All | −18.0 | 3.1 |

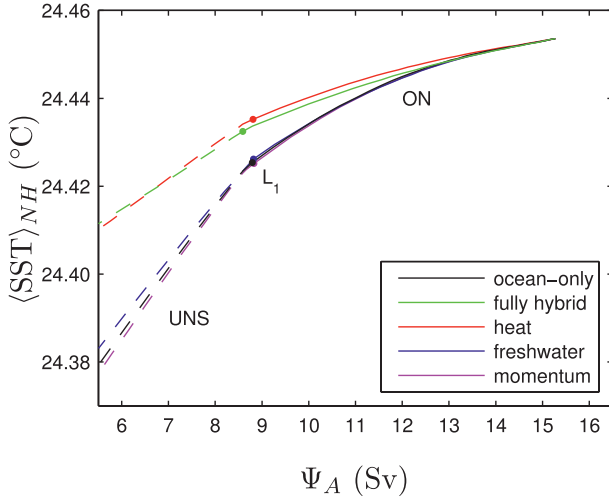


FIG. 5. The NH averaged SST vs Atlantic MOC strength for the computations with the ocean-only model, and the hybrid model with (partially) interactive fluxes as indicated. Solid (dashed) line indicates (un)stable.

is defined by the Atlantic MOC strength, which must be computed at a fixed depth and latitude (denoted by $\Psi_A^{\phi,z}$). To explain the shift of saddle-node bifurcation L_1 we consider simultaneous infinitesimal variations of the strength of one of the feedbacks and of the amplitude of the freshwater perturbation γ^p in such a way that the steady-state overturning remains constant. We start from the ocean-only solution (all ξ_C set to zero) at $\gamma^p = 0.2$ Sv, which is close to bifurcation point L_1 . For this solution the maximum of the Atlantic overturning streamfunction occurs at $\phi = 54^\circ\text{N}$ and $z = 1247$ m.

In view of the insignificance of the wind feedback on the ON state stability we limit the analysis to the heat and freshwater feedbacks. For each case let the homotopy parameter be the independent variable, so $\gamma^p = \gamma^p(\xi_C)$ and the steady-state $\bar{\mathbf{x}} = \bar{\mathbf{x}}[\gamma^p(\xi_C), \xi_C]$, where C may indicate either heat (Q) or net evaporation (E). The requirement of constant $\Psi_A^{\phi,z}$ can be expressed as

$$\frac{d\Psi_A^{\phi,z}}{d\xi_C} = \left(\frac{\partial\Psi_A^{\phi,z}}{\partial\mathbf{x}} \right)^T \left(\frac{\partial\bar{\mathbf{x}}}{\partial\gamma^p} \frac{d\gamma^p}{d\xi_C} + \frac{\partial\bar{\mathbf{x}}}{\partial\xi_C} \right) = 0, \quad (6)$$

where $\partial\Psi_A^{\phi,z}/\partial\mathbf{x}$ denotes the N -dimensional vector whose entries are formed by taking the derivative of $\Psi_A^{\phi,z}$ with respect to each state variable. The definition of $\Psi_A^{\phi,z}$ implies that the only entries that are nonzero are those corresponding to the meridional velocities in the Atlantic at $\phi = 54^\circ\text{N}$. The goal of our analysis is to compute $d\gamma^p/d\xi_C$, which is the shift of bifurcation point L_1 per unit feedback strength. Note, though, that $d\gamma^p/d\xi_C$ may vary along the path in parameter space from $\xi_C = 0$ to

$\xi_C = 1$. So, $d\gamma^p/d\xi_C$ must be considered as a first-order estimate of the total shift of L_1 .

The following step is to compute the derivative of Eq. (4) with respect to ξ_C , and to evaluate it at $\xi_C = 0$:

$$\Phi \left(\frac{\partial\bar{\mathbf{x}}}{\partial\gamma^p} \frac{d\gamma^p}{d\xi_C} + \frac{\partial\bar{\mathbf{x}}}{\partial\xi_C} \right) + \frac{\partial\mathcal{F}}{\partial\gamma^p} \frac{d\gamma^p}{d\xi_C} + \mathbf{L}_C(\bar{\mathbf{x}} - \bar{\mathbf{x}}^R) = \mathbf{0}, \quad (7)$$

where Φ is the Jacobian of the ocean-only operator \mathcal{F} introduced in Eq. (4). Next, we define the adjoint sensitivity vector \mathbf{y} as the solution of

$$\mathbf{y}^T \Phi = - \left(\frac{\partial\Psi_A^{\phi,z}}{\partial\mathbf{x}} \right)^T. \quad (8)$$

Finally, we combine Eqs. (6)–(8) to find

$$\frac{d\Psi_A^{\phi,z}}{d\xi_C} = \mathbf{y}^T \frac{\partial\mathcal{F}}{\partial\gamma^p} \frac{d\gamma^p}{d\xi_C} + \mathbf{y}^T \mathbf{L}_C(\bar{\mathbf{x}} - \bar{\mathbf{x}}^R) = 0. \quad (9)$$

The sensitivity of $\Psi_A^{\phi,z}$ to simultaneous variations of γ^p and ξ_C is now elegantly expressed as the inner product of the sensitivity to a change in forcing and a change in forcing per unit feedback strength. The change in forcing per unit feedback strength is split into a contribution from the anomalous freshwater flux and a contribution from the empirical atmosphere model. The compensation between these two defines the shift per unit feedback strength. By examining the spatial patterns of the adjoint sensitivity vector \mathbf{y} and the forcing per unit feedback strength, we get insight into the physical mechanisms that are responsible for the shift of the saddle-node bifurcation.

