

# Delayed Advective Oscillation of the Atlantic Thermohaline Circulation

Sang-Ki Lee<sup>1,2</sup> and Chunzai Wang<sup>2</sup>

<sup>1</sup>Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami FL

<sup>2</sup>Atlantic Oceanographic and Meteorological Laboratory, NOAA, Miami FL

Revised to Journal of Climate as a Note

September 2009

Corresponding author address: Dr. Sang-Ki Lee, NOAA/AOML, 4301 Rickenbacker Causeway,  
Miami, FL 33149, USA. E-mail: [Sang-Ki.Lee@noaa.gov](mailto:Sang-Ki.Lee@noaa.gov).

## **ABSTRACT**

A simple dynamic model is proposed to illustrate the multidecadal oscillation of the Atlantic thermohaline circulation. The proposed oscillation relies on alternating actions of positive and negative feedbacks, which are operated by a slow adjustment of the ocean circulation and the associated time delay in the advective flux response to a change in meridional density gradient. The key element of the oscillation is the time delay, which is conceptually related to the basin-crossing time of long Rossby waves in the high-latitude North Atlantic. For a sufficiently long time delay, the solution becomes unstable in some regions of model parameter space and oscillates with a period of about twice the delay time.

## 1. Introduction

The objective of this note is to propose a simple dynamic model for the multidecadal oscillation of the Atlantic thermohaline circulation (THC), which is commonly manifested in general circulation model simulations (e.g., Delworth et al. 1993; Dai et al. 2005; Knight et al. 2005). By definition, a self-sustained oscillation requires alternating actions of positive and negative feedbacks. Here, it is proposed that the required feedbacks for the multidecadal oscillation of the THC are provided by a slow adjustment of the ocean circulation and the associated time-delayed advective flux in response to a change in meridional density gradient. A four-box model is used to illustrate the proposed oscillation.

## 2. The four-box model

The North Atlantic Ocean is simplified to four boxes with two-layer structures in the high- and low-latitudes, as shown in Figure 1. The meridional density gradient is always positive, thus the volume transport must be northward in the upper layer and southward in the lower layer. Conservation of mass dictates that the volume transport at the mid-depth is downward in the high-latitude and upward in the low-latitude. Therefore, volume integration of the density conservation equation for each box yields

$$\frac{d\rho_1}{dt} = V(\rho_4 - \rho_1) - q - r\rho_1^3 - \frac{k_v}{H^2}(\rho_1 - \rho_4) + F_1, \quad (1)$$

$$\frac{d\rho_2}{dt} = V(\rho_1 - \rho_2) + q - r\rho_2^3 - \frac{k_v}{H^2}(\rho_2 - \rho_3) + F_2, \quad (2)$$

$$\frac{d\rho_3}{dt} = V(\rho_2 - \rho_3) - r\rho_3^3 + \frac{k_v}{H^2}(\rho_2 - \rho_3) + F_3, \quad (3)$$

$$\frac{d\rho_4}{dt} = V(\rho_3 - \rho_4) - r\rho_4^3 + \frac{k_v}{H^2}(\rho_1 - \rho_4) + F_4, \quad (4)$$

where  $\rho_1$ ,  $\rho_2$ ,  $\rho_3$  and  $\rho_4$  are densities of the upper low-latitude box, upper high-latitude box, lower high-latitude box, and lower low-latitude box, respectively;  $V$  is the volume transport (per unit volume) in response to the meridional density gradient;  $q$  is density flux into the upper high-latitude box (or out of the upper low-latitude box);  $r$  is a damping coefficient;  $k_v$  is a vertical diffusion coefficient;  $H$  is the model ocean depth divided by 2;  $F_1$ ,  $F_2$ ,  $F_3$  and  $F_4$  represent other forcing terms such as horizontal diffusion and convective mixing. Note that a cubic form of dissipation is used instead of a linear form to prevent infinite growth of linearly unstable solutions and model drifts as in Suarez and Schopf (1988) for the ENSO oscillator.

Now, a separate equation for  $V$  is required to solve the equations (1)-(4). Using the geostrophic balance and hydrostatic relation, the zonal baroclinic (i.e., upper minus lower layer) velocity in the mid-latitude corresponding to the north-south density gradient can be written as

$$\hat{u} = \frac{gH}{\rho_o f_o L_y} (\rho_{23} - \rho_{14}), \quad (5)$$

where  $g$  is the gravitational acceleration;  $\rho_o$  is a reference density;  $f_o$  is the planetary vorticity in the mid-latitude;  $L_y$  is the meridional length of the model domain divided by 2;  $\rho_{14} = (\rho_1 + \rho_4)/2$ , and  $\rho_{23} = (\rho_2 + \rho_3)/2$ . As discussed in Killworth (1985), the baroclinic meridional motions are established after the adjustment time, which depends on the basin-crossing time of long baroclinic Rossby wave. Therefore, in this study, the meridional volume transport (per unit volume) is represented by using the equation (5) with a time delay:

$$V = \alpha \frac{gH}{2\rho_o f_o L_y^2} [\rho_{23}(t - \delta) - \rho_{14}(t - \delta)], \quad (6)$$

where  $\alpha$  is a constant on the order of 1; and  $\delta$  is the time delay, which is about the basin-crossing time of long baroclinic Rossby waves.

Scaling time by  $2f_oL_y^2/(\hat{g}H)$ ,  $\rho$  by  $\sqrt{\hat{g}H/(2rf_oL_y^2)}$ , and  $q$  by  $\sqrt{r^{-1}[\hat{g}H/(2f_oL_y^2)]^3}$ , the

following non-dimensional equations can be derived:

$$\frac{d\rho_1}{dt} = \alpha[\rho_{23}(t-\delta) - \rho_{14}(t-\delta)](\rho_4 - \rho_1) - q - \rho_1^3 - k_o(\rho_1 - \rho_4) + F_1, \quad (7)$$

$$\frac{d\rho_2}{dt} = \alpha[\rho_{23}(t-\delta) - \rho_{14}(t-\delta)](\rho_1 - \rho_2) + q - \rho_2^3 - k_o(\rho_2 - \rho_3) + F_2, \quad (8)$$

$$\frac{d\rho_3}{dt} = \alpha[\rho_{23}(t-\delta) - \rho_{14}(t-\delta)](\rho_2 - \rho_3) - \rho_3^3 + k_o(\rho_2 - \rho_3) + F_3, \quad (9)$$

$$\frac{d\rho_4}{dt} = \alpha[\rho_{23}(t-\delta) - \rho_{14}(t-\delta)](\rho_3 - \rho_4) - \rho_4^3 + k_o(\rho_1 - \rho_4) + F_4, \quad (10)$$

where  $\hat{g} = g\Delta\rho/\rho_o$  ( $\Delta\rho$  is the scale of north-south density difference) and the non-dimensional vertical diffusivity  $k_o = 2k_v f_o L_y^2 / (\hat{g}H^3)$ .

