Geophysical Research Letters

RESEARCH LETTER
10.1029/2022GL099800

Key Points:

- Multidecadal Atlantic Meridional Overturning Circulation (AMOC) variability exists in revised Stommel's Two-Box Model with advective time delay and coupled freshwater feedback included
- Regimes/periods of the AMOC delayed oscillator depend on the advective time lag of Arctic salinity signal and freshwater feedback strength
- The model suggests an important role of the Arctic salinity and delayed negative freshwater feedback in the multidecadal AMOC variability

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

X. Wei and R. Zhang,
xinyuew@princeton.edu;
rong.zhang@noaa.gov

Citation:

Wei, X., & Zhang, R. (2022). A simple conceptual model for the self-sustained multidecadal AMOC variability. *Geophysical Research Letters*, *49*, e2022GL099800. https://doi.org/10.1029/2022GL099800

Received 31 MAY 2022
Accepted 7 JUL 2022

A Simple Conceptual Model for the Self-Sustained Multidecadal AMOC Variability

X. Wei[1] and R. Zhang[1,2]

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA, [2]NOAA/OAR/GFDL, Princeton, NJ, USA

Abstract Multidecadal variability of Atlantic Meridional Overturning Circulation (AMOC) has been reconstructed by various proxies, simulated in climate models, and linked to multidecadal Arctic salinity variability. Here we construct a simple conceptual model to understand the two-way interactions of the Arctic with multidecadal AMOC variability through a delayed oscillator mechanism. We revise Stommel's Two-Box Model by including an advective time delay for the Arctic density/salinity anomalies to reach the subpolar North Atlantic and a coupled negative feedback between the AMOC and the freshwater flux entering the Arctic through atmosphere and/or sea ice responses. Self-sustained multidecadal AMOC oscillations exist in the revised Stommel's Two-Box model if the oceanic advective time delay is longer than the oscillation threshold, and the periods of the AMOC delayed oscillator depend crucially on this advective time delay. The coupled freshwater feedback provides additional delayed negative feedback and reduces the advective time delay threshold required for the oscillations.

Plain Language Summary Stommel's Two-Box Model provides a pioneering and powerful theoretical framework to study the mechanisms of steady Atlantic Meridional Overturning Circulation (AMOC) states and abrupt AMOC changes. However, Stommel's Two-Box Model itself does not include a self-sustained multidecadal AMOC oscillation solution. In our study, we revise Stommel's Two-Box model by considering the delayed oceanic advective time lag for the Arctic density/salinity anomalies to reach the subpolar North Atlantic and the dependence of the freshwater flux entering the Arctic on the AMOC strength. The revised Stommel's Two-Box Model is able to obtain the AMOC delayed oscillator at multidecadal timescales and suggests the important role of the Arctic salinity anomalies and associated delayed negative feedback in the multidecadal AMOC variability. The regimes and multidecadal periods of the AMOC delayed oscillator depend crucially on the advective time scale of the Arctic signal reaching the subpolar North Atlantic. The AMOC-related coupled freshwater feedback provides additional delayed negative feedback and reduces the threshold of the advective time delay needed for the multidecadal AMOC oscillations.

1. Introduction

The Atlantic Meridional Overturning Circulation (AMOC) plays a significant role in climate (e.g., Broecker et al., 1985; Manabe & Stouffer, 1997; Rahmstorf, 2002; R. Zhang et al., 2019). Multidecadal AMOC variability has been reconstructed over the modern period using various observed AMOC fingerprints/proxies (e.g., X. Chen & Tung, 2018; Fraser & Cunningham, 2021; Rossby et al., 2020; Yan et al., 2017; R. Zhang, 2007, 2008) and over the past several centuries using paleo records (Mjell et al., 2016). The directly observed AMOC decline from the RAPID program over the recent decade (Smeed et al., 2018) is shown to be part of the multidecadal AMOC variability (L. C. Jackson et al., 2016; Yan et al., 2017), which affects the observed multidecadal Labrador Sea water properties (Thomas & Zhang, 2022). Multidecadal variability has also been observed in the Arctic salinity (Polyakov et al., 2008). Using a high-resolution climate model constrained by the observed hydrographic climatology, R. Zhang and Thomas (2021) show that the Arctic Ocean (rather than the Greenland Sea) is the northern terminus of the AMOC, and the subpolar AMOC is sensitive to the salinity/density anomaly along the East Greenland Current (EGC). The observed Great Salinity Anomaly events with reduced upper-ocean salinity/density along the Arctic outflow (e.g., EGC) might have weakened the AMOC in the 1970s and 1980s (Belkin et al., 1998; Dickson et al., 1988; R. Zhang & Vallis, 2006).

Climate models also simulate the low-frequency AMOC variability, which has very different periods (from multidecadal to centennial) across different models (Keenlyside et al., 2016) and is closely related to the Arctic

salinity anomaly (T. Delworth et al., 1993; T. L. Delworth et al., 1997; Hawkins & Sutton, 2007; L. Jackson & Vellinga, 2013; Jiang et al., 2021; Jungclaus et al., 2005). It is unclear what physical processes/feedbacks in these models control or influence the different periods of the AMOC variability. Unlike the inter-annual El Niño-Southern Oscillation (ENSO) that has been explained through simple conceptual models, the mechanisms of the multidecadal AMOC variability and its two-way interactions with the Arctic salinity anomaly, as well as the factors affecting the periods and amplitudes of the multidecadal AMOC variability are not well understood from the theoretical perspective using simple conceptual models.