In the following two subsections we discuss the term $\mathbf{y}^T \mathbf{L}_C(\bar{\mathbf{x}} - \bar{\mathbf{x}}^R)$ in Eq. (9), first for heat ($C = Q$) and then for net evaporation ($C = E$). The vector $\mathbf{L}_C(\bar{\mathbf{x}} - \bar{\mathbf{x}}^R)$ is the derivative of the empirical atmosphere model flux with respect to the homotopy parameter. Only the entries corresponding to surface variables are nonzero, and follow from Eq. (3):

$$\begin{aligned} \frac{\partial F_C(i,j)}{\partial\xi_C} &= p_1^{F_C}(i,j)[\text{SST}(i,j) - \text{SST}^R(i,j)] \\ &+ p_2^{F_C}(i,j)[\langle\text{SST}\rangle_{\text{NH}} - \langle\text{SST}^R\rangle_{\text{NH}}]. \end{aligned} \quad (10)$$

The temperature contrast $\text{SST} - \text{SST}^R$ for $\gamma^p = 0.2$ Sv is shown in Fig. 6. The most pronounced signal is seen in the Atlantic basin, where the reduction of the overturning strength (Ψ_A decreases by about 5 Sv; Fig. 2a) causes strong and widespread cooling in the North Atlantic

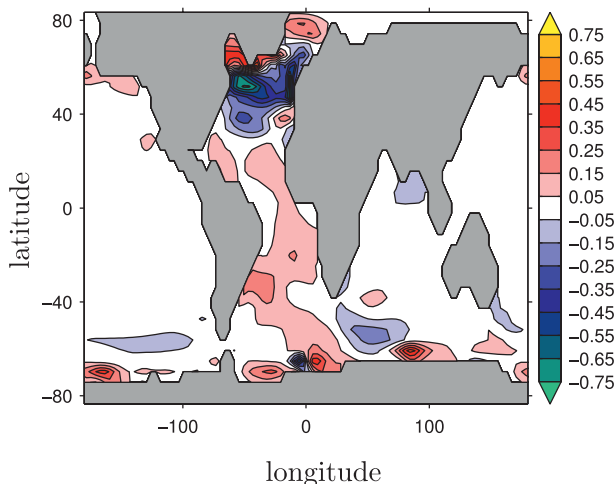


FIG. 6. Difference between the SST field of the ON state solution at $\gamma^p = 0.2$ Sv and the reference field (Fig. 1c).

between 30° and 60°N , and a moderate warming in the region south of 30°N . The increase in SST in the seas surrounding Greenland is related to the application of the freshwater perturbation. The local (equilibrium) response to this perturbation is enhanced upwelling of relatively saline water. Because the deeper water is also relatively warm, the SSTs increase and more heat is lost to the atmosphere. Outside the Atlantic, the SST contrast is confined to a number of small areas around Antarctica. Averaged over the Northern Hemisphere, the SST difference is $\langle \text{SST} \rangle_{\text{NH}} - \langle \text{SST}^R \rangle_{\text{NH}} = -0.018^\circ\text{C}$.

c. Heat feedback

Figure 7a shows the part of \mathbf{y} [Eq. (8)] that corresponds to the sensitivity of $\Psi_A^{\phi,z}$ to local changes in the surface heat flux, which is defined as positive into the ocean. As demonstrated by Dijkstra (2008) there is a linear relation between Ψ_A and the large-scale north-south density difference, which helps to explain the patterns in Fig. 7a. In examining a variety of different locations on the grid, we find that if at a grid point (i, j) a small, persistent perturbation is applied to the background heat flux, the first-order (steady state) response is restricted to the local sea surface temperature, $\text{SST}(i, j)$. The anomaly does not spread horizontally because the time scale associated with the standard atmospheric damping [Eq. (2)] is short compared to the advective time scale. It appears that such a heat flux perturbation can only bring about a significant change in the large-scale density difference, and hence the overturning strength, if there is deep-reaching downwelling at location (i, j) . Consistent with this idea, the sensitivity of $\Psi_A^{\phi,z}$ to changes in the heat flux is strongly negative in the

northeastern part of the Atlantic basin, which is the region where dense water sinks in the model. The strongly positive sensitivity in the region along the coast of South America also coincides with a zone of downwelling. In this area the vertical velocity is negative from the surface downward to a depth of about 1000 m because of the convergence of the northward-flowing extension of the Antarctic Circumpolar Current (the model's representation of the Malvinas Current) and the southward-flowing western boundary current of the subtropical gyre (the Brazil Current). In the largest part of the Atlantic that we did not discuss so far, the adjoint sensitivity of $\Psi_A^{\phi,z}$ to changes in the heat flux is weakly negative. In these areas there is no deep injection of surface waters. When a perturbation is applied here, it affects the MOC only indirectly by changing the surface velocity field.

As discussed more extensively by Cimadoribus et al. (2012), the local regression parameter field $p_1^{F_o}$ (Fig. 7b), as found from the SPEEDO model, can be interpreted as the expression of different physical feedback processes. We remind the reader that the regression parameters include the net effect of changes in sea ice. The strong damping in the (sub)polar areas (negative values in Fig. 7b) is mainly accounted for by the fact that the area of the ocean covered by sea ice decreases with increasing SST, while the ocean is losing heat to the atmosphere in this region. A positive SST anomaly at a certain grid point is thus damped, because it causes a larger fraction of the gridcell area to be exposed to upward heat fluxes. The same applies conversely to negative SST anomalies. So, the positive temperature contrast in the seas surrounding Greenland (Fig. 6) induces cooling (Fig. 7c). The net effect on $\Psi_A^{\phi,z}$, however, is small because the sensitivity of $\Psi_A^{\phi,z}$ to changes in the heat flux is weak and of opposite sign in the Labrador Sea and the Nordic Seas. The SST changes around Antarctica are also damped because of sea ice dynamics, but the net effect on $\Psi_A^{\phi,z}$ is again small.