### 3. The one-equation model

In order to gain some insights on the behavior of the nonlinear system of equations (7)-(10), a simplified one-equation model is derived and evaluated here. Specifically, it is assumed that  $\rho_2$ ,  $\rho_3$ , and  $\rho_4$  are close to each other enough to be represented as a single variable  $\rho_2^* = (\rho_2 + \rho_3 + \rho_4)/3$ . This is not an unreasonable assumption because the convective process maintains  $\rho_3$  as close to  $\rho_2$ , and the advective flux from the box-3 to box-4 also keeps  $\rho_4$  close to  $\rho_3$  for a sufficiently small value of  $k_o$ . Then, the density budget equations (7)-(10) can be simplified to a single equation:

$$\frac{d\hat{\rho}}{dt} = -\frac{2}{3}\alpha\hat{\rho}(t-\delta)\hat{\rho} + \frac{4}{3}q - \frac{4}{3}k_o\hat{\rho} - \hat{\rho}^3, \quad (11)$$

where  $\hat{\rho} = \rho_2^* - \rho_1$ , and the nonlinear dissipation term is slightly modified from its original form.

If  $\delta = 0$ , this equation has a positive steady state solution. By neglecting the nonlinear dissipation term for the sake of illustration, the positive steady solution  $\hat{\rho}_o$  is

$$\hat{\rho}_o = \frac{(-k_o + \sqrt{k_o^2 + 2\alpha q})}{\alpha}. \quad (12)$$

The linear stability of  $\hat{\rho}_o$  in the equation (11) can be studied by replacing  $\hat{\rho}(t)$  in (11) with the sum of the stationary solution  $\hat{\rho}_o$  and the perturbation  $\hat{\rho}'(t)$ . Retaining only the linear terms, the perturbation equation can be written as

$$\frac{d\hat{\rho}'}{dt} = -\frac{2}{3}(\alpha\hat{\rho}_o + 2k_o + 4.5\hat{\rho}_o^2)\hat{\rho}' - \frac{2}{3}\alpha\hat{\rho}_o\hat{\rho}'(t - \delta). \quad (13)$$

Seeking solutions of the form  $\hat{\rho}' \propto \exp(\sigma t)$  with  $\sigma = \sigma_r + i\sigma_i$ , the following equations can be derived:

$$\sigma_r = -\frac{2}{3}(\alpha\hat{\rho}_o + 2k_o + 4.5\hat{\rho}_o^2) - \frac{2}{3}\alpha\hat{\rho}_o e^{-\sigma_r\delta} \cos(\sigma_i\delta), \quad (14)$$

$$\sigma_i = \frac{2}{3}\alpha\hat{\rho}_o e^{-\sigma_r\delta} \sin(\sigma_i\delta). \quad (15)$$

If  $\sigma_r$  is positive, the first term in the RHS of (14) is always larger than the second term, thus the equation (14) cannot be satisfied. Therefore,  $\sigma_r$  is negative regardless of  $\sigma_i$ . This means that the simplified density budget equation (11) is always stable around  $\hat{\rho}_o$ . However, it is shown in the next section that the nonlinear system of equations (7)-(10) is unstable for a sufficiently large value of  $\delta$ . This suggests that the density budgets in the lower layers (box-3 and -4) must be fully incorporated to resolve the multidecadal THC oscillation.

#### 4. Numerical solutions

The behavior of the nonlinear equations (7) - (10) is explored numerically using a fourth-order Runge-Kutta scheme. When the upper layer is heavier than the layer below, the convective mixing is achieved by completely mixing the two layers. Horizontal diffusion is turned off for simplicity because its linear damping effect on the meridional density gradient does not add much value. Figure 2 shows three model solutions for  $\delta = 0, 7$  and  $20$ , when  $\alpha, q$  and  $k_o$  are set to  $2.0, 0.1$  and  $0.2$ , respectively. The dashed lines are the statistical equilibrium values. For small values of  $\delta$ , the solution achieves a stable state. For a larger value of  $\delta$ , on the other hand, the solution oscillates with a period of about twice the delay time (hereafter referred to as *delayed advective oscillator*).

Obviously, at issue is why the solution oscillates when the advective delay is sufficiently long. To answer this question, the equation for the density gradient between the high-latitude boxes and low-latitude boxes is diagnosed here:

$$\underbrace{\frac{d(\rho_{23} - \rho_{14})}{dt}}_{\text{Storage}} = \underbrace{\alpha[\rho_{23}(t - \delta) - \rho_{14}(t - \delta)](\rho_1 - \rho_3)}_{\text{Advection}} + \underbrace{q}_{\text{Surflux}} - \underbrace{\frac{1}{2}[(\rho_2^3 + \rho_3^3) - (\rho_1^3 + \rho_4^3)]}_{\text{Dissipation}}. \quad (16)$$

Figure 3 shows  $[\rho_{23} - \rho_{14}](t)$  and  $[\rho_{23} - \rho_{14}](t - \delta)$  in the upper panel. The storage term in the LHS of equation (16) and the advective density gradient flux, surface density flux gradient and dissipation terms in the RHS of equation (16) are shown in the lower panel. In this case,  $\alpha, q$  and  $k_o$  are set to  $2.0, 0.1$  and  $0.2$ , respectively, and the time delay  $\delta$  is set to  $20$ .