Stommel (1961) provides a pioneering and powerful theoretical framework to study the steady AMOC states and abrupt AMOC changes using a simple two-box model. However, Stommel's Two-Box Model itself does not have self-sustained multidecadal AMOC oscillation solutions. The revised Stommel's Three-Box Model (R. Zhang et al., 2002), Four-Box Model (Colin de Verdière, 2007), or Six-Box Model (Colin de Verdière et al., 2006) has been combined with Welander-type heat–salt oscillators (Welander, 1982; Whitehead et al., 2005) to explain ocean model simulated self-sustained millennial-scale AMOC relaxation oscillations (e.g., Colin de Verdière et al., 2006; Whitehead, 2018; Winton & Sarachik, 1993; R. Zhang et al., 2002) in the paleo context. The revised Stommel's Four-Box Model has been employed to explain the stochastic forced damped multidecadal AMOC oscillations (Griffies & Tziperman, 1995). In this study, we construct an AMOC delayed oscillator to understand the important role of the Arctic salinity anomaly in the multidecadal AMOC variability. We revise Stommel's Two-Box Model by including the oceanic advective time delay for the Arctic density/salinity anomalies to reach the subpolar North Atlantic and the coupled negative feedback between the AMOC and the freshwater flux entering the Arctic through the atmosphere and/or sea ice response. Both damped and self-sustained multidecadal AMOC oscillations can exist in the revised Stommel's Two-Box Model. The regimes and periods of the AMOC oscillations depend on the advective time delay and the strength of the coupled freshwater feedback.

2. Revised Stommel's Two-Box Model

The recent AMOC observations, that is, Overturning in the Subpolar North Atlantic Program (OSNAP) (F. Li, Lozier, Bacon, et al., 2021; F. Li, Lozier, Holliday, 2021; Lozier et al., 2017, 2019), provide important constraints to calibrate the modern climate steady state of our revised Stommel's Two-Box Model (Text S1 and Table S1 in Supporting Information S1). Starting from Stommel's Two-Box Model, we consider a revised two boxes of the ocean currents along the AMOC pathways separated by the OSNAP section: a low-latitude box of the AMOC inflow/outflow extending from the tropical North Atlantic to the OSNAP section and a high-latitude box of the AMOC inflow/outflow extending from the OSNAP section to the Arctic. We define the AMOC q of the revised two-box model as the maximum density-space AMOC across the OSNAP section (the boundary separating the two-boxes), which depends on the west-east density contrast across the section through both thermal wind and horizontal gyre contributions (R. Zhang & Thomas, 2021). The density around OSNAP western boundary is affected by the AMOC outflow from the Arctic (R. Zhang & Thomas, 2021), and the density around OSNAP eastern boundary is affected by the AMOC inflow along the Gulf Stream from the subtropics (Sutton & Allen, 1997) (Figure S1 in Supporting Information S1). Hence we simply assume the AMOC $q(t)$ is linearly related to the density difference between OSNAP western and eastern boundaries through a constant k, that is, $q(t) = k(\rho_H(t-\tau_H) - \rho_L(t-\tau_L))$. Here $\rho_H(t-\tau_H)$ is the density advected from the Arctic north of Greenland to OSNAP western boundary with the mean advective time delay τ_H, and $\rho_L(t-\tau_L)$ is the density advected from the subtropics to OSNAP eastern boundary with the mean advective time delay τ_L, respectively. These advective time delays are also included in the temperature and salinity equations when calculating ocean heat and salt transport between the two boxes across the OSNAP section. The original Stommel's Two-Box Model assumes a linear relationship between the AMOC and the instantaneous density difference between high- and low-latitude boxes. Here we include the advective time lags, that is, it takes a mean advective timescale τ_H (τ_L) for water properties (T, S, and ρ) in high-latitude (low-latitude) box to reach OSNAP western (eastern) boundary to affect the AMOC across the OSNAP section, thus the AMOC q and water properties (T, S, and ρ) at OSNAP western and eastern boundaries at time t depend on water properties at time $t-\tau_H$ and $t-\tau_L$ in each box, respectively.