Outside the polar areas, local heat flux changes are dominated by changes in latent heat release, followed in importance by changes in sensible heat and long-wave radiation (Cimadoribus et al. 2012). In the (northern) midlatitudes and in the equatorial region, the local heat feedback enhances the standard atmospheric damping of SST anomalies. In the (sub)tropics, on the other hand, the heat feedback tends to amplify SST anomalies, probably as a result of the convection–evaporation feedback mechanism (Zhang et al. 1995; Chang et al. 1997). On climatically relevant time scales, positive SST anomalies in these regions give rise to enhanced convection and thus to increased low-level convergence, weaker surface winds, and hence reduced latent heat

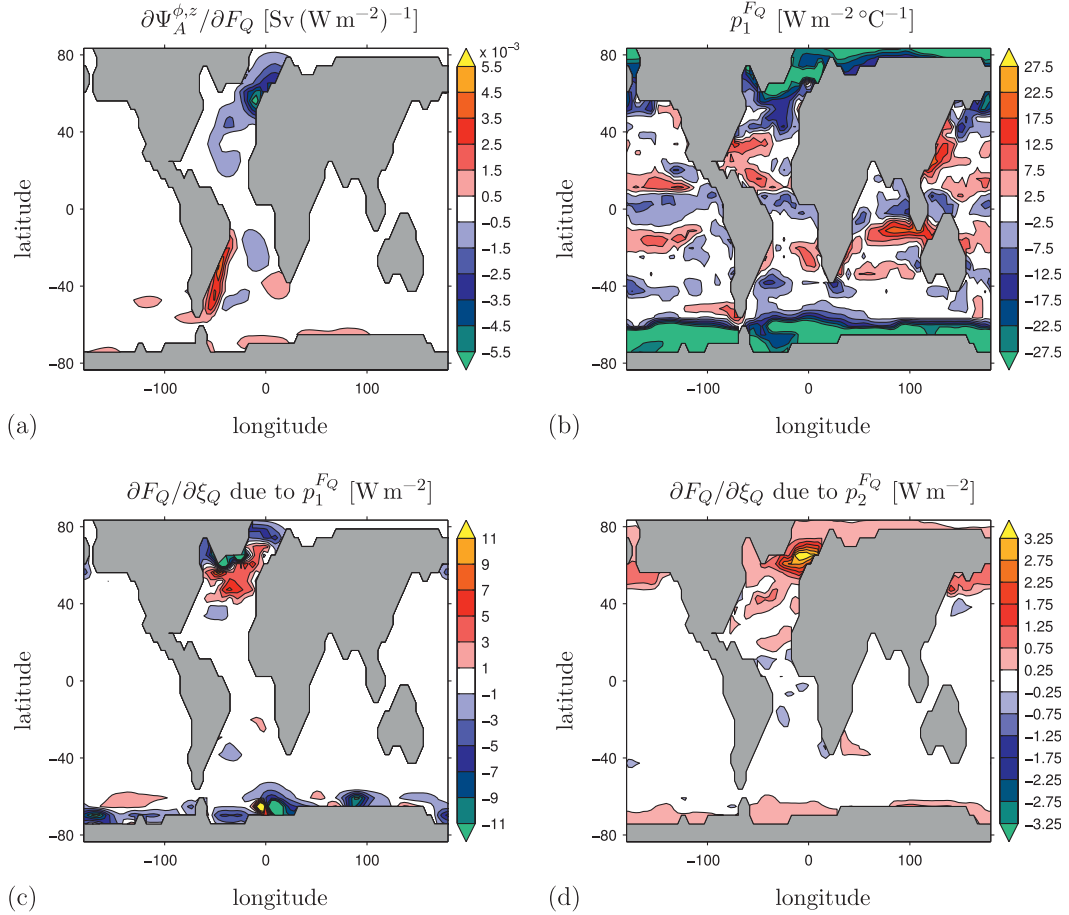


FIG. 7. Analysis of the heat feedback for the ON state solution at $\gamma^p = 0.2$ Sv. All heat fluxes are positive downward. (a) Map of the adjoint sensitivity of the steady-state Atlantic MOC strength with respect to local variations in heat flux. (b) Local regression parameter. (c),(d) Derivative of the heat flux with respect to the homotopy parameter as in Eq. (10), with the flux due to the (c) local regression parameter and (d) large-scale coupling.

loss by the ocean. It is worth noting that the different responses of latent heat are roughly consistent with the pattern of the local freshwater regression parameter $p_1^{F_E}$ (Fig. 8b). The strongly negative temperature contrast in the midlatitude North Atlantic mainly projects onto negative values of $p_1^{F_Q}$ and is therefore counteracted by local heating. Because of the negative adjoint sensitivity of $\Psi_A^{\phi,z}$ to changes in the heat flux, this finally results in a decrease in overturning strength. The heat fluxes associated with the positive heat contrast south of 30°N are relatively weak, and have little effect on $\Psi_A^{\phi,z}$.

The combined effect of the different physical feedbacks leads to an overall sensitivity of $\Psi_A^{\phi,z}$ to the local heat feedback strength of -0.44 Sv per unit feedback strength.