Table 1 summarizes the temporal evolutions of the meridional density gradient, the advective density gradient flux, and the sign of advective feedback during one cycle of the oscillation between the five points (A, B, C, D, and E) as indicated in Figure 3. As illustrated Table 1, the delayed advective oscillator is maintained by alternating actions of amplification (i.e., positive feedback) and stabilization (i.e., negative feedback) through the delayed advective density

gradient flux. In one cycle of the oscillation, there are two periods of amplification separated by two periods of stabilization.

The first amplification occurs when the meridional density gradient increases above the equilibrium point because the advective density gradient flux is smaller than the surface density flux gradient (A-B). This is followed by the first stabilization period, during which the meridional density gradient swings back from its maximum point toward the equilibrium point because the advective density gradient flux is larger than the surface density flux gradient (B-C). Then, the second amplification occurs, during which the meridional density gradient decreases below the equilibrium point because the advective density gradient flux is larger than the surface density flux gradient (C-D). During the second stabilization period, the meridional density gradient swings back from its minimum point toward the equilibrium point because the advective density gradient flux is smaller than the surface density flux gradient (D-E). This cycle then repeats. In summary, the meridional density gradient anomaly and advective density gradient flux anomaly are negatively (positively) correlated during the positive (negative) feedback periods as indicated in Table 1.

To have an idea about the actual time scale of the delayed advective oscillator, let us consider typical parameter values for the North Atlantic:  $L_x = 5 \times 10^6 \text{m}$ ;  $L_y = 2.0 \times 10^6 \text{m}$ ;  $\hat{g} = 10^{-2} \text{m s}^{-1}$ ;  $f_o = 10^{-4} \text{s}^{-1}$ ;  $\beta = 2 \times 10^{-11} \text{m}^{-1} \text{s}^{-1}$ ; and  $H = 2000 \text{m}$ . According to previous general circulation model simulations, the density fluctuation associated with the multidecadal oscillation of THC is mainly in the high-latitude North Atlantic around  $40^\circ\text{N} \sim 65^\circ\text{N}$  (e.g., Delworth et al. 1993) and propagates slowly to the west (e.g., Dijkstra et al. 2006). Therefore, the long baroclinic Rossby wave speed  $c$  ( $= g'H\beta/(2f_o^2)$ ; where  $g'$  is the reduced gravity) is computed separately based on typical parameter values for the high-latitude North Atlantic around  $50^\circ\text{N}$ :  $g' = 5 \times 10^{-3} \text{m s}^{-1}$ ; and

$f_o = 1.1 \times 10^{-4} \text{ s}^{-1}$ . Then, the basin-crossing time can be estimate by  $L_x/c$ , which is about 20 years in this case. Assuming that the delay time is on the order of the basin-crossing time, the period of delayed advective oscillator is about 40 years, which is twice the delay time. Using the time scale,  $2f_o L_y^2 / (\hat{g}H) = 1.3 \text{ years}$ , the nondimensional delay time  $\delta$  is approximately 15.

## 5. Impacts of external forcing

Now, the responses of delayed advective oscillator to low frequency external forcing patterns, such as freshwater flux into the high-latitude North Atlantic, and to high frequency external forcing patterns, such as weather noise related to the North Atlantic Oscillation (NAO), are explored here. Three low-frequency forcing experiments are performed, i.e., (a) density flux out of the high-latitude North Atlantic (i.e., freshening or warming), (b) complete shutdown of the THC, and (c) density flux into the high-latitude North Atlantic (i.e., cooling or net evaporation). For each experiment, the following form of density flux is used only in the high-latitude (i.e., box-2):

$$q_2 = q + q_o \exp\left(\frac{t^* - 40}{20}\right)^2, \quad (17)$$

where  $t^* = t/\delta$ , and  $q_o$  is set to  $-q/2$ ,  $-2q$ , and  $q/2$  for experiments (a), (b) and (c), respectively.

Note that the surface flux in the low-latitude is kept constant in these experiments.

Figure 4a shows that the THC may slow down due to the density flux out of the high-latitude (i.e., freshening or warming), but it swings back to the equilibrium solution once the external forcing is removed. An interesting feature is that the amplitude of the delayed advective oscillator is reduced substantially and its recovery is extremely slow. As shown in Figure 4b, the THC completely shuts down as the surface flux in the high-latitude box is reduced to match that

in the low-latitude box. The THC swings back to the equilibrium solution after the external forcing is removed. However, the delayed advective oscillator is completely disrupted. Interestingly, the external density flux out of the high-latitude North Atlantic (i.e., cooling or net evaporation) has only a minor impact on the THC strength as shown in Figure 4c. However, the delayed advective oscillation is weakened (but not as much as in Figure 4a) and slowly recovers once the external forcing is removed.

Next, the impact of high frequency forcing on the delayed advective oscillator is explored. It is widely believed that the high frequency portion of the NAO originates from weather noise. Nevertheless, the NAO has a coherent spatial structure with a dipole-like meridional pattern of the sea level pressure (Hurrell 1995). Due to this coherent spatial pattern, if the high-latitude is cooled (warmed), the mid-latitude is warmed (cooled) during a positive (negative) phase of the NAO. Therefore, the high frequency forcing of the NAO is represented here as a random noise surface flux with anti-symmetric meridional pattern; that is, the sign of random forcing is opposite in the two latitude boxes but with the same amplitude. Note that the random noise forcing does not produce a net surface flux into or out of the system. The amplitude of the random forcing is set to  $q/2$ .

Figure 5 shows the model solutions (a) with and (b) without the random forcing, and (c) with both the low- and high-frequency forcing for  $\alpha=1.2$ ,  $q=0.1$ ,  $k_o=0.2$ , and  $\delta=20$ . In the last case, the low frequency forcing is added only in the high-latitude (i.e., box-2) using the equation (17) with  $q_o = -1.5q$ . As shown in Figure 5b, the oscillation is damped out without the weather noise for the given parameter values. Interestingly, if the weather noise is introduced, the delayed advective oscillator with the period of  $\sim 2\delta$  can sustain its amplitude of up to about 35% of the mean. It is also interesting to note that, under the weather noise forcing, the amplitude of the

oscillation fluctuates at a very low-frequency, which amounts to the multicentennial time scale using realistic parameter values for the North Atlantic. Figure 5c shows that even when the external forcing is large enough to nearly shutdown the THC, the weather noise can invigorate the delayed advective oscillator once the external forcing is removed.

