A Theory for Self-Sustained Multi-Centennial Oscillation of the Atlantic Meridional Overturning Circulation

Yang Li\(^1\), and Haijun Yang\(*^{2,3}\)

\(^1\)Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, 100871, China.

\(^2\)Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Science and CMA-FDU Joint Laboratory of Marine Meteorology, Fudan University, Shanghai, 200438, China.

\(^3\)Shanghai Scientific Frontier Base for Ocean-Atmosphere Interaction Studies, Fudan University, Shanghai 200438, China.

Journal of Climate

1\textsuperscript{st} submission on September 1, 2021

1\textsuperscript{st} revision on January 15, 2022

2\textsuperscript{nd} revision on May 4, 2022

*Corresponding author address: Haijun Yang, Department of Atmospheric and Oceanic Sciences, Fudan University, 2005 Songhu Road, Shanghai China, 200438.

Email: yanghj@fudan.edu.cn
Abstract

In this work, a single-hemisphere 4-box model is used to study the low-frequency variability of the Atlantic meridional overturning circulation (AMOC). We introduce an enhanced mixing mechanism in the subpolar ocean to balance the positive salinity advection feedback, so that the AMOC in the 4-box model exhibits a self-sustained multi-centennial oscillation. The enhanced mixing mechanism is proposed based on results from a coupled climate model, which show the eddy-induced mixing or diffusion in the subpolar ocean is always enhanced when the AMOC anomaly is large, namely, the enhancement is due to weak stratification when the AMOC is strong, and is due to meso- and submeso-scale eddies when the AMOC is weak. Without the enhanced mixing, the 4-box model system can be either stable or unstable, but cannot realize a self-sustained stable oscillation. With the enhanced mixing, the 4-box model can be interpreted approximately as a reduced 3-box model, so that the theoretical solution to the multi-centennial oscillation can be obtained. The oscillation period is determined by the eigenvalue of the system, which is fundamentally controlled by the turnover time of the upper ocean. We also illustrate that the multi-centennial oscillation can be excited by stochastic freshwater forcing. This study suggests that the Atlantic Ocean has an intrinsic multi-centennial mode, which may help us understand this class of variability identified in paleoclimatic proxy data.

Keywords: AMOC, Self-sustained oscillation, Multi-centennial oscillation, Enhanced mixing, Box model
1. Introduction

The Earth's climate system has multi-centennial timescale variability, evidenced by paleoclimatic reconstruction data from tree rings, ice core, sediments in lakes and oceans as well as historical documents from Europe, Asia, and America (e.g., Sirén 1961; Soutar and Isaacs 1969; Johnsen et al. 1970; LaMarche 1974; Fisher 1982; Gajewski 1988; Stuiver and Braziunas 1989; Stocker and Mysak 1992; Chapman and Shackleton 2000; Risebrobakken et al. 2003; Zheng et al. 2010). Multi-centennial variability has been observed in land precipitation, ocean temperature, salinity and currents. Particularly, many proxies obtained in the North Atlantic region are thought to be associated with historical events such as the Medieval Climatic Anomaly (MCA) and the Little Ice Age (LIA) (e.g., Cronin et al. 2010; Sicre et al. 2011; Miettinen et al. 2015; Moffa-Sánchez et al. 2015), which, in turn, suggest multi-centennial variability of the Atlantic meridional overturning circulation (AMOC). In this paper, multi-centennial timescale refers to the timescale of about 500±300 years. Centennial variability (about 100 years) and millennial variability are not our focuses.

The origin and mechanism of multi-centennial variability have been challenging topics for scientists, due to the lack of observational data and complexities of internal processes, external forcing and their interactions. For example, although many coupled model simulations showed that solar irradiance dominates internal variabilities at multi-decadal and longer timescales and at large spatial scales (Cubasch et al. 1997; Drijfhout et al. 1999; Rind et al. 1999), studies also suggested that at the multi-centennial timescale the internal feedback processes, especially the delayed advection-salinity feedback, seem to play an important role, and the solar irradiance acts more as an initial stimulation (Weber et al. 2004). More specifically, many studies supported the idea that the LIA was triggered by external forcing, such as solar irradiance variation and volcanic cluster (Berner et al. 2011; Miller et al. 2012). However, results from coupled model sensitivity experiments suggested that volcanic forcing is not a prerequisite (Moreno-Chamarro et al. 2017). Modeling studies suggested that
the forced signal only accounts for 10%−35% of the total variance of the signals longer than 10 years in the past two millennia (Moffa-Sánchez et al. 2019).

Assessing the relative roles of external forcing and internal processes in multi-centennial variability is critical to unravel the mystery of historical events such as the MCA and LIA. Above all, an in-depth investigation of multi-centennial internal variability is needed first. This also helps climate prediction, as the magnitude of global warming over the last 6-7 decades may have been damped by internal variability from the Atlantic Ocean (Bonnet et al. 2021). The ocean is believed to govern the multi-centennial variability, due to the integration of short-term atmosphere fluctuations and its internal dynamics related to the thermohaline circulation (THC) or the AMOC (Mikolajewicz and Maier-Reimer 1990; Msadek and Frankignoul 2008). Besides, a better understanding and representation of the multi-centennial AMOC variability may help improve the decadal predictability of Earth’s climate (Collins and Sinha 2003).

There have been many studies on the multi-centennial mode of the climate system, using a hierarchy of models of different complexities, ranging from simple energy balance models to coupled general circulation models (GCMs) (e.g., Welander 1982; Mikolajewicz and Maier-Reimer 1990; Mysak et al. 1993; Sévellec et al. 2006; Sévellec and Fedorov 2014; Jiang et al. 2021). In an earlier study using an ocean GCM (OGCM), Mikolajewicz and Maier-Reimer (1990) identified a 320-year mode, which stands out most clearly in the strength of the Antarctic circumpolar current, and its period agrees well with the overall flushing time of the Atlantic Ocean. This mode is attributed to salinity anomaly advection by the AMOC, driven by the deep-water formation in the Weddell Sea. Mysak et al. (1993) found a similar mode under random freshwater flux in a 2-dimensional ocean model representing the Atlantic Ocean; they also emphasized the importance of the transport of
salinity anomalies. In this paper, we are mainly interested in the multi-centennial mode of the AMOC in simple box models, because we aim for fundamental understanding via theoretical study.

In modeling studies using simple models, low-frequency oscillations can emerge from either external forcing or internal processes. The former can be classified as the forced oscillatory, in which the low-frequency variability is attributed to the response of the AMOC to stochastic atmosphere forcing (Griffies and Tziperman 1995; Roebber 1995). Both the spatial pattern and spectrum of atmospheric forcing can influence the period of AMOC oscillation (Tziperman and Ioannou 2002). The latter is classified as the self-sustained oscillation, in which the ocean itself can generate self-sustained oscillation (e.g., Rivin and Tziperman 1997; hereafter RT97). In fact, in the earlier study of Welander (1982), a self-sustained oscillation was realized by introducing a flip-flop local convection mechanism, which was then applied to describe millennial thermohaline oscillation (e.g., Zhang et al. 2002; Colin de Verdière et al. 2006). While these studies concentrated on multiple equilibrium states, they reminded us that local vertical density profile may influence the large-scale overturning circulation. Salinity advection feedback can also be an energy source; for example, a strong AMOC will make the subpolar ocean more saline and thus denser, which in turn leads to a stronger AMOC. In RT97, a self-sustained oscillation was identified after introducing a weakly nonlinear velocity closure, assuming that the large-scale overturning circulation will become insensitive to the large-scale meridional density gradient when perturbation is large. In addition, self-sustained oscillation can also be realized in a one-dimensional Howard-Malkus loop model (Welander 1986), when considering temperature-salinity advection.