Figure 7d shows the heat flux due to the large-scale feedback, calculated as the product of $\langle \text{SST} \rangle_{\text{NH}} - \langle \text{SST}^R \rangle_{\text{NH}} = -0.018^\circ\text{C}$ and the large-scale regression parameter field $p_2^{F_Q}$. The large-scale heat feedback is

entirely dominated by changes in insulation due to changing sea ice cover. Similar to the expression of this effect in the local regression parameter $p_1^{F_Q}$ (Fig. 7b), this causes $p_2^{F_Q}$ to be negative and hence the large-scale interactive heat flux to be positive in the subarctic region. Regions where the adjoint sensitivity of $\Psi_A^{\phi,z}$ to changes in the heat flux is strongly negative roughly overlap with regions of strong large-scale heat fluxes, resulting in an overall sensitivity of -0.22 Sv per unit feedback strength. The total response of $\Psi_A^{\phi,z}$ to enabling the heat feedback is thus $\mathbf{y}^T \mathbf{L}_Q (\bar{\mathbf{x}} - \bar{\mathbf{x}}^R) = -0.66$ Sv per unit feedback strength.

d. Freshwater feedback

The part of \mathbf{y} [Eq. (8)] corresponding to the sensitivity of $\Psi_A^{\phi,z}$ to local changes in the net evaporative flux (Fig. 8a) shows a less localized pattern than the sensitivity to the heat flux (Fig. 7a), which is due to the fact that sea surface salinity anomalies are not damped by

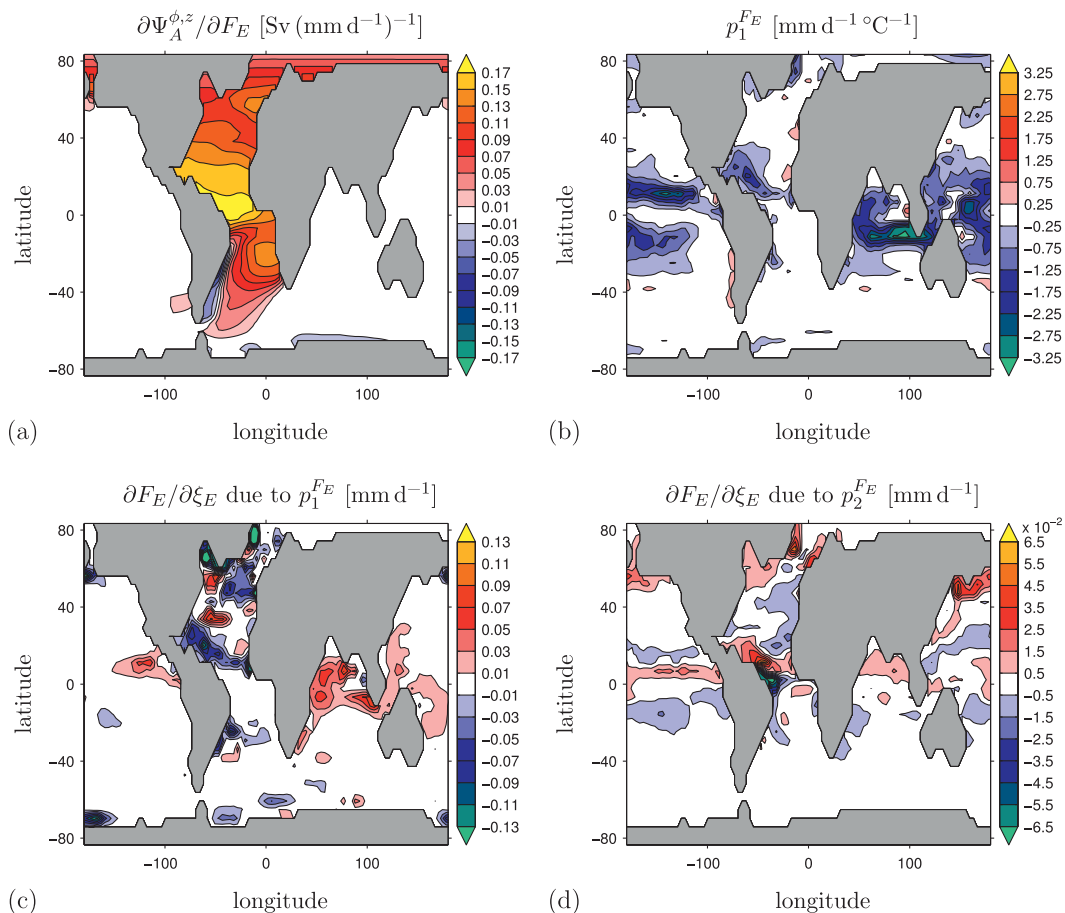


FIG. 8. As in Fig. 7, but for the freshwater feedback. Positive fluxes represent net evaporation.

the atmosphere. By applying a small, persistent perturbation to the local freshwater flux at a number of different locations, we find for each case that such a perturbation produces a steady state with a large-scale anomaly in the sea surface salinity, in contrast to the localized response of SST to local heat flux perturbations. A positive perturbation (enhanced evaporation) typically causes sea surface salinities to increase upstream, and to decrease downstream of the location where the perturbation is applied. It thus appears that perturbations in the freshwater flux can influence the MOC strength directly by exerting a remote influence on the salinity field in the regions of strong downwelling, and hence on the large-scale density contrast. Consistently, the sensitivity of $\Psi_A^{\phi,z}$ to changes in the net evaporative flux is positive in the Atlantic basin, except in the region close to the southern coast of South America. Since gridcell area decreases with latitude, a smaller change in net evaporation (which is measured in mm day^{-1}) is needed at the equator than at higher latitudes to produce the same change in the total amount of salt in a grid cell. The fact that the strongest sensitivities

are found at the equator therefore fits with the idea that evaporation changes have a remote but direct effect on the salinity in the downwelling region and hence on the strength of the MOC.