In summary, the THC is remarkably stable because it always swings back to its original state once external forcing is removed (this conclusion may not be valid if both the temperature and salinity are considered as in Stommel 1961). The delayed advective oscillator is, on the other hand, very fragile. If external forcing is large enough, it can virtually wipe out the delayed advective oscillator. The recovery of the delayed advective oscillator is extremely sluggish, suggesting that the growth rate of the delayed advective oscillator is very small. However, these are characteristics of the delayed advective oscillator in its pure form. If the NAO-like weather noise forcing is added to the system, the behavior of the delayed advective oscillator is drastically changed. In particular, relatively large amplitude of the weather noise (50% of the mean is used in the experiment) can sustain an active delayed advective oscillation of an otherwise stable system. The THC can still shut down if external forcing is large enough. However, the weather noise can quickly invigorate the delayed advective oscillator once the external forcing is removed. Finally the weather noise can also produce a very low-frequency fluctuation of the delayed advective oscillation at the multicentennial time scale.

An important question is why the delayed advective oscillator is excited by the NAO-like weather noise. The simple stochastic climate model of Hesselmann (1976) provides a plausible explanation for this question. It is well known after Hesselmann (1976) that a random noise atmospheric forcing produces a red noise spectrum of ocean temperature via ocean memory. If this theory is applied to the four-box model, it means that a random surface forcing can produce

large amplitude signals in the meridional density gradient field,  $(\rho_{23} - \rho_{14})$ , at low frequencies including at the frequency of delayed advective oscillator,  $\omega \sim 0.5\delta^1$ . Therefore, the delayed advective oscillation can be excited and maintained by the weather noise even if it is subject to a damped oscillation as shown in Figure 5a.

## 6. Linear stability analysis

To better understand how the four parameters,  $\alpha$ ,  $q$ ,  $k_o$ , and  $\delta$  influence the delayed advective oscillator, a linear stability analysis of the nonlinear system of equations (7)-(10) is performed. First, the stationary solutions with  $\delta=0$  are obtained by numerically integrating the equations (7)-(10). Replacing the solutions with the sum of stationary solution and perturbation, and retaining only the linear terms can derive the perturbation equations (not shown). Seeking solutions of the form  $\rho'_k = c_k \exp(\sigma t)$  with  $k=1, 2, 3$  and  $4$ , and  $\sigma = \sigma_r + i\sigma_i$ , a matrix equation,  $\mathbf{A} \cdot \mathbf{C} = 0$ , can be derived. The determinant of the matrix  $\mathbf{A}$  must vanish in order for nontrivial eigenfunctions to exist: this yields an equation for the calculation of the complex eigenvalue  $\sigma$  for chosen values of  $\alpha$ ,  $\delta$ ,  $k_o$  and  $q$ . Since the matrix  $\mathbf{A}$  contains the eigenvalue  $\sigma$  and its exponential form,  $\exp(-\sigma\delta)$ , an iterative Muller's method is used to obtain the eigenvalue  $\sigma$ .

Figure 6 shows the neutral curves (a) on the  $\alpha$ - $\delta$  plane for  $q=0.1$  and  $k_o=0.2$ , (b) on the  $k_o$ - $\delta$  plane for  $\alpha=2.0$  and  $q=0.1$ , and (c) on the  $q$ - $\delta$  plane for  $\alpha=2.0$  and  $k_o=0.2$ . For a given value of  $\delta$ , increasing  $\alpha$  and decreasing  $k_o$  destabilize the system. These results are not surprising since  $\alpha$  is proportional to the meridional volume transport, which provides the positive feedback required to maintain the delayed advective oscillator, thus serves as a growth rate, and  $k_o$  serves as a damping rate. However, it is important to note that the neutral curve on the  $q$ - $\delta$  plane is not monotonic. For a given value of  $\delta$ , the system is unstable only when  $q$  is within a certain range.

This strongly supports an idea that the delayed advective oscillator exists via a delicate balance of various terms, and also nicely explains why the delayed advective oscillator (in its pure form) is so fragile under the impacts of external forcing as illustrated in Figure 4.

## **7. Summary and Discussions**

Perhaps, the four-box model presented here is one of the simplest dynamic models for the multidecadal oscillation of the THC, which is commonly manifested in general circulation model simulations. Despite the overly simplified nature of the model, this minimal complexity model describes effectively the mechanism of the delayed advective oscillator, which appears to be an important stepping-stone toward our understanding of the THC and its multidecadal oscillation.

The key element of the delayed advective oscillator is the time delay in the advective flux response to a change in meridional density gradient. This time delay, which is conceptually related to the basin-crossing time of long baroclinic Rossby waves at the high-latitude North Atlantic, allows alternating actions of positive and negative advective feedbacks and thus gives rise to a self-sustained oscillation. An important signature (or indicator) of the delayed advective oscillator is that the meridional density gradient anomaly and meridional advective density gradient flux anomaly are negatively (positively) correlated during the positive (negative) feedback periods.

Related to the fundamental issue of decadal predictability of the THC, an apparently important and practical question is whether the multidecadal oscillation of the THC is self-sustained (e.g., Greatbatch and Zhang 1995) or damped, in which case the NAO-like weather noise may sustain the oscillation (e.g., Capotondi and Holland 1997; Eden and Greatbatch 2003). The linear stability analysis of the delayed advective oscillator may help us better understand

under what conditions (or regions of the model parameter space for  $\alpha$ ,  $k_o$ ,  $q$  and  $\delta$ ) the THC switches between an unstable state to a damped regime.

The delayed advective oscillator can be compared to the delayed action oscillator for ENSO (Suarez and Schopf 1988) because the key element in both cases is the oceanic Rossby wave transit effects. However, there is an important distinction between the two oscillators. In the delayed action oscillator, the equatorial oceanic Rossby wave transit effects provide a negative delayed feedback to an otherwise linearly amplifying system. In the delayed advective oscillator, the advective density gradient flux, which is delayed by the oceanic Rossby wave transit effects, provides both the negative and positive feedbacks for the oscillation.