In this paper, we demonstrate that a combination of salinity advection and “enhanced mixing” can lead to a self-sustained multi-centennial oscillation in a 4-box model. The enhanced mixing mechanism is parameterized in our box model, which is based on the results from a coupled model.
control run. We show that in a coupled model, the diffusion or mixing in the subpolar Atlantic can be enhanced due to the weak stratification when the AMOC is strong, or due to the active meso- and submeso-scale eddies when the AMOC is weak. The enhanced mixing can restrain the positive salinity advection feedback, which can eventually lead to a self-sustained stable oscillation. Linear stability analyses further reveal that the oscillation period is determined by the eigenvalue of the system, which is controlled by the turnover time of the upper ocean. Without the enhanced mixing, the AMOC in the 4-box model is unsteady, exhibiting an oscillation with either growing or decaying amplitude.

This work differs from previous studies in the following aspects. First, only salinity equations are employed in our model for simplicity, and density perturbation is linearly determined by salinity perturbation. Second, we focus on the multi-centennial timescale, different from the decadal to centennial timescales in RT97 and millennial timescale in Colin de Verdière et al. (2006). Third, only small oscillation of the AMOC around its stable equilibrium state is studied, and the step- or delta-function-like oscillations (Huang 1994; Sakai and Peltier 1999; Zhang et al. 2002; Sévellec and Fedorov 2014) are not our concern. The latter is more related to the climate regime shift.

The paper is organized as follows. In Section 2, the multi-centennial oscillation of the AMOC from a coupled model control run is analyzed briefly. In Section 3, the 4-box model is introduced, and eigenvalues of the linear system are obtained. In Section 4, enhanced mixing is parameterized in the salinity equation, based on the coupled model results presented in Section 2. A self-sustained oscillation in the box model is exhibited, and its existence is explained theoretically. In Section 5, effects of model geometry, external stochastic forcing, and nonlinear salinity advection on the multi-centennial variability are discussed, followed by a summary in Section 6.
2. Multi-centennial oscillation in CESM

In this section, we investigate the low-frequency variability of the AMOC in a coupled climate model. Temporal evolutions of vertical temperature, salinity, and density structures in the North Atlantic are examined. Salinity diffusion and mixing are calculated. Analyses of the coupled model results provide the rationale for the assumptions used in the simple box model in Section 3.

A long simulation was carried out using the Community Earth System Model (CESM, version 1.0) of the National Centre for Atmospheric Research (NCAR). The ocean horizontal grid has a uniform 1.125° spacing in the zonal direction, and non-uniform spacing in the meridional direction, whose resolution is 0.27° near the equator, extending to the maximum 0.65° at 60°N/S and then shrinking gradually to the poles. Detailed description can be found in Yang et al. (2015). The control run starts from the rest with the standard configuration for the preindustrial condition, and is integrated for 2500 years. The outputs during the last 1500 years are used for analysis. In the CESM, the velocity can be divided into Eulerian-mean and eddy-induced components; the latter consists of bolus velocity and sub-mesoscale velocity (Gent and McWilliams 1990; Fox-Kemper and Ferrari 2008; Fox-Kemper et al. 2008, 2011). Each velocity component has its corresponding mass transport (i.e., streamfunction).

A multi-centennial oscillation of the AMOC is identified in the CESM control run. Figure 1 shows the AMOC index and its spatial pattern. The AMOC index is defined as the maximum of the residual streamfunction in the region of 20°–70°N between 200 and 3000 m in the Atlantic. The lowpass-filtered AMOC index exhibits clearly a multi-centennial oscillation with a period between 300 and 400 years (Fig. 1a). The climatological mean of the total AMOC is about 24 Sv, and the low-frequency variation is about 10% of the mean value. Figures 1b and c show the climatological mean patterns of Eulerian-mean AMOC and eddy-induced AMOC, respectively. The Eulerian-mean
AMOC has the maximum strength of about 24 Sv around 40°N at the depth of ~1000 m. The eddy-induced AMOC has the maximum strength of about 8 Sv between 35° and 70°N at a shallower depth of 100–500 m, corresponding to the North Atlantic Deep Water formation region with strong meridional density gradient and deep convection (Yang et al. 2015).

The low-frequency Eulerian-mean AMOC shows a meridional coherent change in the whole Atlantic basin (Fig. 1d). The low-frequency eddy-induced AMOC, however, has remarkable change only locally in the subpolar Atlantic (Fig. 1e), revealed by their regression patterns on the mean-normalized AMOC index. Furthermore, the change in eddy-induced AMOC has a comparable magnitude but opposite sign in the subpolar deep convection region to the change in Eulerian-mean AMOC. This implies that a strong AMOC is concomitant with weakened eddy-induced AMOC; or when the AMOC is weak, the eddy-induced mixing is strong. Although the causality between the mean and eddy-induced AMOC is still unclear, Fig. 1 provides a clue that the eddy-induced transport may play a role in the multi-centennial AMOC oscillation.

At the multi-centennial timescale, the change in upper-ocean density in the subpolar Atlantic (Fig. 2b) is determined predominantly by the change in salinity (Fig. 2c), while temperature change (Fig. 2d) acts against density change. When the AMOC is strong (weak) (Fig. 1a), the density stratification is weak (strong) (Fig. 2b), which is mainly due to a more saline (fresh) upper ocean (Fig. 2c), suggesting a well (weakly) ventilated structure in the subpolar upper ocean. Lag/lead correlation analyses reveal that the surface density, salinity, and temperature changes in the subpolar Atlantic all lead the AMOC change by about 10 years (figure not shown). This suggests a positive salinity feedback in triggering AMOC change, while the temperature change has a slight damping effect at the same time.
The evolution of anomalous salinity gradient (Fig. 2a, expressed as $\Delta S'$) between the subpolar upper ocean (above 1000 m) and deeper ocean resembles well the AMOC variation (Fig. 1a, expressed as $q'$); and the former leads the latter by about five years. It is straightforward that there is a linear relationship between them (i.e., $q' \sim \Delta S'$). Both eddy-induced mixing and diffusion can change $\Delta S'$ in the subpolar ocean, and thus may have important influence on $q'$. These two processes are different: the eddy-induced mixing is related to meso- and submeso-scale eddies at a relatively shallower depth and limited region as illustrated in Fig. 1e, while salinity diffusion is of molecular to turbulent scales, which occurs over a broader horizontal and vertical region. Hereafter, we will use “salinity mixing” for simplicity, instead of “salinity diffusion and eddy-induced mixing.”

To unravel the relationship between salinity mixing and $\Delta S'$ (equivalently, $q'$), scattering plots between them are shown in Fig. 3. Figure 3a suggests that the anomalous salinity in the subpolar upper ocean is damped more effectively when $\Delta S'$ is large. That is, when $\Delta S'$ is positively (negatively) large, the salinity mixing is negatively (positively) large so that the upper ocean saline (fresh) water can be removed efficiently. This nonlinear feature is mainly caused by the vertical salinity mixing when the AMOC is strong (i.e., $\Delta S' > 0$, $q' > 0$) (Fig. 3c), and by both the horizontal and vertical salinity mixing when the AMOC is weak (i.e., $\Delta S' < 0$, $q' < 0$) (Fig. 3b). This nonlinear feature can be physically understood as follows. When the stratification in the subpolar ocean is weak and the AMOC is strong, upper-ocean salinity anomaly can be mixed into the deeper ocean more efficiently by salinity diffusion. When the stratification in the subpolar ocean is strong and the AMOC is weak, the salinity diffusion is weak but the eddy-induced salinity mixing can enhance the downward mixing. Instead of canceling each other completely, the combined effect of diffusion and eddy-induced mixing is that the salinity mixing is enhanced when the AMOC anomaly is large. From now on, we simply use “enhanced mixing” for “enhanced eddy-induced mixing or diffusion when the AMOC anomaly is large.”
Here, the salinity mixing is obtained based on the following equation:

\[ V \frac{\partial S'}{\partial t} = F_{surf} + F_{Euler} + F_{mixing} \]  

(1)

where \( S' \) is salinity anomaly; \( F_{surf} \) and \( F_{Euler} \) represent the surface freshwater effect and Eulerian-mean advection effect, respectively; and \( F_{mixing} \) includes both horizontal and vertical mixing and both eddy-induced mixing and diffusion. \( V \) is the total volume of the subpolar ocean above 1000 m in depth, with the horizontal range outlined in Fig. 3d. \( V \frac{\partial S'}{\partial t} \), \( F_{surf} \), and \( F_{Euler} \) (units: \( \text{psu} \cdot \text{Sv} \)) can be calculated accurately offline based on model outputs. Therefore, \( F_{mixing} \) is obtained as the residual of these three terms, which is more reliable than directly calculating the mixing effect. In addition, we calculate directly the horizontal mixing. The vertical mixing is obtained as the residual of the total \( F_{mixing} \) and the horizontal component. Note that the way to decompose \( F_{mixing} \) into horizontal and vertical components is not accurate, but the total \( F_{mixing} \) is reliable. We need to bear in mind that large uncertainty exists in simulated eddies in ocean models, so the eddy-induced effect may vary among different models. Results here are meant to show the enhanced mixing mechanism may exist in coupled models.