The local freshwater regression parameter $p_1^{F_E}$ (Fig. 8b) is largest in magnitude in the subtropics and tropics. The evaporation–convection feedback (Zhang et al. 1995) is at work in these areas: higher SSTs lead to enhanced convection and stronger precipitation. The SST contrast is positive in the (sub)tropical part of the Atlantic where $p_1^{F_E}$ is negative. This means that the evaporation–convection feedback leads to a net input of freshwater into the Atlantic (Fig. 8c), and hence to a decrease in overturning strength. In the midlatitudes the parameter $p_1^{F_E}$ is weakly positive, implying that net evaporation is enhanced at higher SSTs. A significant part of the negative midlatitude SST contrast projects onto the area of positive $p_1^{F_E}$ and thus leads to a smaller overturning rate. The positive SST contrast at the equator does not have a significant effect on $\Psi_A^{\phi,z}$ though. In the polar areas and south of Greenland the local feedback causes an enhanced freshwater input at higher SSTs,

which is probably a combined effect of enhanced discharge from the land surface and increased precipitation. The positive temperature contrast in the Labrador and Nordic Seas therefore results in a net freshwater input and hence in a decrease in $\Psi_A^{\phi,z}$. The overall sensitivity of $\Psi_A^{\phi,z}$ to the local freshwater feedback strength equals -0.72 Sv per unit feedback strength.

The dominant signal in the large-scale freshwater regression parameter $p_2^{F_E}$ is the signature of the southward shift of the intertropical converge zone (Yin et al. 2006) in response to a weaker overturning circulation (Fig. 8d). However, because the adjoint sensitivity of $\Psi_A^{\phi,z}$ to changes in the net evaporative flux is relatively constant just north of the equator, the ITCZ shift hardly affects the Atlantic overturning rate. The sensitivity to the large-scale feedback strength is therefore only 0.03 Sv per unit feedback strength. The total sensitivity of $\Psi_A^{\phi,z}$ to the freshwater feedback is hence $\mathbf{y}^T \mathbf{L}_E (\bar{\mathbf{x}} - \bar{\mathbf{x}}^R) = -0.69$ Sv per unit feedback strength.

e. Shift of saddle-node bifurcation L_1

As explained in section 4b, the final aim of our analysis is to calculate $d\gamma^p/d\xi_C$ from Eq. (9):

$$\frac{d\gamma^p}{d\xi_C} = -\mathbf{y}^T \mathbf{L}_C (\bar{\mathbf{x}} - \bar{\mathbf{x}}^R) \left(\mathbf{y}^T \frac{\partial \mathcal{F}}{\partial \gamma^p} \right)^{-1}, \quad (11)$$

which provides a first-order estimate of the shift of bifurcation point L_1 . The term $\mathbf{y}^T \mathbf{L}_C (\bar{\mathbf{x}} - \bar{\mathbf{x}}^R)$ was computed in sections 4c and 4d for feedbacks acting through the heat ($C = Q$) and freshwater ($C = E$) fluxes, respectively. The term $\mathbf{y}^T (\partial \mathcal{F} / \partial \gamma^p)$ represents the sensitivity of $\Psi_A^{\phi,z}$ to γ^p and can be thought of as the slope of the tangent to the bifurcation diagram at $\gamma^p = 0.2$ Sv. It follows from Eq. (1) that $(\partial \mathcal{F} / \partial \gamma^p)$ corresponds to the perturbation pattern $-F_E^p$. Using the adjoint sensitivity to net evaporative fluxes (Fig. 8a), the sensitivity of Ψ_A to γ^p is found to be -0.0682 Sv mSv^{-1} . The total response of Ψ_A to enabling the heat feedback (-0.66 Sv per unit feedback strength) can thus be compensated by a shift in γ^p of $d\gamma^p/d\xi_Q = -9.7$ mSv per unit feedback strength. The actual shift is smaller (-5.2 mSv; see Table 1) because of the direct negative feedback between the interactive forcing and SST. The total response of Ψ_A to enabling the freshwater feedback (-0.69 Sv per unit feedback strength) can be counteracted by a shift in γ^p of $d\gamma^p/d\xi_E = -10.1$ mSv per unit feedback strength, which is very close to the actual shift of -8.7 mSv (Table 1). The fact that the first-order estimates of the shifts are close to the actual values indicates the relevance of the physical mechanisms described above for explaining the changes in the stability of the ON state.

5. Summary and discussion

Using a global hybrid ocean–atmosphere model and techniques from numerical bifurcation theory to solve for exact steady states of this model, we investigated the changes in the multiple equilibria regime of the Atlantic meridional overturning circulation (MOC) due to atmospheric feedbacks. The main new element of this study is a systematic assessment of how different interactive ocean–atmosphere fluxes affect the position of the two saddle-node bifurcations that bound the multiple equilibria regime of the MOC.

Atmospheric feedbacks cause the saddle-node bifurcation that terminates the present-day (ON) branch of solutions (L_1) to occur at a smaller magnitude of the freshwater perturbation (γ^p) than in the ocean-only model (Fig. 4). Feedbacks acting through the freshwater flux appear to contribute most to the destabilization of the ON state. Their effect is probably enhanced by the direct impact of interactive heat fluxes on the sea surface temperature (SST) field. The implication of this result is that, if the freshwater input into the northern part of the North Atlantic increases gradually in time, the threshold value for collapse will be reached sooner as a result of air–sea interactions. If, instead, the present-day state of the ocean is perturbed by a temporary freshwater pulse, a smaller magnitude or shorter duration of the pulse will suffice to force a transition to the collapsed (OFF) state (Dijkstra et al. 2004).

In general, the MOC will recover from a collapse induced by a temporary perturbation if the unperturbed state is in the unique regime, but it will remain collapsed when it is in the bistable regime. This means that the bifurcation point that terminates the OFF branch of solutions (L_2) must also be considered when assessing the nonlinear sensitivity of the MOC to freshwater perturbations. The hybrid coupled model predicts that L_2 occurs at a larger value of γ^p . Although its position on the bifurcation diagram is uncertain, the consequence might be that the present-day state of the climate is brought to the unique regime as a result of atmospheric feedbacks.