Finally, it is important to point out that, for a given positive (negative) anomaly of meridional density gradient, if there is no time delay, the meridional salt advection has a destabilizing effect whereas the meridional heat transport has a stabilizing effect. Griffies and Tziperman (1995) showed that a phase-lag between the salt and heat advection could support a damped oscillation. This mechanism, which is obviously missing in our four-box model, can be readily explored by expanding the density equations (7)-(10) to two sets of equations for salinity and temperature equations.

**Acknowledgements.** We wish to thank R. Greatbatch, Y.-G. Park, and two anonymous reviewers for insightful comments. This work was supported by a grant from the National Oceanic and Atmospheric Administration's Climate Program Office and by a grant from the National Science Foundation (ATM-0850897). The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the funding agency.

## References

- Capotondi, A., and W. R. Holland, 1997: Decadal variability in an idealized ocean model and its sensitivity to surface boundary conditions. *J. Phys. Oceanogr.*, 27, 1072-1093.
- Dai, A., A. Hu, G. A. Meehl, W. M. Washington, and W. G. Strand, 2005: Atlantic thermohaline circulation in a coupled general circulation model: unforced variations versus forced changes. *J. Climate*, 18, 3270–3293.
- Delworth, T. L., S. Manabe, and R. Stouffer, 1993: Interdecadal Variations of the Thermohaline Circulation in a Coupled Ocean-Atmosphere Model. *J. Climate*, 6, 1993–201.
- Dijkstra, H. A., L. te Raa, M. Schmeits, and J. Gerrits, 2006: On the physics of the Atlantic Multidecadal Oscillation. *Climate Dyn.*, 56, 36-50, doi-10.1007/s10236-005-0043-0.
- Eden C., and R. J. Greatbatch, 2003: A damped decadal oscillation in the North Atlantic climate system, *J. Climate*, 16, 4043-4060.
- Greatbatch R. J., and S. Zhang, 1995: An interdecadal oscillation in an idealized ocean basin forced by constant heat flux. *J. Climate*, 8, 81-91.
- Griffies, S. M., and E. Tziperman, 1995: A linear thermohaline oscillator driven by stochastic atmospheric forcing. *J. Climate*, 8, 2440–2453.
- Hasselmann, K., 1976: Stochastic climate models, Part I, Theory. *Tellus*, 28, 473-485.
- Hurrell, J. W, 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation, *Science*, 269, 676–679.
- Killworth, P. D., 1985: A two-level wind and buoyancy driven thermocline model. *J. Phys. Oceanogr.* **15**, 1414-1432.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann, 2005: A signature of persistent natural thermohaline circulation cycles in observed climate, *Geophys. Res. Lett.*, 32, L20708, doi:10.1029/2005GL024233.

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

Suarez, M. J., and P. S. Schopf, 1988: A delayed action oscillator for ENSO. *J. Atmos. Sci.*, 3283-3287.

Table 1. The meridional density gradient ( $[\rho_{23}-\rho_{14}](t)$ ), the advective density gradient flux and the sign of advective feedback during the periods between the five points (A, B, C, D, and E) as indicated in Figure 3. “ $\uparrow$ ” and “ $\downarrow$ ” indicate that the value is above and below the equilibrium solution, respectively. The second and third columns are negatively correlated during (+) feedback periods, and positively correlated during (-) feedback periods.

Time	$[\rho_{23}-\rho_{14}](t)$	-Advection	Feedback
(A) - (B)	$\uparrow$	$\downarrow$	(+)
(B) - (C)	$\uparrow$	$\uparrow$	(-)
(C) - (D)	$\downarrow$	$\uparrow$	(+)
(D) - (E)	$\downarrow$	$\downarrow$	(-)

Figure 1. The North Atlantic Ocean is simplified with four boxes. Since the meridional density gradient is always positive, the volume transport (per unit volume)  $V$  is always northward in the upper layer and southward in the lower layer.  $\rho_1$ ,  $\rho_2$ ,  $\rho_3$  and  $\rho_4$  are densities of the upper low-latitude box, upper high-latitude box, lower high-latitude box, and lower low-latitude box, respectively.  $q$  is density flux into the upper high-latitude box (or out of the upper low-latitude box).

Figure 2. Three model solutions,  $(\rho_{23} - \rho_{14})$ , for (a)  $\delta = 0$ , (b)  $\delta = 7$  and (c)  $\delta = 20$ .  $\alpha$ ,  $q$  and  $k_o$  are set to 2.0, 0.1, and 0.2, respectively. The dashed lines are the statistical equilibrium values.

Figure 3. The model solutions  $[\rho_{23} - \rho_{14}](t)$  and  $[\rho_{23} - \rho_{14}](t - \delta)$  are shown in the upper panel. The storage term in the LHS of equation (16) and the advective density gradient flux, surface flux gradient and dissipation terms in the RHS of equation (16) are shown in the lower panel.  $\delta$ ,  $\alpha$ ,  $q$  and  $k_o$  are set to 20, 2.0, 0.1, and 0.2, respectively. The symbols (+) and (−) represent the period of positive and negative feedback, respectively. See text and Table 1 for the description of density budgets between the points, (A), (B), (C), (D) and (E).

Figure 4. The responses of the delayed advective oscillator to the external forcing of (a) density flux out of the high-latitude North Atlantic (i.e., freshening or warming), (b) complete shutdown of the THC, and (c) density flux into the high-latitude North Atlantic (i.e., cooling or net evaporation). The amplitude of the external forcing  $q_o$  is set to  $-q/2$ ,  $-2q$ , and  $q/2$  for (a), (b) and (c), respectively.  $\delta$ ,  $\alpha$ ,  $q$  and  $k_o$  are set to 20, 2.0, 0.1, and 0.2, respectively.

Figure 5. The behavior of delayed advective oscillator (a) with and (b) without the random noise forcing, and for (c) a case with both the low- and high-frequency forcing. The amplitude of the random forcing is set to  $q/2$  for (a) and (c). For (c),  $q_o = -1.5q$ .  $\delta$ ,  $\alpha$ ,  $q$  and  $k_o$  are set to 20, 1.2, 0.1, and 0.2, respectively.

Figure 6. Neutral stability curves (a) on the  $\alpha$ - $\delta$  plane for  $q=0.1$  and  $k_o=0.2$ , (b) on the  $k_o$ - $\delta$  plane for  $\alpha=2.0$  and  $q=0.1$ , and (c) on the  $q$ - $\delta$  plane for  $\alpha=2.0$  and  $k_o=0.2$ .