Finally, based on the cubic regression between salinity mixing and \( \Delta S' \) (red curve in Fig. 3a) and considering the linear relationship between \( \Delta S' \) and \( q' \), the enhanced mixing effect can be expressed as follows,

\[ F_{mixing} \sim - (\Delta S')^3 \sim - (q')^2 \Delta S' \sim - k_m \Delta S' \]  

(2)

where \( k_m \sim (q')^2 \), parameterizing the enhanced mixing effect when the AMOC is large. Here, we emphasize that the detailed forms of Eq. (2) and \( k_m \) are not important. The regression between salinity mixing and \( \Delta S' \) can be linear or even the power of five. The cubic relationship is just a
reasonable outcome from the coupled model we used in this work. In summary, the enhanced mixing mechanism will be used in the box model in Section 4.

3. Hemispheric box model

The conceptual model used here consists of four ocean boxes (Fig. 4a). With a zonal width of 60°, the meridional scales of these boxes are 45° and 25° for the tropical and subpolar basins, respectively. This selection is based on the horizontal patterns of surface density variances from the CESM results (Fig. 3d). The AMOC is included explicitly, with sinking in the extratropics and rising in the tropics. Such box model was first presented in Stommel (1961), and has been used in different forms by many researchers (e.g., Marotzke 1990; Huang et al. 1992; Nakamura et al. 1994; Tziperman et al. 1994; Marotzke and Stone 1995; Yang et al. 2016; Zhao et al. 2016). Since the relaxation time for temperature is substantially shorter than that for salinity (Huang 1994), the salinity effect provides the main control on low-frequency variation of density in the subpolar ocean (Rooth 1982), which is confirmed by our CESM control run (Fig. 2). Therefore, only salinity changes are considered in the box model. The salinity equations for the 4-box ocean model are written as follows:

\[ V_1 \dot{S}_1 = q(S_4 - S_1) + F_w \]  
\[ V_2 \dot{S}_2 = q(S_1 - S_2) - F_w \]  
\[ V_3 \dot{S}_3 = q(S_2 - S_3) \]  
\[ V_4 \dot{S}_4 = q(S_3 - S_4) \]

where \( V_i \) (\( i = 1, 2, 3, 4 \)) is the volume of each box, \( q \) is the volume transport by the AMOC, and \( F_w \) is freshwater flux. The equilibrium state of the box model is,
\[
\bar{q}(\bar{S}_1 - \bar{S}_2) = \bar{F}_w, \quad \bar{S}_2 = \bar{S}_3 = \bar{S}_4.
\] (4)

Here, \(\bar{q}\) is given by a constant of 10 Sv in the box model, which is a typical value in previous single hemispheric studies (Winton and Sarachik 1993; Cessi 1994; Roebber 1995). The upper ocean in the box model is 500 m in thickness, so that the turnover time in the upper ocean is similar to that in the CESM. Based on the CESM results, \(\bar{S}_1\) and \(\bar{S}_2\) are set to 36 and 33.5 psu, respectively, which gives 
\[
\bar{F}_w = 2.5 \times 10^7 \text{ psu} \cdot m^3 s^{-1}.
\] Please refer to Table 1 for other parameters.

We assume a small amplitude of oscillation with respect to the equilibrium climate, and do not consider surface freshwater flux change (i.e., \(F_w = \bar{F}_w\)). The total volume transport \(q\) consists of a mean state \(\bar{q}\) and salinity-driven perturbation \(q'\). The latter is assumed to be linearly proportional to meridional density difference caused only by salinity change:

\[
q = \bar{q} + q' = \bar{q} + \lambda \Delta \rho'
\] (5)

where

\[
\Delta \rho' = \rho_0 \beta [\delta(S_2' - S_1') + (1 - \delta)(S_3' - S_4')], \quad \text{and} \quad \delta = \frac{V_1}{V_1 + V_4} = \frac{V_2}{V_2 + V_3} = \frac{D_1}{D}
\] (6)

Here, \(\lambda\) is a linear closure parameter, depicting the efficiency that the meridional perturbation salinity gradient affects the AMOC; \(\rho_0\) is the reference seawater density; \(\beta\) is the saline expansion coefficient of seawater; \(D_1\) and \(D\) are upper level and total ocean depths, respectively. Their values are listed in Table 1. The system of Eqs. (3)–(6) is similar to that of RT97, except that the subpolar box is divided into two boxes (Fig. 4a), for further consideration of vertical mixing in the subpolar ocean. In addition, while temperature is not explicitly shown in the equations, it virtually determines the mean transport \(\bar{q}\) and freshwater flux \(\bar{F}_w\), which have indirect impacts on the eigenmodes of the system.
The linearized version of Eq. (3) can be written as follows:

\[ V_1 \dot{S}_1' = q'(\overline{S}_4 - \overline{S}_1') + \overline{q}(S_4' - S_1') \] (7a)

\[ V_2 \dot{S}_2' = q'(\overline{S}_1 - \overline{S}_2') + \overline{q}(S_1' - S_2') \] (7b)

\[ V_3 \dot{S}_3' = \overline{q}(S_2' - S_3') \] (7c)

\[ V_4 \dot{S}_4' = \overline{q}(S_3' - S_4') \] (7d)

where \( S_i' \ (i = 1, 2, 3, 4) \) is the perturbation salinity of each ocean box.

Eigenvalues can be calculated numerically from the linearized Eq. (7). With parameters in Table 1, there are a pair of conjugate eigenvalues \([0.31 \pm 5.83i] \cdot 10^{-10} \text{ s}^{-1}\), corresponding to a 340-year period and 1025-year e-folding time. This is a marginally unstable mode because of the small positive real part. The eigenvector \([S_1', S_2', S_3', S_4']\) is \([-0.11 + 0.74i, 0.55 - 0.08i, 0.20 - 0.29i, -0.04 - 0.04i]\), showing a dipole pattern between the tropical and subpolar upper oceans by either the real part or the imaginary part. The other two eigenvalues are 0 and \(-37.4 \times 10^{-10} \text{ s}^{-1}\), respectively, suggesting a zero mode (caused by the total salt conservation) and a purely damped mode with e-folding time of about 8.5 years.

4. Self-sustained oscillation in the box model

4.1 Unstable oscillation

Results from forward numerical integration of Eq. (7) are shown in Fig. 5. The fourth-order Runge-Kutta method is used to solve Eq. (7), with an initial perturbation of \( S_2' = -0.02 \text{ psu} \) and \( S_1' = S_3' = S_4' = 0 \). The integration time step is 7.2 days, and the total integration length is 10000 years.
Given the velocity closure parameter \( \lambda = 12 \text{ Sv} \cdot m^2 k g^{-1} \), the time series of salinity anomalies show a periodic oscillation of about 340 years, and a gradually enhanced amplitude (Fig. 5a), which are exactly predicted by the eigenvalue discussion in Section 3. \( S_1' \) leads \( S_2' \) by about \( \frac{2\pi}{3} \), suggesting roughly a dipole pattern in tropical-subpolar surface salinity distribution. \( S_2' \) only leads \( S_3' \) slightly, suggesting an efficient vertical advection by the mean AMOC. The perturbed AMOC (\( q' \)) shows a similar oscillation, with its amplitude growing gradually (Fig. 5b). \( q' \) lags \( S_1' \) by about \( \frac{2\pi}{3} \), and is roughly out of phase with \( (S_1' - S_2') \). Besides, \( q' \) lags \( (S_2' - S_3') \) by about \( \frac{\pi}{2} \). This phase lag can be reduced by including vertical mixing between the two subpolar boxes with a constant mixing coefficient. Note that the integral of \( \frac{d (S_1')^2}{dt} \) in a period is positively correlated to the growth rate. To identify the mechanism for amplitude change, The three terms in Eq. (7b) are multiplied by \( S_2' \) and shown in Fig. 5c. We can see the growing perturbation comes from the positive perturbation advection feedback \( (q'(S_1' - S_2')) \), and the negative mean advection feedback \( (\bar{q}(S_1' - S_2')) \) tends to restrain the growth, which is consistent with the results in Marotzke (1996).