Atmospheric feedbacks arise because changes in the MOC induce changes in the atmospheric momentum, heat, and freshwater fluxes, which in turn influence the strength and pattern of the MOC. Using adjoint methods we identified which physical mechanisms are responsible for the shift of L_1 . We analyzed the sensitivity of the MOC to the feedback strengths of heat and freshwater, which dominate the response (section 4). The impact of interactive heat fluxes on the MOC stability is roughly limited to the region of the North Atlantic subpolar gyre (Fig. 7a), where the regression coefficients of the

hybrid model represent atmosphere and sea ice dynamics that act to reduce the negative SST anomalies (Fig. 6). Interactive freshwater fluxes, on the other hand, have an effect everywhere in the Atlantic basin (Fig. 8a). We identified the (sub)tropical evaporation–convection feedback (Zhang et al. 1995; Chang et al. 1997), the mid-latitude coupling between SST and net evaporation, and the polar changes in precipitation and runoff as important mechanisms for the MOC stability. Because of the SST signature of the MOC decrease (Fig. 6), their net effect is to shift the saddle-node L_1 to smaller values of γ^p . The shift of the ITCZ captured by the large-scale regression parameter was found to have little effect on the ON state MOC stability.

The qualitative results on the shifts of the saddle-node bifurcations and the mechanisms derived from the ad-joint sensitivity are more important than the quantitative shifts themselves. We need to recall here that the hybrid model is based on regression coefficients that were derived with reference to the ON state with $\gamma^p = 0$. Furthermore, we noted that the small response to winds being interactive is due to the relatively high lateral viscosity of the ocean model. The results of Arzel et al. (2011) suggest that this deficiency only affects the position of L_2 . In addition to possible direct effects, which only modify the strength of the overturning but not its pattern, interactive momentum fluxes influence the OFF state SST field (and hence all atmospheric fluxes) indirectly by modifying the surface expression of the overturning. We found that this effect is not reproduced by THCM. As a consequence of these limitations, predictions for the shift of L_1 will be more reliable than for the shift of L_2 .

Another weakness of the hybrid model approach is that certain feedbacks may be underestimated or even ignored. First, although the linearized effect of sea ice on the heat flux feedback is taken into account through the regression parameters (Cimatoribus et al. 2012), the full nonlinear response due to sea ice dynamics is not captured. Furthermore, it should be realized that the effect of the hydrological cycle on the stability of the MOC is a very complex issue. It is possible that for a different mean climate, other feedbacks dominate over the ones that were captured in the regression coefficients, which were derived with respect to the present climate. For example, Tziperman and Gildor (2002) pointed out that the temperature–precipitation feedback, according to which there is less precipitation over higher latitudes during colder periods, may stabilize the thermohaline circulation, in contrast to our results. Additional feedbacks may also arise if other parameters, such as CO_2 concentration, are varied besides the high-latitude input of freshwater. For instance, Latif et al. (2000) showed

how tropical air–sea interactions may stabilize the MOC in a greenhouse warming scenario.

Our motivation to study the effect of atmospheric feedbacks stems from the apparent absence of a multiple equilibria regime in state-of-the-art IPCC AR4 models. The results in this paper show (i) that atmospheric feedbacks reduce the size of the multiple equilibria regime, or hysteresis width, of the Atlantic MOC, and (ii) that the present-day state is more vulnerable to long-lasting increases in the northern high-latitude freshwater input when air–sea interactions are taken into account. Emphasizing that the results are mainly qualitative, it is not expected that atmospheric feedbacks remove the multiple equilibria regime. In agreement with Hofmann and Rahmstorf (2009), we argue that the feedbacks are too weak to induce substantial changes in the positions of the saddle-node bifurcations. The extent of the multiple equilibria regime in state-of-the-art GCMs is hence more likely affected by biases in the GCMs (Huisman et al. 2010; Drijfhout et al. 2011) than by the effect of atmospheric feedbacks.

Acknowledgments. The authors thank the three anonymous reviewers for their valuable comments. Computations were done on the Huygens IBM p6 supercomputer at SARA Amsterdam. Use of these computing facilities was sponsored by the National Computing Facilities Foundation (NCF) under Project SH084-08 with financial support from the Netherlands Organization for Scientific Research (NWO). This work was further supported by an NWO Top talent Grant to one of the authors (MdT).

APPENDIX

Steady-State Continuation of the Hybrid Coupled Model

Steady-state solutions of Eq. (4) are found using Newton's method, which involves solving linear systems with the Jacobian matrix

$$\tilde{\Phi} = \Phi + \sum_{C=Q,S,u,v} \xi_C \mathbf{L}_C. \quad (\text{A1})$$

The matrix Φ is the Jacobian matrix of the ocean-only operator \mathcal{F} . Because the stencil used for space discretization is compact, this matrix is sparse. As a result, systems with Φ can be solved relatively efficiently using preconditioning and iterative solvers (Weijer et al. 2003). The matrices \mathbf{L}_C , however, are dense because of the averaging operations implied by the definitions of $\langle \text{SST} \rangle_{\text{NH}}$ and ΔF_E . This means that $\tilde{\Phi}$ cannot be handled efficiently with iterative solution techniques.