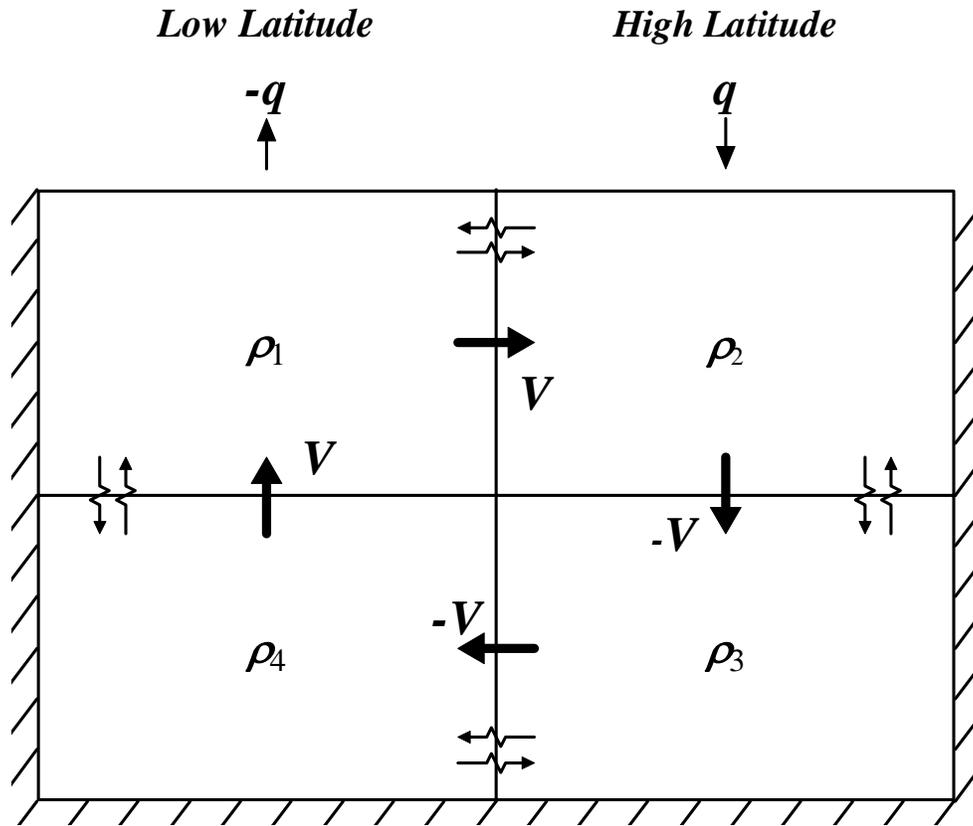


Figure 1. The North Atlantic Ocean is simplified with four boxes. Since the meridional density gradient is always positive, the volume transport (per unit volume)  $V$  is always northward in the upper layer and southward in the lower layer.  $\rho_1$ ,  $\rho_2$ ,  $\rho_3$  and  $\rho_4$  are densities of the upper low-latitude box, upper high-latitude box, lower high-latitude box, and lower low-latitude box, respectively.  $q$  is density flux into the upper high-latitude box (or out of the upper low-latitude box).

### Delayed Advective Oscillator

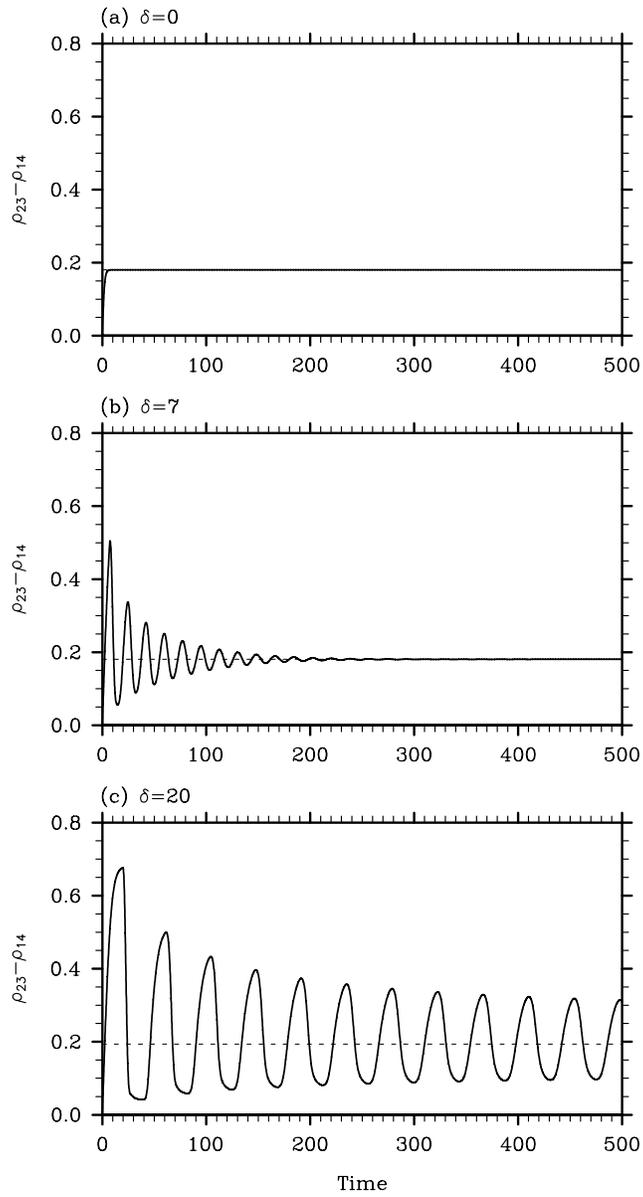


Figure 2. Three model solutions,  $(\rho_{23} - \rho_{14})$ , for (a)  $\delta=0$ , (b)  $\delta=7$  and (c)  $\delta=20$ .  $\alpha$ ,  $q$  and  $k_o$  are set to 2.0, 0.1, and 0.2, respectively. The dashed lines are the statistical equilibrium values.