4.2 Self-sustained oscillation with enhanced mixing

In this subsection, we show that including the enhanced mixing process in the subpolar ocean leads to a stable self-sustained oscillation. Terms representing the enhanced vertical mixing in the subpolar ocean are added to Eqs. (7b) and (7c) as follows:

\[
V_2 S_2' = q'(\bar{S}_1 - \bar{S}_2) + \bar{q}(S_1' - S_2') - k_m(S_2' - S_3') \quad (8a)
\]

\[
V_3 S_3' = \bar{q}(S_2' - S_3') + k_m(S_2' - S_3') \quad (8b)
\]
where $k_m$ (units: $m^3/s$) is proportional to vertical mixing coefficient. In this paper, $k_m$ is defined as a function of the overturning rate anomaly squared:

$$k_m = \kappa q'^2$$

(9)

Here, $\kappa$ is a positive constant (units: $m^{-3}s$), so that vertical mixing in the subpolar ocean is always enhanced when perturbation grows. Specifically, we choose $\kappa = 10^{-3} m^{-3}s$.

We would like to emphasize several points here. First, the oscillation behaviors are similar for different forms of $k_m$, as long as the enhanced mixing mechanism can be explicitly expressed. For example, we can also define $k_m = 100|q'|$. Second, the value of $\kappa$ affects the amplitude of the oscillation but not the period. Third, a background vertical mixing can be included, say, $k_m = \bar{k}_m + k'_m$, which will not qualitatively change the results as long as $\bar{k}_m$ is not too big. The background vertical mixing is neglected for simplicity, although using a non-zero $\bar{k}_m$ may give more realistic results. Fourth, while only vertical mixing is included in the box model, horizontal mixing is also important for the enhanced mixing mechanism as revealed in the CESM results (Fig. 3b).

A speedup of the AMOC can lead to an increase of $S'_2$. This would increase the subpolar upper-ocean density and reduce stratification, enhancing downward diffusion, which in turn helps remove upper-ocean saline water. A slowdown of the AMOC can lead to a fresher subpolar upper ocean. While the stratification becomes more stable, the vertical mixing can become stronger due to eddies’ effect, as examined in Section 2. The enhanced eddy activities help mix the anomalous freshening signal in the upper ocean downward, so the signal can flow out of the subpolar faster. A self-sustained oscillation of the system exists under this mechanism (Fig. 6). Now, the growing perturbation due to the positive perturbation advection feedback ($q'(\bar{S}_1 - \bar{S}_2)$) can be eventually balanced by the
combined negative feedback of mean advection \( \langle q (S'_1 - S'_2) \rangle \) and enhanced vertical mixing \( (k_m (S'_3 - S'_2)) \) (Fig. 6c). The latter provides a critical mechanism to stabilize the system.

4.3 Understanding the self-sustained oscillation

The enhanced mixing helps stabilize the system by removing the subpolar surface salinity anomaly out of the region. To further understand the self-sustained oscillation, let us assume an extreme situation: the vertical mixing is so strong that the salinity in the subpolar upper ocean and that in the subpolar lower ocean almost instantly becomes the same. Therefore, the 4-box model (Fig. 4a) can be interpreted as a 3-box model (Fig. 4b), making it comparable to that of RT97. Now, Eqs. (7) and (8) become,

\[
\begin{align*}
V_1 \dot{S}'_1 &= q' ( \bar{S}'_4 - \bar{S}'_1 ) + \bar{q} (S'_4 - S'_1) \\
V_2 \dot{S}'_2 &= q' ( \bar{S}'_1 - \bar{S}'_2 ) + \bar{q} (S'_1 - S'_2) \\
V_4 \dot{S}'_4 &= \bar{q} (S'_2 - S'_4)
\end{align*}
\]

where

\[
q' = \lambda \Delta \rho' = \lambda \rho_0 \beta [S'_2 - \delta S'_1 - (1 - \delta) S'_4], \quad \text{and} \quad \delta = \frac{V_1}{V_1 + V_4} = \frac{\rho_1}{\bar{\rho}}
\]

The 3-box model has the same equilibrium as the 4-box model. By applying the total salt conservation \( V_1 S'_1 + V_2 S'_2 + V_4 S'_4 = \text{constant} \), the 3-box system can be further simplified, and its eigenvalues can be obtained theoretically. Here, we define the following quantities:

\[
\begin{align*}
S_d &\equiv \frac{F_w}{\bar{q}} = \bar{S}_1 - \bar{S}_4 = \bar{S}_1 - \bar{S}_2; \quad \rho_d = \delta \beta \rho_0 S_d \\
T_n &= \frac{V_1 + V_2 + V_4}{\bar{q}}; \quad \delta_i = \frac{V_i}{V_1 + V_2 + V_4}, \quad i = 1, 2, 4 \\
\tau &= \frac{t}{T_n}; \quad a' = \frac{s'_1 - s'_2}{S_d}; \quad h' = \frac{\Delta \rho}{\rho_d}; \quad M = \frac{\rho_4}{\bar{q}} \lambda
\end{align*}
\]
Eq. (10) can be reduced to:

\[
\frac{da'}{d\tau} = -C_3 a' - (C_4 + C_1 M) h' \\
\frac{dh'}{d\tau} = C_2 a' + C_2 M h'
\] (12a)

(12b)

Assuming the form of solution \(a' = A e^{\omega \tau}\), Eq. (12) has an eigenvalue:

\[
\omega = \frac{1}{2} \left[ (C_2 M - C_3) \pm \sqrt{(C_2 M - C_3)^2 - 4C_2 C_4 (1 - M)} \right] 
\] (13)

where \(C_1 = \frac{1}{\delta_1} + \frac{1}{\delta_2}, C_2 = \frac{1}{\delta_1 \delta_2}, C_3 = \frac{1}{\delta_1} + \frac{1}{\delta_2} + \frac{1}{\delta_4}, \) and \(C_4 = \frac{1}{\delta_4}.

Based on (13), the essential stability condition for the system is

\[
M \leq \min \left( \frac{C_1}{C_2}, 1 \right) 
\] (14)

and the oscillation condition of the system is

\[
M_1 < M < \min(M_2, 1) 
\] (15)

where \(M\) is non-dimensional linear closure parameter, and

\[
M_1 = \frac{C_3 - 2C_4}{C_2} - \frac{2}{C_2} \sqrt{C_4^2 + C_4 (C_2 - C_3)}, \quad
M_2 = \frac{C_3 - 2C_4}{C_2} + \frac{2}{C_2} \sqrt{C_4^2 + C_4 (C_2 - C_3)}.
\]

Therefore, \(\lambda_{1,2} = \frac{\pi}{\rho_d M_{1,2}}\).