We therefore separate out the averaging operations and rewrite the hybrid model as

$$\mathbf{M} \frac{d\mathbf{x}}{dt} = \tilde{\mathcal{F}}(\mathbf{x}, \gamma^p, \xi_C) = \mathcal{F}'(\mathbf{x}, \gamma^p, \xi_C) + p\mathbf{P}(\xi_C) + q\mathbf{Q}, \quad (\text{A2})$$

where the additional variables p and q are the Northern Hemisphere averaged temperature, that is,

$$\langle \text{SST} \rangle_{\text{NH}} - p = 0, \quad (\text{A3})$$

and the freshwater correction, that is,

$$\Delta F_E - q = 0. \quad (\text{A4})$$

The vector \mathbf{P} contains the large-scale regression coefficients $p_2^{Fc}(i, j)$, while \mathbf{Q} corresponds to a spatially uniform correction to the sea surface salinity field. The remaining sparse parts of the matrices $\xi_C \mathbf{L}_C$, which correspond to the local coupling through the regression coefficients $p_1^{Fc}(i, j)$, are added to \mathcal{F} , and the result is indicated by \mathcal{F}' . As will be explained below, the advantage of introducing the additional variables p and q is that the

calculation of Newton updates now only requires solving systems with the Jacobian matrix of \mathcal{F}' (denoted by Φ'), which is a sparse matrix.

Steady states are tracked in parameter space by a pseudoarclength parameterization of the branches (Keller 1977). So, yet another additional parameter, s , is introduced, which is used to parameterize a branch of solutions $[\bar{\mathbf{x}}(s), \bar{p}(s), \bar{q}(s), \gamma^p(s)]$. The set of equations is closed by adding an equation of the form

$$r(\mathbf{x}, p, q, \gamma^p) = \dot{\mathbf{x}}_0^T(\mathbf{x} - \mathbf{x}_0) + \dot{p}_0(p - p_0) + \dot{q}_0(q - p_0) + \dot{\gamma}_0^p(\gamma^p - \gamma_0^p) - \Delta s = 0. \quad (\text{A5})$$

The subscript 0 is used to denote a previous solution on the branch, dots indicate the derivative with respect to s , and Δs is the step size of the arclength parameter.

In addition to the N -dimensional state vector \mathbf{x} , three additional variables, p , q , and s , were introduced, so that the total number of equations that must be solved is $N + 3$. Omitting the Newton iteration counter k in the Jacobian, the Newton updates for this extended system [Eqs. (A2)–(A5)] are found from

$$\begin{pmatrix} \Phi' & \mathbf{P} & \mathbf{Q} & \frac{\partial \mathcal{F}'}{\partial \gamma^p} \\ \left(\frac{\partial \langle \text{SST} \rangle_{\text{NH}}}{\partial \mathbf{x}} \right)^T & -1 & 0 & 0 \\ \left(\frac{\partial \Delta F_E}{\partial \mathbf{x}} \right)^T & \frac{\partial \Delta F_E}{\partial p} & -1 & \frac{\partial \Delta F_E}{\partial \gamma^p} \\ \dot{\mathbf{x}}_0^T & \dot{p}_0 & \dot{q}_0 & \dot{\gamma}_0^p \end{pmatrix} \begin{pmatrix} \mathbf{x}^{k+1} - \mathbf{x}^k \\ p^{k+1} - p^k \\ q^{k+1} - q^k \\ \gamma^{p,k+1} - \gamma^{p,k} \end{pmatrix} = \begin{pmatrix} -\tilde{\mathcal{F}}^k \\ p^k - \langle \text{SST} \rangle_{\text{NH}} \\ q^k - \Delta F_E \\ -r^k \end{pmatrix}.$$

The extended Jacobian thus consists of the sparse Jacobian matrix Φ' , bordered by three dense rows and columns. We exploit the sparsity of Φ' by first solving the four systems $\Phi' \mathbf{z}_1 = -\tilde{\mathcal{F}}$, $\Phi' \mathbf{z}_2 = \mathbf{P}$, $\Phi' \mathbf{z}_3 = \mathbf{Q}$, and $\Phi' \mathbf{z}_4 = (\partial \mathcal{F}' / \partial \gamma^p)$ for \mathbf{z}_i , $i = 1, \dots, 4$, and then constructing the Newton update from the \mathbf{z}_i .

REFERENCES

- Arzel, O., M. H. England, and O. A. Saenko, 2011: The impact of wind stress feedback on the stability of the Atlantic meridional overturning circulation. *J. Climate*, **24**, 1965–1984.
- Bryan, F., 1986: High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature*, **323**, 301–304.
- Cessi, P., and W. R. Young, 1992: Multiple equilibria in two-dimensional thermohaline circulation. *J. Fluid Mech.*, **241**, 291–309.
- Chang, P., L. Ji, and H. Li, 1997: A decadal climate variation in the tropical Atlantic Ocean from thermodynamic air-sea interactions. *Nature*, **385**, 516–518.
- Cimadoribus, A. A., S. S. Drijfhout, and H. A. Dijkstra, 2012: A global hybrid coupled model based on atmosphere-SST feedbacks. *Climate Dyn.*, **38**, 745–760.
- Clark, P. U., N. G. Pisias, T. F. Stocker, and A. J. Weaver, 2002: The role of the thermohaline circulation in abrupt climate change. *Nature*, **415**, 863–869.
- den Toom, M., H. A. Dijkstra, and F. W. Wubs, 2011: Spurious multiple equilibria introduced by convective adjustment. *Ocean Modell.*, **38** (1–2), 126–137.
- Dijkstra, H. A., 2005: The thermohaline circulation. *Nonlinear Physical Oceanography: A Dynamical Systems Approach to the Large Scale Ocean Circulation and El Niño*, H. A. Dijkstra, Ed., Springer, 267–368.
- , 2008: Scaling of the Atlantic meridional overturning circulation in a global ocean model. *Tellus*, **60A**, 749–760.
- , and W. Weijer, 2005: Stability of the global ocean circulation: Basic bifurcation diagrams. *J. Phys. Oceanogr.*, **35**, 933–948.
- , L. Te Raa, and W. Weijer, 2004: A systematic approach to determine thresholds of the ocean's thermohaline circulation. *Tellus*, **56A**, 362–370.
- Dixon, K. W., T. L. Delworth, M. J. Spelman, and R. J. Stouffer, 1999: The influence of transient surface fluxes on North Atlantic