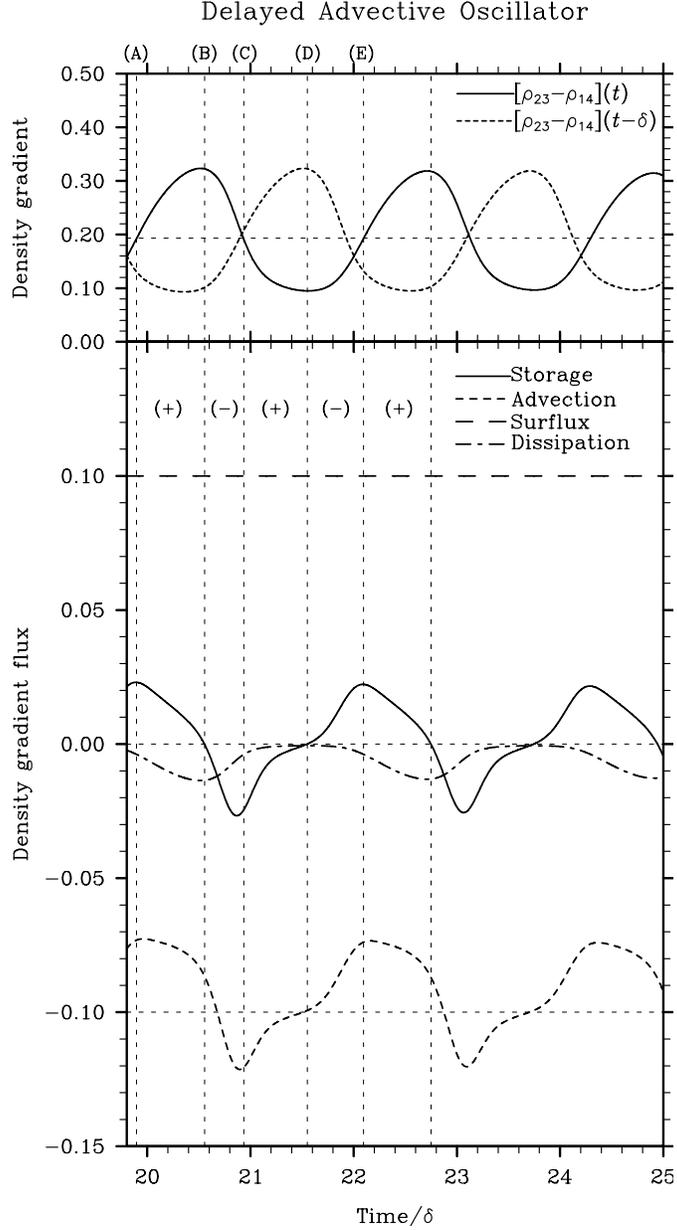


Figure 3. The model solutions  $[\rho_{23} - \rho_{14}](t)$  and  $[\rho_{23} - \rho_{14}](t - \delta)$  are shown in the upper panel. The storage term in the LHS of equation (16) and the advective flux divergence, surface flux and dissipation terms in the RHS of equation (16) are shown in the lower panel.  $\delta$ ,  $\alpha$ ,  $q$  and  $k_o$  are set to 20, 2.0, 0.1, and 0.2, respectively. The symbols (+) and (-) represent the period of positive and negative feedback, respectively. See text and Table 1 for the description of density budgets between the points, (A), (B), (C), (D) and (E).

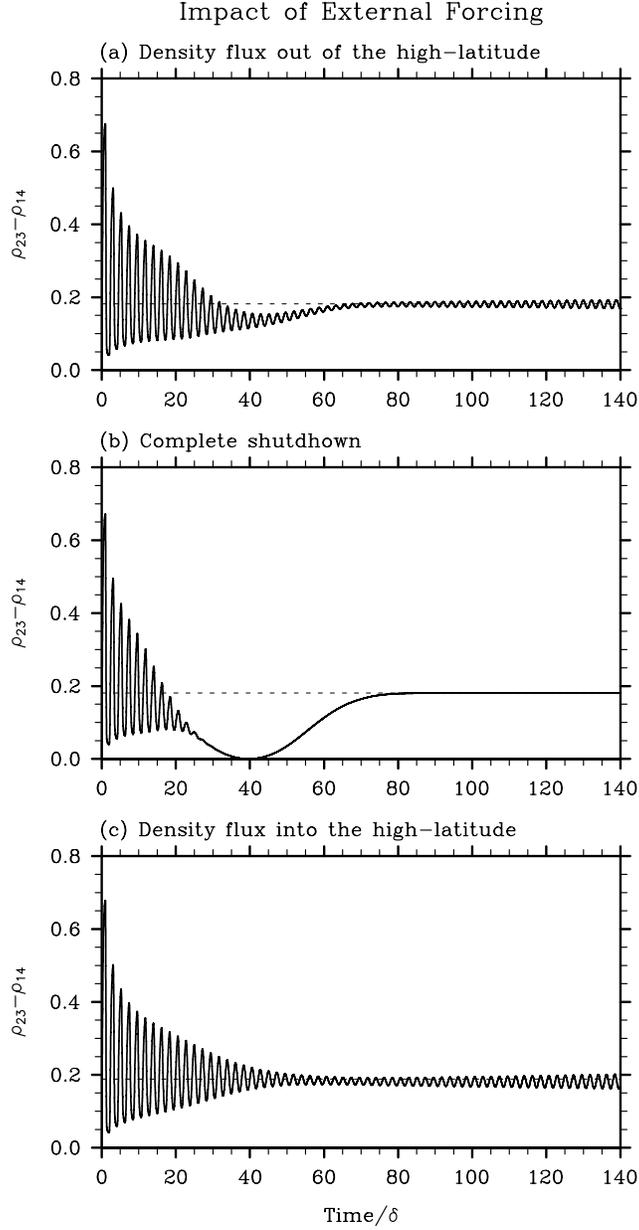


Figure 4. The responses of the delayed advective oscillator to the external forcing of (a) density flux out of the high-latitude North Atlantic (i.e., freshening or warming), (b) complete shutdown of THC, and (c) density flux into the high-latitude North Atlantic (i.e., cooling or net evaporation). The amplitude of the external forcing  $q_o$  is set to  $-q/2$ ,  $-2q$ , and  $q/2$  for (a), (b) and (c), respectively.  $\delta$ ,  $\alpha$ ,  $q$  and  $k_o$  are set to 20, 2.0, 0.1, and 0.2, respectively.