The unrealistic surface boundary condition, that is, restoring salinity S to a prescribed value S^*, in the original Stommel's Two-Box Model is replaced with a net freshwater flux F added into the high-latitude box and removed from the low-latitude box. The freshwater flux F is usually assumed uncoupled to the AMOC for Stommel-type box-models, but here F includes both the steady state F_e and the anomaly F' which we assume is coupled linearly

with the AMOC strength anomaly, that is, $F' = c|q(t-\tau_C)|'$, where c and τ_C are the coupled feedback strength and time delay. The absolute sign only takes effect in the haline AMOC mode in the paleo context, and for the modern climate ($q > 0$) focused in this study, we have $F' = cq'(t-\tau_C)$. For example, a stronger AMOC can lead to enhanced poleward ocean heat transport (OHT) and heat/moisture released from the ocean into the high-latitude atmosphere, resulting in stronger atmospheric moisture transport into the Arctic, hence the river runoff into the Arctic increases linearly with the AMOC anomaly with a few-year time lag (Jungclaus et al., 2005). Meanwhile, the intensified AMOC and associated warmer high latitude temperature may increase the high-latitude atmospheric blocking and reduce the Arctic freshwater export (Lonita et al., 2016; Peings & Magnusdottir, 2014), and the AMOC-induced enhanced poleward OHT can cause more Arctic sea ice melting and less Arctic sea ice export (Jiang et al., 2021; Jungclaus et al., 2005; D. Li et al., 2018; R. Zhang, 2015). These processes result in a positive Arctic freshwater flux anomaly lagging the AMOC anomaly with a simplified time delay τ_C.

The dynamic equations for our revised two-box model become nonlinear delay differential equations (DDE) with multiple delays (τ_H, τ_L, τ_C):

$$\frac{dT_L}{dt} = \gamma_L\left(T_L^* - T_L\right) + \frac{|q|}{V_L}\left[T_H(t-\tau_H) - T_L(t-\tau_L)\right] - a{\rho'_L}^2 T'_L \tag{1}$$

$$\frac{dT_H}{dt} = \gamma_H\left(T_H^* - T_H\right) + \frac{|q|}{V_H}\left[T_L(t-\tau_L) - T_H(t-\tau_H)\right] - a{\rho'_H}^2 T'_H \tag{2}$$

$$\frac{dS_L}{dt} = \frac{S_0 F}{V_L} + \frac{|q|}{V_L}\left[S_H(t-\tau_H) - S_L(t-\tau_L)\right] - a{\rho'_L}^2 S'_L \tag{3}$$

$$\frac{dS_H}{dt} = -\frac{S_0 F}{V_H} + \frac{|q|}{V_H}\left[S_L(t-\tau_L) - S_H(t-\tau_H)\right] - a{\rho'_H}^2 S'_H \tag{4}$$

$$q = k(\rho_H\,(t-\tau_H) - \rho_L(t-\tau_L)) \tag{5}$$

$$\begin{aligned}&\rho_H(t-\tau_H) - \rho_L(t-\tau_L) = \\ &\rho_0(\beta_S(S_H\,(t-\tau_H) - S_L(t-\tau_L)) - \beta_T(T_H\,(t-\tau_H) - T_L(t-\tau_L)))\end{aligned} \tag{6}$$

$$F = F_e + F' = F_e + c|q(t-\tau_C)|' \tag{7}$$

where $T_L, S_L, \rho_L, V_L, T_L^*, \gamma_L$ and $T_H, S_H, \rho_H, V_H, T_H^*, \gamma_H$ are ocean temperature, salinity, density, volume, prescribed atmosphere temperature, and thermal restoring rate for low- and high-latitudes boxes respectively. Here S_0 and ρ_0 are the reference salinity and density, β_S and β_T are saline/thermal expansion coefficients, and the prime symbol ($'$) represents anomalies relative to the steady state. The nonlinear damping, that is, the last term in Equations 1–4, represents the nonlinear dissipation, for example, submesoscale/mesoscale mixing of the AMOC inflow/outflow temperature/salinity anomalies with the interior ocean, where the mixing coefficient is proportional to the square of density anomalies through a damping coefficient a (L. Chen and Young, 1995). See extended description and parameters of the revised two-box model in Text S1 and Table S1 in Supporting Information S1.

3. Results

3.1. Multidecadal AMOC Oscillations in the Revised Stommel's Two-Box Model

Self-sustained multidecadal AMOC oscillations are obtained both with and without the coupled freshwater feedback in the revised Stommel's Two-Box Model. Figure 1a shows examples of the numerically solved results at $\tau_H = 20$ years, $\tau_L = 5$ years, and $\tau_C = 5$ years with $c = 0.005$ and $c = 0$ respectively. When the coupled freshwater feedback is included ($c = 0.005$), the standard deviation of the anomalous freshwater flux F' (0.009 Sv) is similar to the observed variations in the Arctic freshwater flux (Carmack et al., 2016). The AMOC anomalies (Figures 1a and 1c) have a 10-year low-pass filtered standard deviation of ∼1.8 Sv, similar to the reconstructed estimates (Fraser & Cunningham, 2021; Yan et al., 2018). The oscillation period is ∼70 years, also consistent with the reconstructed multidecadal timescales of the AMOC variability (Fraser & Cunningham, 2021). The temperature/salinity/density anomalies in the two boxes have opposite phases and their amplitudes in the high-latitude box are much larger than in the low-latitude box (Figures 1d–1f) mainly due to the much larger volume of the low-latitude