- overturning in a coupled GCM climate change experiment. *Geophys. Res. Lett.*, **26**, 2749–2752.
- Drijfhout, S. S., S. L. Weber, and E. van der Waluw, 2011: The stability of the MOC as diagnosed from model projections for pre-industrial, present and future climates. *Climate Dyn.*, **37**, 1575–1586.
- Ganopolski, A., and S. Rahmstorf, 2001: Rapid changes of glacial climate simulated in a coupled climate model. *Nature*, **409**, 153–158.
- Goosse, H., and T. Fichefet, 1999: Importance of ice-ocean interactions for the global ocean circulation: A model study. *J. Geophys. Res.*, **104** (C10), 23 337–23 355.
- Hawkins, E., R. S. Smith, L. C. Allison, J. M. Gregory, T. J. Woollings, H. Pohlmann, and B. de Cuevas, 2011: Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport. *Geophys. Res. Lett.*, **38**, L10605, doi:10.1029/2011GL047208.
- Hofmann, M., and S. Rahmstorf, 2009: On the stability of the Atlantic meridional overturning circulation. *Proc. Natl. Acad. Sci.*, **106**, 20 584–20 589.
- Huisman, S. E., M. den Toom, H. A. Dijkstra, and S. Drijfhout, 2010: An indicator of the multiple equilibria regime of the Atlantic meridional overturning circulation. *J. Phys. Oceanogr.*, **40**, 551–567.
- Keller, H. B., 1977: Numerical solution of bifurcation and nonlinear eigenvalue problems. *Applications of Bifurcation Theory*, P. H. Rabinowitz, Ed., Academic Press, 359–384.
- Latif, M., E. Roeckner, U. Mikolajewicz, and R. Voss, 2000: Tropical stabilization of the thermohaline circulation in a greenhouse warming simulation. *J. Climate*, **13**, 1809–1813.
- Levitus, S., R. Burgett, and T. P. Boyer, 1994: *Salinity*. Vol. 3, *World Ocean Atlas 1994*, NOAA Atlas NESDIS 3, 99 pp.
- Meehl, G. A., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 747–846.
- Rahmstorf, S., and Coauthors, 2005: Thermohaline circulation hysteresis: A model intercomparison. *Geophys. Res. Lett.*, **32**, L23605, doi:10.1029/2005GL023655.
- Schmittner, A., M. Latif, and B. Schneider, 2005: Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations. *Geophys. Res. Lett.*, **32**, L23710, doi:10.1029/2005GL024368.
- Severijns, C. A., and W. Hazeleger, 2009: The efficient global primitive equation climate model SPEEDO. *Geosci. Model Dev. Discuss.*, **2**, 1115–1155.
- Stommel, H., 1961: Thermohaline convection with two stable regimes of flow. *Tellus*, **13**, 224–230.
- Stouffer, R. J., and Coauthors, 2006: Investigating the causes of the response of the thermohaline circulation to past and future climate changes. *J. Climate*, **19**, 1365–1387.
- Thual, O., and J. C. McWilliams, 1992: The catastrophe structure of thermohaline convection in a two-dimensional fluid model and a comparison with low-order box models. *Geophys. Astrophys. Fluid Dyn.*, **64**, 67–95.
- Trenberth, K. E., J. G. Olson, and W. G. Large, 1989: A global ocean wind stress climatology based on ECMWF analyses. National Center for Atmospheric Research Tech. Rep. NCAR/TN-338+STR, Boulder, CO, 93 pp.
- Tziperman, E., and H. Gildor, 2002: The stabilization of the thermohaline circulation by the temperature–precipitation feedback. *J. Phys. Oceanogr.*, **32**, 2707–2714.
- Vellinga, M., 1998: Multiple equilibria in ocean models as a side effect of convective adjustment. *J. Phys. Oceanogr.*, **28**, 621–633.
- , and R. A. Wood, 2002: Global climatic impacts of a collapse of the Atlantic thermohaline circulation. *Climatic Change*, **54**, 251–267.
- , and P. Wu, 2004: Low-latitude freshwater influence on centennial variability of the Atlantic thermohaline circulation. *J. Climate*, **17**, 4498–4511.
- , R. A. Wood, and J. M. Gregory, 2002: Processes governing the recovery of a perturbed thermohaline circulation in HadCM3. *J. Climate*, **15**, 764–780.
- Weijer, W., and H. A. Dijkstra, 2001: A bifurcation study of the three-dimensional thermohaline ocean circulation: The double hemispheric case. *J. Mar. Res.*, **59**, 599–631.
- , —, H. Oksuzoglu, F. W. Wubs, and A. C. de Niet, 2003: A fully-implicit model of the global ocean circulation. *J. Comput. Phys.*, **192**, 452–470.
- Wubs, F. W., A. C. de Niet, and H. A. Dijkstra, 2006: The performance of implicit ocean models on b- and c-grids. *J. Comput. Phys.*, **211**, 210–228.
- Yin, J., M. E. Schlesinger, N. G. Andronova, S. Malyshev, and B. Li, 2006: Is a shutdown of the thermohaline circulation irreversible? *J. Geophys. Res.*, **111**, D12104, doi:10.1029/2005JD006562.
- Zhang, G. J., V. Ramanathan, and M. J. McPhaden, 1995: Convection–evaporation feedback in the equatorial Pacific. *J. Climate*, **8**, 3040–3051.