### Impact of Weather Noise

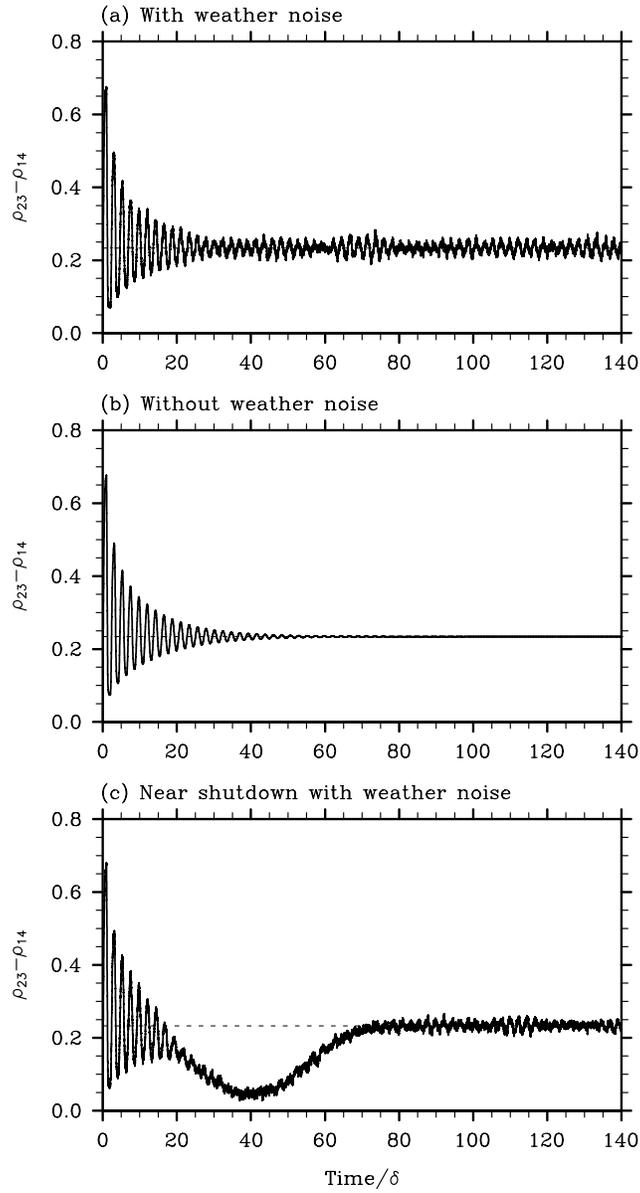


Figure 5. The behavior of delayed advective oscillator (a) with and (b) without the random noise forcing, and for (c) a case with both the low- and high-frequency forcing. The amplitude of the random forcing is set to  $q/2$  for (a) and (c). For (c),  $q_o = -1.5q$ .  $\delta$ ,  $\alpha$ ,  $q$  and  $k_o$  are set to 20, 1.2, 0.1, and 0.2, respectively.

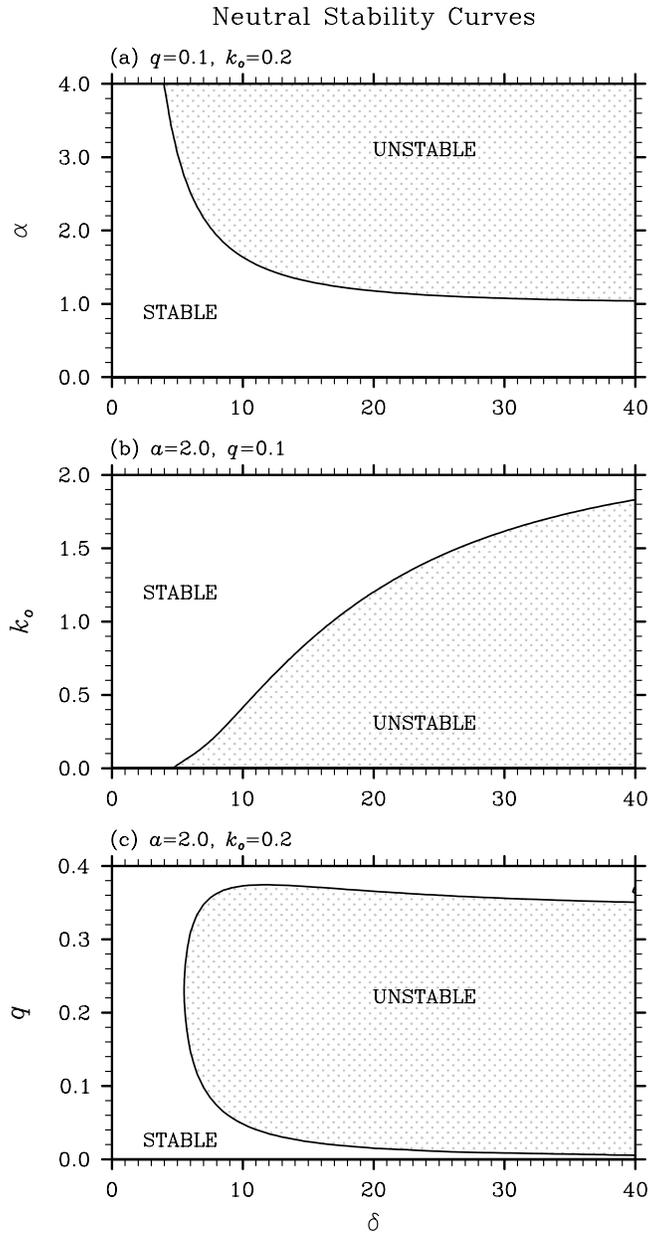


Figure 6. Neutral stability curves (a) on the  $\alpha$ - $\delta$  plane for  $q=0.1$  and  $k_o=0.2$ , (b) on the  $k_o$ - $\delta$  plane for  $\alpha=2.0$  and  $q=0.1$ , and (c) on the  $q$ - $\delta$  plane for  $\alpha=2.0$  and  $k_o=0.2$ .