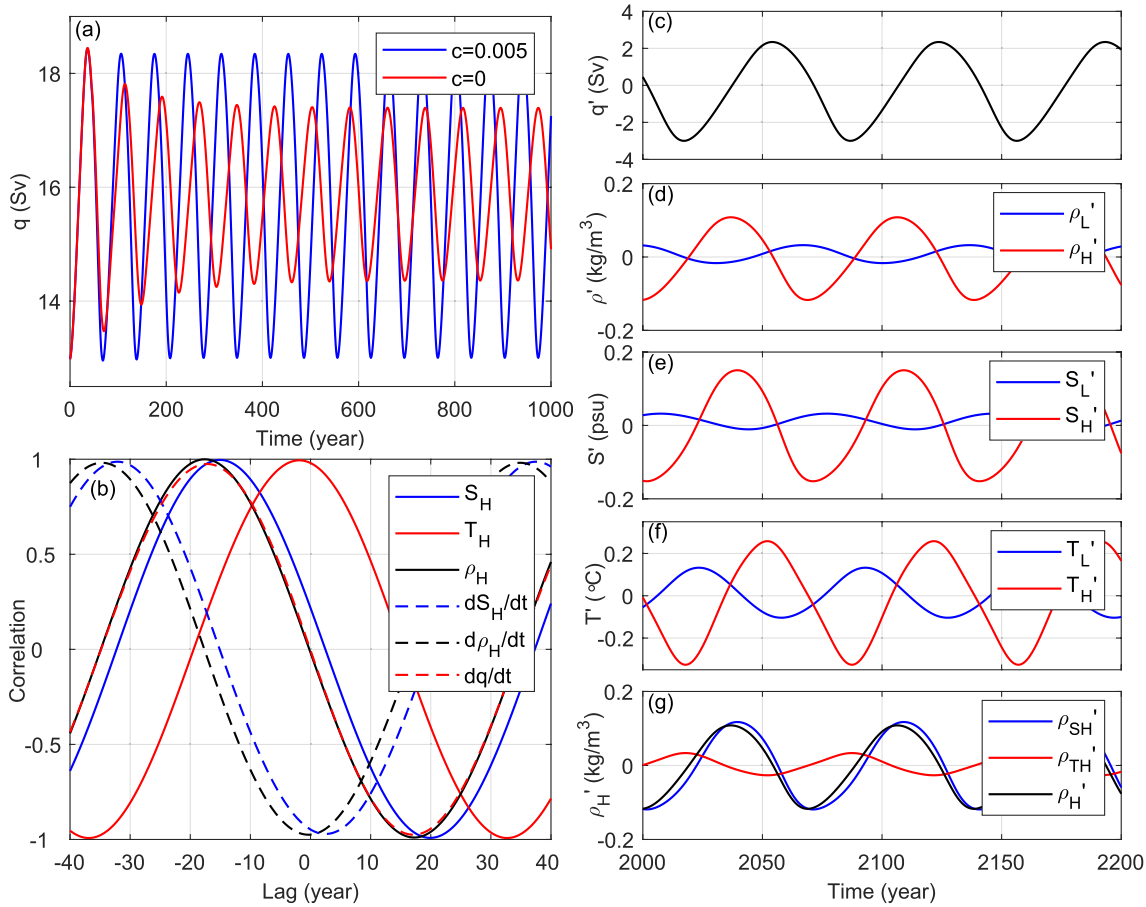

Figure 1. Numerical solutions of the revised Stommel's Two-Box Model. (a) Atlantic Meridional Overturning Circulation (AMOC) time series q over the model years 0–1000, $\tau_H = 20$ years, $\tau_L = 5$ years, $\tau_C = 5$ years, with (blue line, $c = 0.005$) or without (red line, $c = 0$) coupled freshwater feedback. (b–g) The example with $\tau_H = 20$ years, $\tau_L = 5$ years, $\tau_C = 5$ years and $c = 0.005$. (b) Correlations with the AMOC at different time lags. Positive lags means that the AMOC leads. (c–g) Anomalies of AMOC, high- and low-latitude density, salinity, temperature, and high-latitude density (black) with haline (blue) and thermal (red) components over the model years 2000–2200.

box compared to that of the high-latitude box (Table S1 in Supporting Information S1). Although we are in the thermally dominant modern climate in the steady state ($q > 0$), the high-latitude salinity anomaly S'_H is the major contributor to the high-latitude density anomaly ρ'_H both in amplitude and phase (Figure 1g), and therefore a major contributor to the AMOC anomaly q'.

The AMOC q lags ρ_H and S_H by ~18 years ($\sim\tau_H$) while is almost in phase with T_H (Figures 1b, 1c and 1f). The negative anomalies of the tendencies of ρ_H and S_H lag the positive ρ_H and S_H anomalies by ~1/4 period (~17 years), providing a delayed negative feedback (Figure 1b). The advective time delay τ_H plays a crucial role in the multidecadal AMOC oscillation. For example, the AMOC oscillation switches from self-sustained to damped when τ_H is shortened from 20 to 10 years (Figure S2 in Supporting Information S1) and disappears when $\tau_H = 0$. The coupled freshwater feedback also contributes to the self-sustained oscillation, but the self-sustained oscillation may still exist even when $c = 0$, with a smaller amplitude and longer period (Figure 1a).