Figure 7 shows the dependence of real and imaginary parts of \(\omega\) on the non-dimensional parameter \(M\). Given the parameters listed in Table 1, \(M_1 \approx -0.02\) and \(M_2 \approx 0.51\). When \(M > M_2\) or \(M < M_1\), there are only two real positive or negative \(\omega\) (Fig. 7a), suggesting the system has only exponentially growing (decaying) modes. An oscillatory mode occurs when \(M_1 < M < M_2\). With the
decrease of $M$, the oscillation mode changes from an unstable oscillation ($Re[\omega] > 0$) to a stable oscillation ($Re[\omega] < 0$) (Fig. 7b). When $M = 0$, $q' = 0$, that is, the AMOC does not respond to any salinity change, the oscillation mode is a strong damped mode, with the oscillation timescale of about 1000 years and the damping timescale of about 50 years. The peak value of the imaginary part of $\omega$ corresponds to the shortest timescale (Fig. 7b), which is about 350 years ($\frac{2\pi(v_1v_2)^{\frac{1}{2}}}{q}$) with the damping timescale of 340 years when $M \approx 0.245$.

Given an eigenvalue $\omega$, the corresponding eigenvector is $\left[\frac{\omega}{C_2} - M, 1\right]^T \equiv [p, 1]^T$. When ignoring the term $e^{\omega t}$, the theoretical solutions to $S'_1$ and $q'$ can be written as follows,

$$S'_1 = S_d p + S'_2, \quad S'_2 = \delta_1 S_d, \quad S'_4 = \frac{\delta}{\delta - 1} S_d (1 + p) + S'_2 \quad (16)$$

$$q' = \lambda \delta \beta \rho_0 S_d = \bar{q} M \quad (17)$$

Therefore, based on Eqs. (10b), (16), and (17), the salinity tendency in the subpolar box is

$$V_2 \dot{S}'_2 \sim (q' S_d + \bar{q} S_d p) = \bar{q} S_d (M + p) \quad (18)$$

whose sign is determined by $Re[\Omega + p]$. Here, the salinity tendency caused by perturbation flow tends to have the same sign as the salinity anomaly (here $M > 0$ in most cases, as shown in Fig. 7b).

The salinity tendency caused by mean flow relies on $p$, which can be of opposite sign when $Re[\omega] < C_2 M$. Further, if $Re[\omega] < 0$, we have $Re[p] < -M$; therefore, the negative feedback by the mean flow would dominate over the positive feedback by the perturbation flow, and the 3-box model is stable. In summary, the perturbation advection of mean salinity has a positive feedback, and the mean advection of salinity anomaly has a negative effect. A self-sustained oscillation can be realized when these two feedbacks are balanced.
Mathematically, a stable oscillation in the 3-box model can exist when $Re[\omega] < 0$. Therefore, the stability condition can be expressed in dimensional form as follows:

$$\lambda < \lambda_C \equiv \frac{\bar{q}^2}{\rho_o \beta \bar{F}_w} \left[ 1 + \frac{\delta_2}{\delta (1-\delta)} \right]$$  \hspace{1cm} (19)

where $\lambda_C$ (marked in Fig. 7a) is the critical linear closure parameter when $Re[\omega] = 0$, which is determined by mean AMOC strength $\bar{q}$, surface freshwater flux $\bar{F}_w$ as well as basin geometry. Eq. (19) shows that a stronger $\bar{F}_w$ and a weaker $\bar{q}$ give a smaller $\lambda_C$, implying higher possibility for an unstable oscillation, since the background meridional salinity gradient in this situation will be stronger. This is qualitatively consistent with previous studies showing that increasing freshwater flux could lead to a regime shift (e.g., Zhang et al. 2002). In addition, salinity anomalies also spend more time at the surface with a weaker mean AMOC. This will also make the system more unstable from a Lagrangian viewpoint, and this is why we have a quadratic term of $\bar{q}$. Overall, these are similar to the results of Sevellec et al. (2006), that is, a competition between the mean AMOC and freshwater flux determines the salinity behavior. Eq. (19) also shows a bigger volume of the subpolar ocean ($\delta_2$) gives a larger $\lambda_C$, implying a higher probability for a stable oscillation. In this situation, the salinity difference anomaly between subpolar and tropical upper oceans is larger under the same $q'$, and thus the mean advection of salinity anomaly is stronger, which would result in a stronger stabilizing effect. This is consistent with the analysis on Eq. (18).

Linear stability analyses on the 4-box model (Eq. (7)) without the enhanced vertical mixing can be done numerically. The results are also plotted in Fig. 7 for comparison. Generally, the eigenvalues in the 3-box and 4-box models are very close to each other. However, there is a subtle but very important difference between them: The 3-box model is more stable than the 4-box model, and the critical linear closure parameter in the 4-box model, noted as $\lambda'_C$, is slightly smaller than $\lambda_C$ in the 3-
A neutral oscillation can be observed in the 4-box model when $\lambda = \lambda'_C$ and in the 3-box model when $\lambda = \lambda_C$.

Figure 8 shows the oscillation behaviors of the AMOC in the 3-box model and 4-box model with and without the enhanced vertical mixing under different $\lambda$. When $\lambda < \lambda'_C$, the AMOC in all models exhibits a decaying oscillation (Fig. 8a). When $\lambda > \lambda_C$, the AMOC in all models shows an unstable oscillation with a fast-growing amplitude (Fig. 8c), which implies that even an extreme vertical mixing is still inadequate to restrain the growing perturbation. Here, we emphasize that a self-sustained oscillation can only exist in the 4-box model with the enhanced vertical mixing. The self-sustained oscillation is meant to be a growing oscillation due to the linear instability that can be constrained by itself as shown in Fig. 6. While neutral oscillations ($Re[\omega] = 0$) can be found in the 3-box model and 4-box model without the enhanced vertical mixing, they cannot grow up from small perturbation.

A self-sustained oscillation can exist in the 4-box model with the enhanced vertical mixing when $\lambda'_C < \lambda < \lambda_C$. Within this parameter range, the 3-box system is stable, while the 4-box system without the enhanced vertical mixing is unstable (Fig. 7a). As we know from linear analyses that energy is provided by the mean salinity gradient in the upper ocean, salinity anomaly accumulates in the subpolar upper ocean and then flows downward and southward. When the AMOC is close to its mean state (i.e., $q' \sim 0$), the system is unstable because the enhanced mixing plays no role ($k_m \sim 0$) in the removal of salinity anomalies from the subpolar ocean. When the AMOC anomaly is large, the enhanced vertical mixing accelerates the outflow of salinity anomaly in the subpolar upper ocean. When the vertical mixing is too strong, salinity anomaly will be removed before it grows, and the system becomes stable again, similar to the 3-box model as the limit.
5. Discussion

5.1 Oscillation dependence on model geometry

Model geometry can affect both oscillation period and e-folding time of the system (Fig. 9). Under a reasonable range of the basin size, the multi-centennial timescale is notable (Fig. 9a), since this timescale is fundamentally determined by the background parameters, such as the mean AMOC ($\bar{q}$) and the total volume of the upper ocean ($V_1 + V_2$). These two parameters suggest an upper-ocean turnover time ($[(V_1 + V_2)/\bar{q}]$) of about 54 years, given the values listed in Table 1. This turnover time corresponds to an oscillation period of 340 years ($2\pi \times 54$), which is similar to the period of the eigenvalue through the linear stability analysis on Eq. (7) in Section 3. The consistency between the turnover time of overturning circulation and the oscillation period of the system was revealed theoretically in Sévellec et al. (2006), in which analytical expressions for the oscillation period were derived for the fixed-temperature case and a more general case including both salinity and temperature changes.

The dependence of the 4-box eigenvalue on model geometry is illustrated in Fig. 9. Here, for simplicity and without losing generality, we keep the equilibrium $\bar{q}$, $\bar{F}_W$, the ocean total volume, and $\lambda$ in Table 1 unchanged. Figure 9a shows that for a given upper-ocean volume ($V_1 + V_2$), the oscillation period is weakly sensitive to the subpolar ocean volume ($V_2 + V_3$); for a given subpolar ocean volume, the oscillation period increases monotonically with the increase of the upper-ocean volume. In other words, for a bigger upper-ocean volume, there would be a longer oscillation mode. Note that under a reasonable range of basin geometry, the millennium oscillation mode has a shorter positive e-folding time (Fig. 9b). This implies that the intrinsic millennium modes would have to be fast-growing modes in our hemispheric box model. Bear in mind that this conclusion is obtained
based on the equilibrium $\overline{q}$ and $\overline{F}_W$ given in Table 1, and a more thorough study under different mean state of climate is needed.