The simple revised Stommel's Two-Box Model can obtain multidecadal AMOC oscillations of a wide range of amplitudes and periods under different parameter regimes (Figure 2). Self-sustained AMOC oscillations exist if τ_H is above certain thresholds (even without the coupled freshwater feedback, $c = 0$) and the oscillation periods depend crucially on τ_H (Figure 2). The oscillation period increases with τ_H and is insensitive to τ_L (Figures 2a and 2c). The oscillation amplitude increases with τ_H and also increases with τ_L when τ_H is large (Figures 2b and 2d). The oscillation period decreases with a larger c when τ_C is small (Figure 2g) and increases in general with a longer τ_C (Figure 2e). The oscillation amplitude also increases with c and is insensitive to τ_C when τ_H is

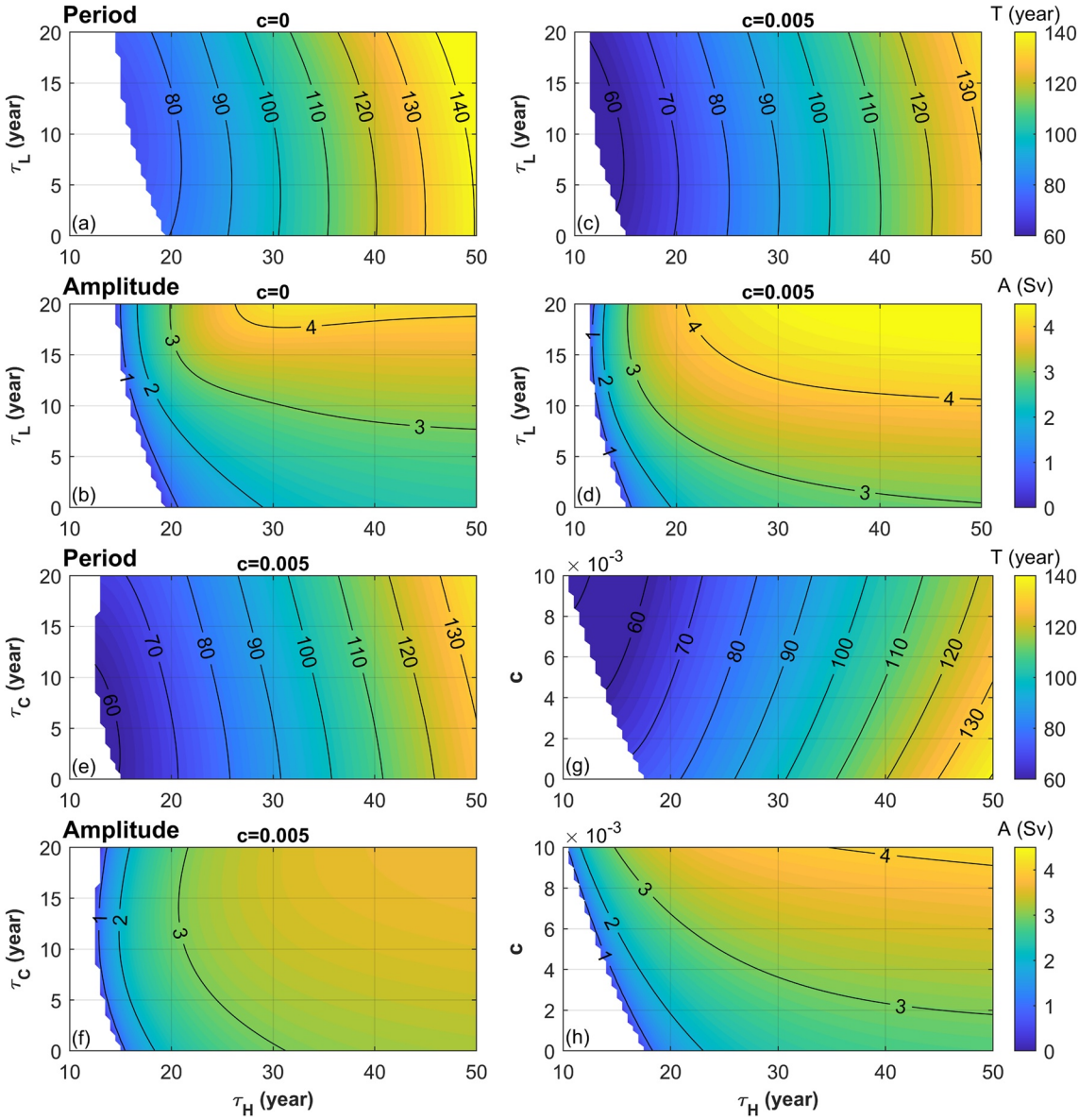

Figure 2. Numerical solutions of periods (a, c, e, and g) and amplitudes (i.e., peak values, (b, d, f, and h)) of the self-sustained oscillations under different parameters. (a–d) τ_H–τ_L diagrams at (a and b) $c = 0$ and (c and d) $c = 0.005$ with $\tau_C = 5$ years. (e and f) $\tau_H - \tau_C$ diagrams at $c = 0.005$ with $\tau_L = 5$ years. (g and h) $\tau_H - c$ diagrams with $\tau_L = \tau_C = 5$ years.

small (Figures 2f and 2h). The τ_H threshold required for the self-sustained oscillation decreases as c increases (Figure 2g), which will be explained in the simplified analytical solution in Section 3.3. When $c = 0$, the delayed negative feedback is mainly provided by the ocean-only process through the ocean salt transport (OST); when $c \neq 0$, the atmosphere-ocean (or ice-ocean) coupled negative freshwater feedback also plays an important role. The oscillation amplitudes decrease as the nonlinear damping coefficient a increases, whereas the oscillation periods are insensitive to a (Figure S3 in Supporting Information S1). It is the nonlinear damping on S'_H (not on S'_L, T'_H, T'_L), same order of magnitude as the anomalous OST and salt flux (Figure S4 in Supporting Information S1), that matters for the oscillation amplitudes.

Different periods of AMOC variability have been found in different climate models (Keenlyside et al., 2016), and each climate model simulates its own equivalent parameter regimes for τ_H and c. For instance, Jungclaus et al. (2005) found a 70–80 years AMOC variability in a climate model and there is linear coupled freshwater feedback in the AMOC-induced anomalous river runoff entering the Arctic ($c \approx 0.004$) with a time lag of a

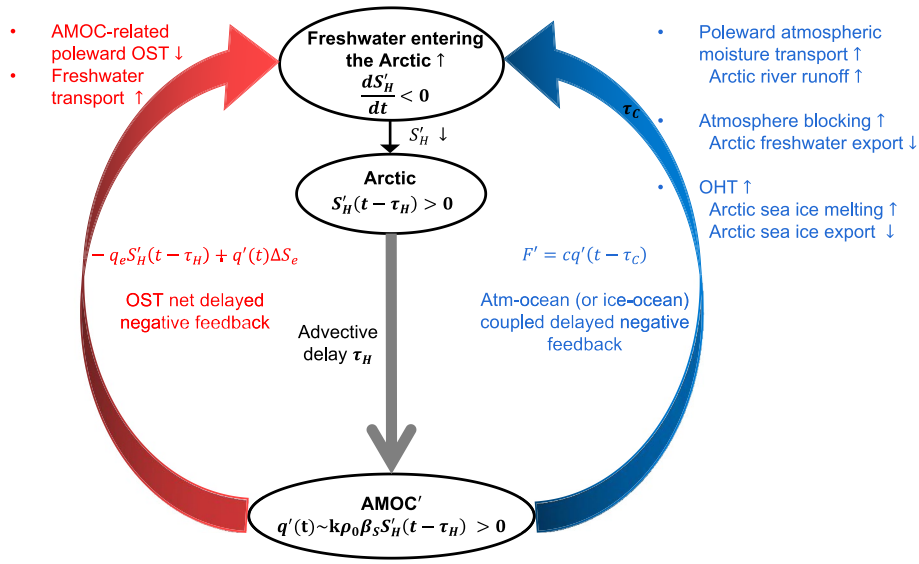

Figure 3. Schematic diagram of the Atlantic Meridional Overturning Circulation (AMOC) delayed oscillator, that is, the delayed feedback of the revised Stommel's Two-Box Model. A positive Arctic salinity anomaly takes an advective time delay τ_H (gray arrow) to induce a positive AMOC anomaly across the Overturning in the Subpolar North Atlantic Program section. Both the ocean salt transport (OST) anomaly (red arrow) and the AMOC-induced atmosphere-ocean (or ice-ocean) coupled freshwater feedback (blue arrow) increases the freshwater flux (F) entering the Arctic and causes a negative salinity tendency there, that is, a delayed negative feedback for the Arctic salinity anomaly.

few years. In this model, it takes ~20–30 years for the Arctic salinity anomaly to reach the subpolar North Atlantic ($\tau_H \approx$ ~20–30 years) and trigger AMOC oscillations with a 70–80 year period, consistent with our revised two-box model results. In another climate model (L. Jackson & Vellinga, 2013), although it also takes ~20–30 years for the Arctic salinity anomaly at the north coast of Greenland to affect the AMOC, the simulated multidecadal AMOC variability has a longer period, likely related to the weak coupled freshwater feedback ($c \approx 0$) in this model. Centennial AMOC variability is found in a recent climate model (Jiang et al., 2021), which takes over 40–50 years for the Arctic salinity anomaly at the north coast of Greenland to affect the AMOC ($\tau_H \approx$ ~40–50 years) and the coupled freshwater feedback is dominated by the AMOC-induced Arctic sea ice export anomaly. With a much longer τ_H, our revised two-box model also gives a centennial timescale for the AMOC oscillation. In contrast, the earlier coarse-resolution climate model simulates a 40–80 years damped AMOC oscillation (T. Delworth et al., 1993; T. L. Delworth et al., 1997), which might be related to a shorter advective time delay ($\tau_H \approx$ ~10–15 years) due to the coarse resolution with simplified land-sea distributions. The damped AMOC oscillation can be obtained in our revised two-box model with such a shorter τ_H. In summary, the revised two-box model could provide a novel theoretical framework/perspective to understand the mechanisms and different characteristics of the low-frequency AMOC variability simulated in various climate models.