The self-sustained oscillation can exist only in region 2, which is enclosed by the two orange lines in Fig. 9. This region denotes the regime where unstable oscillation occurs in the 4-box model while stable oscillation appears in the 3-box model, as discussed in Section 4.3. In region 1, the modes in both 3-box and 4-box models are decayed modes with negative e-folding time (Fig. 9b). In region 3, the modes in both 3-box and 4-box models are unstable, and even extreme mixing cannot lead to a stable oscillation. The orange star in Fig. 9 denotes the mode under the standard parameters in Table 1. The grey region in Fig. 9a denotes the region of no oscillation mode, i.e., the situation when $\lambda \geq \lambda'_2$. In general, under these ranges of parameters, a naturally self-sustained oscillation, relatively, is not that easy to happen, in terms of the ratio of region 2 to the total area. This might explain the fact that, in reality, relative to the other low-frequency climate variabilities, the multi-centennial oscillation is under appreciated and inadequately studied (Stocker and Mysak 1992).

5.2 Multi-centennial oscillation forced by stochastic freshwater flux

External stochastic forcing can also excite a multi-centennial oscillation in the 4-box model (Fig. 10). We test the response of stable modes to stochastic forcing in the 4-box model. The linear closure parameter $\lambda = 9.45 \text{ Sv} \cdot m^3 kg^{-1}$, corresponding to an eigenvalue with a period of 332 years and e-folding time of 285 years. This stable mode is within region 1 of Fig. 9. After adding a stochastic freshwater perturbation, Eq. (8a) of the 4-box model is rewritten as follows,

$$V_2 \dot{S}'_2 = \cdots + V_2 N$$  (20)
where $N$ denotes the random freshwater perturbation, which is generated from a first-order autoregressive model,

$$N_{k+1} = \alpha N_k + \sigma G_k$$  \hfill (21)

The noise $N$ at time step $k + 1$ is generated from the noise at time step $k$ and a standard Gaussian random variable $G_k$. Here, $\sigma$ is given as 0.03, 0.005, and 0.001 psu/year; and correspondingly, $\alpha$ is set to 0.78, 0.98, and 0.998, respectively, which gives an e-folding time of auto-covariance function of about one month, one year, and 10 years, respectively. For each pair of parameters, 100 noise samples are generated. The 4-box model is integrated for 5000 years for each noise forcing, and the last 2000 years are used for analysis.

Forced by stochastic freshwater forcing (Fig. 10c), the 4-box model in region 1, which has only damped internal oscillations, exhibits sustained oscillation even in the presence of damping effect of nonlinear vertical mixing (Fig. 10a). The power spectrum of the forced AMOC anomaly suggests a multi-centennial oscillation (Fig. 10b). The ratio of the AMOC spectrum to the freshwater spectrum is shown in Fig. 10d, which suggests the AMOC responses most efficiently to the noise whose period is about 330 years. This principal period is similar to the period from the linear analysis. Our study suggests that the multi-centennial variation may be an intrinsic feature of the Atlantic Ocean, regardless of external forcing.

5.3 Effect of nonlinear salinity advection

The effect of nonlinear salinity advection (i.e., $q'\Delta S'$) on the oscillation behaviors in the box model is briefly examined in Fig. 11 by directly integrating the original Eq. (3), using the same
parameters listed in Table 1. Without the enhanced vertical mixing, the nonlinearity will lead to a regime shift of the AMOC much faster than that in the linear system (Fig. 11a). However, in the presence of enhanced vertical mixing, the nonlinearity has nearly no effect on the oscillation in both amplitude and period (Fig. 11b). This is because, as mentioned previously, the enhanced vertical mixing is a critical stabilizing factor that can efficiently restrain the growth of perturbations, and the magnitude of $q'\Delta S'$ is always one order smaller than that of the other terms.

6. Summary

A self-sustained multi-centennial oscillation of the AMOC is identified via a simplified modeling study. It is shown that the positive salinity advection feedback is important to the multi-centennial oscillation. The enhanced mixing process in the subpolar ocean can be a critical mechanism to weaken the positive salinity advection feedback and limit the amplitude of the oscillation, ultimately leading to a self-sustained oscillation (Fig. 6). Mathematically, the oscillation period is determined by the eigenvalue of the system. Physically, it is determined by the turnover time of the upper-ocean water.

The enhanced mixing mechanism proposed in this work includes only two components: the mixing due to meso- and submeso-scale eddies, and the diffusion related to the molecular to turbulent scale activities. Other processes, such as the rapid deep convection in the subpolar ocean, may have similar effects on the oscillation as vertical mixing. Deep convection is difficult to quantify in both models and real world, and it can be treated as an extreme vertical mixing event that happens in one time-step in a box model as in Colin de Verdière et al. (2006). Our result suggests the eddy-induced mixing can influence the stability of the AMOC. Recent studies have shown that mesoscale
turbulence may vary on climate timescale (Busecke and Abernathey 2019), and long-term variability can be damped by meso-scale turbulence (Sévellec et al. 2021). Further research is needed to quantify the role of turbulence on large-scale variability.

While both the enhanced mixing used in this paper and the nonlinear velocity closure used in RT97 can lead to a self-sustained oscillation by limiting the salinity advection feedback, these two mechanisms are different. In this paper, the enhanced mixing can transport salinity anomalies out of the subpolar upper ocean before the anomalies accumulate and lead to a bounded growth. The nonlinear velocity closure limits the development of perturbation flow so that it helps the system stay near its equilibrium. RT97 noted that the nonlinearity may come from internal nonlinear relation between large-scale meridional density gradient and AMOC strength. However, we prefer the enhanced mixing mechanism in maintaining the self-sustained oscillation of the system, since mixing processes, to some extent, are detectable and can be quantified in both observations and coupled models. In general, both the enhanced mixing and nonlinear volume transport closure can lead to a self-sustained oscillation. The two mechanisms are not contradictory. A combination of these two in a model may bring out more interesting results.

This study is motivated by the multi-centennial variation of the AMOC found in a CESM control run (Fig. 1), which highlights the roles of internal processes in self-sustained oscillation. Bear in mind that external stochastic forcing can also excite a multi-centennial oscillation in the box models (Fig. 10). It is still far from determined whether the multi-centennial oscillation in our CESM control run is a self-sustained one. The simple model results in this work can be used to explain the oscillations in coupled climate models. However, its physics is too simple. Whether the mechanism of self-sustained oscillation is still valid in a system when more complex physics are considered needs to be investigated. For example, if temperature equation is included in the box model, does the self-
sustained oscillation still exist? As described in the introduction, there are comprehensive studies on the stability and regime shift of the AMOC, and the centennial to millennium oscillations of the AMOC can be well identified (e.g., Sévellec et al. 2006, 2010). However, we would like to search for self-sustained multi-centennial internal modes in a more realistic model, but not the oscillation with decaying or growing amplitude under external forcing. The existence of intrinsic modes of multi-centennial timescale in the Atlantic Ocean may have profound implication for paleoclimatic studies and interpretations of the on-going and future climate change.

Acknowledgement: This research is jointly supported by the NSF of China (Nos. 41725021, 91737204, and 41376007) and by the foundation at the Shanghai Frontiers Science Center of Atmosphere-Ocean Interaction of Fudan University.
References


LaMarche, V.C., 1974: Paleoclimatic Inferences from Long Tree-Ring Records: intersite comparison shows climatic anomalies that may be linked to features of the general circulation. *Science*, 183, 1043-1048.


Table 1  Parameters used in this study.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Physical Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_2$</td>
<td>Volume of upper subpolar Atlantic box</td>
<td>$2.8 \times 10^{15} \text{ m}^3$</td>
</tr>
<tr>
<td>$V_1$, $V_3$, $V_4$</td>
<td>Volumes of upper tropical Atlantic, lower subpolar Atlantic, and lower tropical Atlantic boxes, respectively</td>
<td>$5V_2$, $7V_2$, $35V_2$</td>
</tr>
<tr>
<td>$D_1$, $D_2$, $D$</td>
<td>Depths of upper box, lower box, and total, respectively</td>
<td>$500$, $3500$, $4000 \text{ m}$</td>
</tr>
<tr>
<td>$\overline{S}_1$, $\overline{S}_2$, $\overline{S}_3$, $\overline{S}_4$</td>
<td>Reference salinity values of the four ocean boxes</td>
<td>$36$, $33.5$, $33.5$, $33.5 \text{ psu}$</td>
</tr>
<tr>
<td>$\overline{q}$</td>
<td>Equilibrium AMOC strength</td>
<td>$10 \text{ Sv} \left(10^6 \text{ m}^3\text{s}^{-1}\right)$</td>
</tr>
<tr>
<td>$F_w$</td>
<td>Total virtual salt flux</td>
<td>$2.50 \times 10^7 \text{ psu m}^3\text{s}^{-1}$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Haline contraction coefficient</td>
<td>$7.61 \times 10^{-4} \text{ psu}^{-1}$</td>
</tr>
<tr>
<td>$\rho_0$</td>
<td>Reference density</td>
<td>$1.00 \times 10^3 \text{ kg m}^{-3}$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Linear closure coefficient</td>
<td>$12 \text{ Sv kg}^{-1}\text{m}^3$</td>
</tr>
</tbody>
</table>
Figure captions

**Figure 1** (a) Atlantic meridional overturning circulation (AMOC) index (units: Sv, 1 Sv = 10^6 m^3 s^-1).
Thin vertical colorbars show the annual mean AMOC with respect to its climatological mean of 24 Sv. Thick black curve is the lowpass-filtered AMOC index after applying a Lanczos filter with 121 weights, with a cutoff period of 60 years. The AMOC index is defined as the maximum streamfunction in the region of 20°-70°N between 200 and 3000 m in the Atlantic. (b) and (c) are for climatological patterns of Eulerian-mean and eddy-induced AMOC (units: Sv), respectively. (d) and (e) are for regression patterns of Eulerian-mean and eddy-induced AMOC, respectively, with respect to the mean-normalized filtered AMOC index.

**Figure 2** (a) Time series of anomalous vertical salinity gradient between subpolar upper ocean (above 1000 m) and deeper ocean (units: psu). Thin vertical colorbars show the annual mean anomalies. Thick black curve is the lowpass-filtered anomalies after applying a Lanczos filter with 121 weights, with a cutoff period of 60 years. Temporal evolutions of (b) annual mean anomalous potential density and the contributions due to (c) salinity change and (d) temperature change (units: kg · m^-3). These variables are horizontally averaged over the subpolar Atlantic region defined in Fig. 3d.

**Figure 3** (a) Scattering plots of anomalous salinity mixing (units: psu · Sv) in the subpolar upper ocean (above 1000 m) versus the salinity difference anomaly between subpolar upper ocean and deeper ocean (units: psu). (b) and (c) show the horizontal and vertical components (units: psu · Sv), respectively. The red curve shows the cubic regression result, and all these data are filtered with a cutoff period of 10 years. (d) Sea-surface density anomaly (units: kg · m^-3) regressed onto the mean-normalized filtered AMOC index. Black curves encircle the region defined for the area-
average region, and these curves are parallel to the grid lines of the ocean model. In (a), the expression for cubic regression $F_{mixing} = -285.15 \times (\Delta S')^3 - 2.51 \times (\Delta S')^2 - 1.28 \times (\Delta S') - 0.003$, and the explained variance is about 25%; the effective rank for the scaled Vandermonde coefficient matrix is 4, suggesting the cubic regression polynomial is significant.

**Figure 4** Schematic diagrams of box models: (a) 4-box ocean model and (b) 3-box ocean model. Ocean boxes are denoted by ①, ②, ③, and ④. Boxes 1 and 4 represent the upper and lower layers of the tropical ocean, respectively; boxes 2 and 3, of the subpolar ocean, respectively. $D_1$ and $D_2$ are the depths of upper and lower ocean layers, respectively. $F_w$ is the net freshwater gain (loss) in the subpolar (tropical) ocean. $q$ represents the thermohaline circulation. With the extreme vertical mixing in the subpolar ocean, the 4-box model can be interpreted as a 3-box model.

**Figure 5** Unstable oscillation in the 4-box model without the enhanced vertical mixing considered in the subpolar ocean. (a) Salinity anomalies $S_1', S_2', S_3$ (units: psu); (b) salinity stratification in the subpolar ocean ($S_2' - S_3'$) (units: psu), the meridional salinity gradient ($S_1' - S_2'$) (units: psu), and AMOC anomaly ($q'$) (units: Sv). The vertical coordinate on the right side is for $q'$. (c) shows the terms in Eq. (7b), which are multiplied by $S_2'$ to represent the evolution of amplitude change (units: $\text{psu}^2 \text{Sv}$). The linear closure parameter $\lambda = 12 \text{Sv} \cdot kg^{-1} m^3$, and its non-dimensional value $M = 0.285$.

**Figure 6** Same as Fig. 5, but for self-sustained oscillations when considering the enhanced vertical mixing in the subpolar ocean. In (c), the units are $10^{-1} \text{psu}^2 \text{Sv}$.

**Figure 7** Dependence of eigenvalue $\omega$ in the 3-box and 4-box models on the non-dimensional linear closure parameter $M$. (a) and (b) are for real and imaginary parts of the non-dimensional
eigenvalue $\omega$, respectively. Black curves are for the 3-box model, which are obtained theoretically. Orange curves are for the 4-box model without the enhanced vertical mixing, which are obtained numerically. $\lambda_1$, $\lambda_2$, and $\lambda_C$ are dimensional values of $M$ for the 3-box model; $\lambda'_2$ and $\lambda'_C$ are for the 4-box model. $\lambda'_C = 11.45$ and $\lambda_C = 12.42 \, Sv \cdot kg^{-1} m^3$; and the corresponding $M = 0.272$ and 0.295, respectively.

Figure 8 Oscillations of AMOC anomalies (units: Sv) in the 3-box and 4-box models with and without the enhanced vertical mixing. (a) $M=0.25$ ($\lambda < \lambda'_C$, region 1 in Fig. 9), showing decaying oscillations for both models. (b) $M=0.285$ ($\lambda'_C < \lambda < \lambda_C$, region 2 in Fig. 9), showing a decaying oscillation in the 3-box model, a self-sustained oscillation 4-box model with $k_m$ and a growing oscillation in the 4-box model without $k_m$. The vertical coordinate on the right side is for the 4-box model with $k_m$. (c) $M=0.35$ ($\lambda > \lambda_C$, region 3 in Fig. 9), showing growing oscillations in both models. $\lambda'_C$ and $\lambda_C$ are marked in Fig. 7a.

Figure 9 Relations of (a) period and (b) e-folding time of oscillation modes in the 4-box model without enhanced vertical mixing to the relative geometry of the ocean boxes (units: years). The abscissa and ordinate represent the volume fractions of upper ocean ($V_1 + V_2)/V$ and subpolar ocean ($V_2 + V_3)/V$, respectively. $V$ is the total ocean volume. The grey region in (a) denotes the region of no oscillation modes. Solid and dashed orange curves show the location of critical values $\lambda'_C$ in the 4-box model and $\lambda_C$ in the 3-box model, respectively. The oscillations are decayed in region 1, self-sustained in region 2, and unstable in region 3 when considering the enhanced vertical mixing. The star denotes the standard geometry and mode for parameters listed in Table 1.

Figure 10 (a) Time series of AMOC anomaly (units: Sv) in the 4-box model when considering enhanced vertical mixing in the subpolar ocean, forced by stochastic freshwater flux. Here,
M=0.225 (\lambda = 9.45 \text{ Sv} \cdot kg^{-1}m^3); and the 4-box model is in a stable regime (region 1 in Fig. 9).

(b) and (c) are the power spectra (units: dB) of AMOC anomalies and the noise, respectively. (d) shows the ratios of the AMOC spectrum to the stochastic freshwater spectrum (units: dB) with peaks around 0.003 cpy (330 years). In (b), (c), and (d), thick curves represent ensemble mean, whose standard deviations are denoted by vertical short line segments. The black, green, and orange lines show results under red noises whose damping timescale are 1 month, 1 year, and 10 years, respectively.

Figure 11  Comparison between AMOC anomaly (units: Sv) obtained from time-forward integration of the original nonlinear Eq. (3) (orange) and corresponding linearized Eq. (7) (black). (a) is for oscillation without considering enhanced vertical mixing in the subpolar box, and (b), with enhanced vertical mixing.
**Figure 1** (a) Atlantic meridional overturning circulation (AMOC) index (units: Sv, 1 Sv=10^6 m^3 s^-1). Thin vertical colorbars show the annual mean AMOC with respect to its climatological mean of 24 Sv. Thick black curve is the lowpass-filtered AMOC index after applying a Lanczos filter with 121 weights, with a cutoff period of 60 years. The AMOC index is defined as the maximum streamfunction in the region of 20°-70°N between 200 and 3000 m in the Atlantic. (b) and (c) are for climatological patterns of Eulerian-mean and eddy-induced AMOC (units: Sv), respectively. (d) and (e) are for regression patterns of Eulerian-mean and eddy-induced AMOC, respectively, with respect to the mean-normalized filtered AMOC index.
Figure 2  (a) Time series of anomalous vertical salinity gradient between subpolar upper ocean (above 1000 m) and deeper ocean (units: psu). Thin vertical colorbars show the annual mean anomalies. Thick black curve is the lowpass-filtered anomalies after applying a Lanczos filter with 121 weights, with a cutoff period of 60 years. Temporal evolutions of (b) annual mean anomalous potential density and the contributions due to (c) salinity change and (d) temperature change (units: kg · m⁻³). These variables are horizontally averaged over the subpolar Atlantic region defined in Fig. 3d.
Figure 3  (a) Scattering plots of anomalous salinity mixing (units: $psu \cdot Sv$) in the subpolar upper ocean (above 1000 m) versus the salinity difference anomaly between subpolar upper ocean and deeper ocean (units: $psu$). (b) and (c) show the horizontal and vertical components (units: $psu \cdot Sv$), respectively. The red curve shows the cubic regression result, and all these data are filtered with a cutoff period of 10 years. (d) Sea-surface density anomaly (units: $kg \cdot m^{-3}$) regressed onto the mean-normalized filtered AMOC index. Black curves encircle the region defined for the area-average region, and these curves are parallel to the grid lines of the ocean model. In (a), the expression for cubic regression $F_{mixing} = -285.15 \times (\Delta S')^3 - 2.51 \times (\Delta S')^2 - 1.28 \times (\Delta S') - 0.003$, and the explained variance is about 25%; the effective rank for the scaled Vandermonde coefficient matrix is 4, suggesting the cubic regression is significant.
Figure 4  Schematic diagrams of box models: (a) 4-box ocean model and (b) 3-box ocean model. Ocean boxes are denoted by ①, ②, ③, and ④. Boxes 1 and 4 represent the upper and lower layers of the tropical ocean, respectively; boxes 2 and 3, of the subpolar ocean, respectively. $D_1$ and $D_2$ are the depths of upper and lower ocean layers, respectively. $F_w$ is the net freshwater gain (loss) in the subpolar (tropical) ocean. $q$ represents the thermohaline circulation. With the extreme vertical mixing in the subpolar ocean, the 4-box model can be interpreted as a 3-box model.
Figure 5  Unstable oscillation in the 4-box model without the enhanced vertical mixing considered in the subpolar ocean. (a) Salinity anomalies $S_1', S_2'$, and $S_3'$ (units: psu); (b) salinity stratification in the subpolar ocean ($S_2' - S_3'$) (units: psu), the meridional salinity gradient ($S_1' - S_2'$) (units: psu), and AMOC anomaly ($q'$) (units: Sv). The vertical coordinate on the right side is for $q'$. (c) shows the terms in Eq. (7b), which are multiplied by $S_2'$ to represent the evolution of amplitude change (units: psu$^2$Sv). The linear closure parameter $\lambda = 12 \text{Sv} \cdot kg^{-1} m^3$, and its non-dimensional value $M = 0.285$. 

![Figure 5](THC_MultiCentennial_Theory_Fig_5.pdf)
Figure 6  Same as Fig. 5, but for self-sustained oscillations when considering the enhanced vertical mixing in the subpolar ocean. In (c), the units are $10^{-1} \text{psu}^2 \text{Sv}$.  

Figure; THC_MultiCentennial_Theory_Fig_6.pdf
Figure 7  Dependence of eigenvalue $\omega$ in the 3-box and 4-box models on the non-dimensional linear closure parameter $M$. (a) and (b) are for real and imaginary parts of the non-dimensional eigenvalue $\omega$, respectively. Black curves are for the 3-box model, which are obtained theoretically. Orange curves are for the 4-box model without the enhanced vertical mixing, which are obtained numerically. $\lambda_1, \lambda_2$, and $\lambda_C$ are dimensional values of $M$ for the 3-box model; $\lambda'_2$ and $\lambda'_C$ are for the 4-box model. $\lambda'_C = 11.45$ and $\lambda_C = 12.42 \, Sv \cdot kg^{-1} m^3$; and the corresponding $M = 0.272$ and 0.295, respectively.
**Figure 8** Oscillations of AMOC anomalies (units: Sv) in the 3-box and 4-box models with and without the enhanced vertical mixing. (a) M=0.25 ($\lambda < \lambda'_C$, region 1 in Fig. 9), showing decaying oscillations for both models. (b) M=0.285 ($\lambda'_C < \lambda < \lambda_C$, region 2 in Fig. 9), showing a decaying oscillation in the 3-box model, a self-sustained oscillation 4-box model with $k_m$ and a growing oscillation in the 4-box model without $k_m$. The vertical coordinate on the right side is for the 4-box model with $k_m$. (c) M=0.35 ($\lambda > \lambda_C$, region 3 in Fig. 9), showing growing oscillations in both models. $\lambda'_C$ and $\lambda_C$ are marked in Fig. 7a.
Figure 9  Relations of (a) period and (b) e-folding time of oscillation modes in the 4-box model without enhanced vertical mixing to the relative geometry of the ocean boxes (units: years). The abscissa and ordinate represent the volume fractions of upper ocean \((V_1 + V_2)/V\) and subpolar ocean \((V_2 + V_3)/V\), respectively. \(V\) is the total ocean volume. The grey region in (a) denotes the region of no oscillation modes. Solid and dashed orange curves show the location of critical values \(\lambda'_{C}\) in the 4-box model and \(\lambda_{C}\) in the 3-box model, respectively. The oscillations are decayed in region 1, self-sustained in region 2, and unstable in region 3 when considering the enhanced vertical mixing. The star denotes the standard geometry and mode for parameters listed in Table 1.
Figure 10  (a) Time series of AMOC anomaly (units: Sv) in the 4-box model when considering enhanced vertical mixing in the subpolar ocean, forced by stochastic freshwater flux. Here, $M=0.225 \ (\lambda = 9.45 \ Sv \cdot kg^{-1}m^3)$; and the 4-box model is in a stable regime (region 1 in Fig. 9). (b) and (c) are the power spectra (units: dB) of AMOC anomalies and the noise, respectively. (d) shows the ratios of the AMOC-spectrum to the stochastic freshwater-spectrum (units: dB) with peaks around 0.003 cpy (330 years). In (b), (c), and (d), thick curves represent ensemble mean, whose standard deviations are denoted by vertical short line segments. The black, green and orange lines show results under red noises whose damping timescale are 1 month, 1 year, and 10 years, respectively.
**Figure 11** Comparison between AMOC anomaly (units: $Sv$) obtained from time-forward integration of the original nonlinear Eq. (3) (orange) and corresponding linearized Eq. (7) (black). (a) is for oscillation without considering enhanced vertical mixing in the subpolar box, and (b), with enhanced vertical mixing.