3.2. Delayed Negative Feedbacks

The self-sustained multidecadal AMOC oscillation can be explained by the delayed negative feedback mechanism (Figure 3). A positive salinity anomaly in the high-latitude (Arctic) takes an advective time delay τ_H to reach OSNAP western boundary, where it induces a positive density difference anomaly across the OSNAP section (i.e., the boundary separating the two-boxes) and thus an intensified AMOC $q'(t) = k\left(\rho'_H(t-\tau_H) - \rho'_L(t-\tau_L)\right) \approx k\rho_0\beta_S S'_H(t-\tau_H)$ (Figures 1d and 1g). Changes across the OSNAP section provide a delayed negative feedback to the Arctic salinity tendency through two plausible pathways, the ocean-only OST feedback and the atmosphere-ocean (or ice-ocean) coupled freshwater feedback. The OST anomaly can be approximately decomposed into two linear terms, that is,

$$q'(t)\Delta S_e + q_e\left(S'_L(t-\tau_L) - S'_H(t-\tau_H)\right) \approx k\rho_0\beta_S S'_H(t-\tau_H)\,\Delta S_e - q_e S'_H(t-\tau_H) \quad (8)$$

where $\Delta S_e = S_{Le} - S_{He}$ is the steady-state salinity difference between the low- and high-latitude boxes and q_e is the steady-state AMOC across the OSNAP section. Here $S'_L(t-\tau_L)$ is neglected since it is much smaller than $S'_H(t-\tau_H)$ (Figure 1e). For the first term, the intensified AMOC transports more salinity from the low latitude to the Arctic, which is a delayed positive feedback. For the second term, the salinity export from the Arctic into the low-latitude across the OSNAP section increases therefore more freshwater accumulates in the high-latitude box and thus the Arctic, which is a delayed negative feedback. Since $q_e - k\rho_0\beta_S\Delta S_e$ is positive under the modern climate steady state, the second term has a larger amplitude and thus the total OST anomaly is negative and provides a net delayed negative feedback to the Arctic salinity anomaly.

For the coupled freshwater feedback, that is, the anomalous freshwater flux F' in response to the AMOC anomaly q' with a time delay simply represented as τ_C, several feedback processes might be involved as discussed in Section 2. These atmosphere-ocean (or ice-ocean) coupled feedbacks can contribute to the freshwater accumulation in the Arctic, providing a delayed negative feedback, that is, $F' = cq'(t-\tau_C) \approx ck\rho_0\beta_S S'_H(t-\tau_H-\tau_C)$. In summary, both the ocean-induced salinity/freshwater transport feedback and the coupled freshwater feedback can provide critical delayed negative feedbacks to sustain the multidecadal AMOC oscillation in the revised Stommel's Two-Box Model.

3.3. Analytical Solutions for the Simplified/Linearized System

To understand the fundamental oscillation mechanism of the revised Stommel's Two-Box Model (Figure 3), we simplify/linearize equations to solve for the analytical solution. We keep the linear terms related to S'_H around the steady state S_{He} in Equation 4 and ignore nonlinear perturbation terms. Since τ_H is more critical than τ_L and τ_C for the oscillation (Figure 2), we set $\tau_L = \tau_C = 0$ for simplicity in solving for the analytical solution. Since the anomalous density difference is dominated by the high-latitude density/salinity anomaly, the AMOC anomaly q' can be approximated as:

$$q'(t) \approx k\rho'_H(t-\tau_H) \approx k\rho_0\beta_S S'_H(t-\tau_H) \tag{9}$$

Combining Equation 9 with the linearized Equation 4, the linearized OST (Equation 8), and the expression for the coupled freshwater feedback $F' = cq'(t)$ (with $\tau_C = 0$), the final simplified/linearized equation for S'_H becomes a linear DDE with a single delay (τ_H):

$$\frac{dS'_H}{dt} = -\lambda S'_H(t-\tau_H) \tag{10}$$

where λ is the net delayed negative feedback strength:

$$\lambda = \frac{1}{V_H}(q_e - k\rho_0\beta_S\Delta S_e + k\rho_0\beta_S S_0 c) > 0 \tag{11}$$

With an exponential eigenvector solution $S'_H(t) = Ae^{Bt}$, where $B = \sigma + i\omega$ is the eigenvalue, A is the amplitude, σ is the growth rate, and ω is the oscillation frequency, there is a growing self-sustained oscillation if $\omega \neq 0$ and $\sigma > 0$, and the nonlinear damping term in the full equations will stabilize the oscillation if the amplitude is large. If $\omega \neq 0$ and $\sigma < 0$, then there is a damped oscillation converging toward the stable steady state S_{He}. With this exponential solution and Equation 10, we have

$$\sigma = -\lambda e^{-\tau_H\sigma}\cos(\tau_H\omega) \tag{12}$$

$$\omega = \lambda e^{-\tau_H\sigma}\sin(\tau_H\omega) \tag{13}$$

The threshold of the self-sustained oscillation requires $\sigma = 0$:

$$0 = -\lambda\cos(\tau_H\omega) \tag{14}$$

$$\omega = \lambda sin(\tau_H\omega) \tag{15}$$

Which gives

$$\tau_H\omega = \frac{\pi}{2} \tag{16}$$

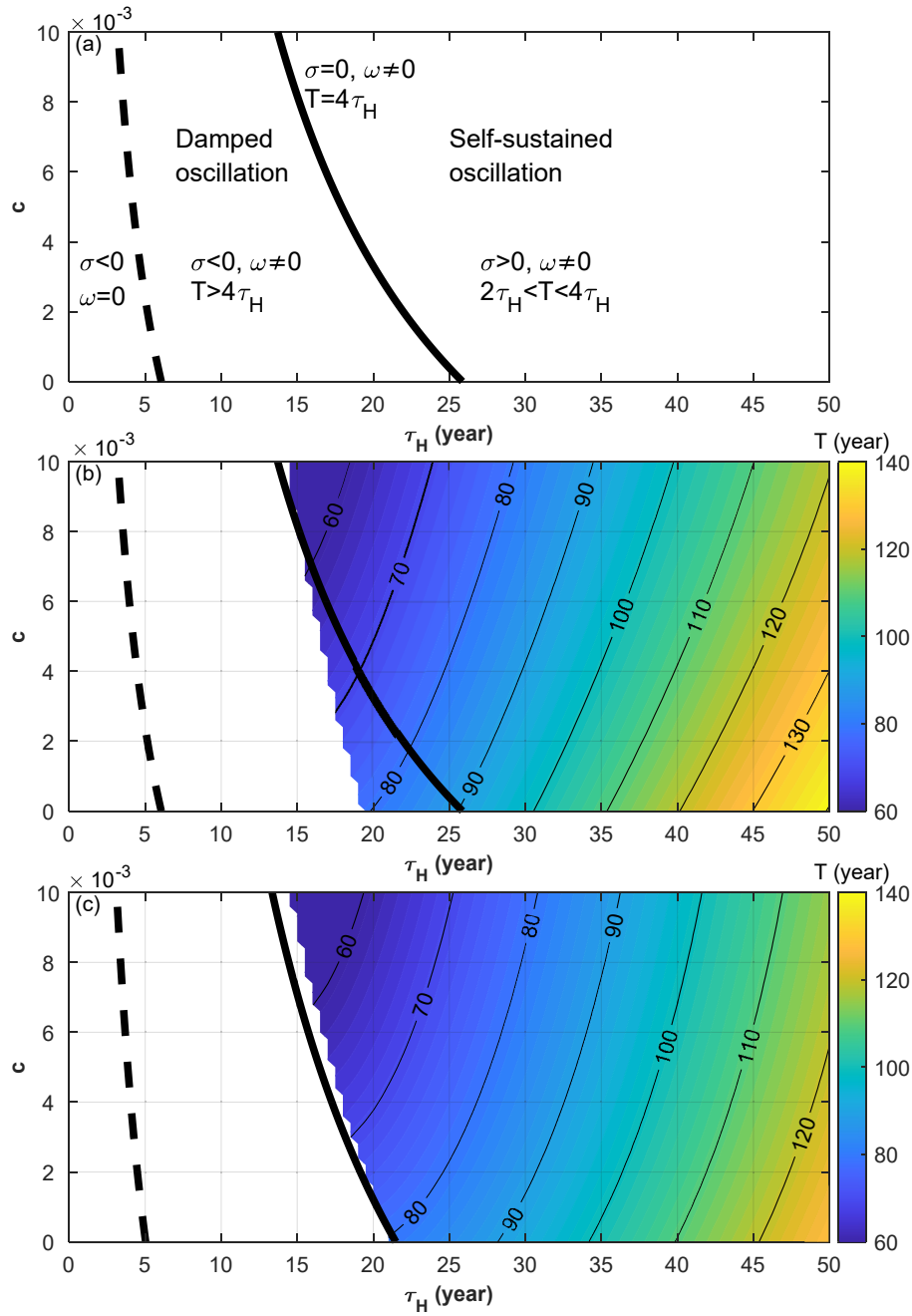

Figure 4. Comparing analytical and numerical solutions of the revised Stommel's Two-Box Model. (a) Damped and self-sustained oscillation regimes on the $\tau_H - c$ diagram (analytical solution). (b) Numerical solutions of the self-sustained oscillation period T (color) on the $\tau_H - c$ diagram when $\tau_L = \tau_C = 0$, overlapped with the analytical damped (dashed thick line) and self-sustained (solid thick line) oscillation thresholds shown in panel (a). (c) as in panel (b), but with $\tau_H = 0$, $a = 0$ for temperature terms in Equations 1, 2 and 5, overlapped with improved analytical solution by considering the T'_H contribution (Text S2 in Supporting Information S1).

$$\omega = \lambda \tag{17}$$

Therefore, we obtain the oscillation period $T = \frac{2\pi}{\omega} = 4\tau_H$ at the self-sustained oscillation threshold. When the oscillation decays, $\cos(\tau_H\omega) > 0$, $T > 4\tau_H$; When the oscillation grows, $\cos(\tau_H\omega) < 0$, $2\tau_H < T < 4\tau_H$ (Figure 4a). Using Equations 11, 16 and 17, we get

$$\tau_H = \frac{\pi}{2\omega} = \frac{\pi}{2\lambda} = \frac{\pi V_H}{2\left(q_e - k\rho_0\beta_S\Delta S_e + k\rho_0\beta_S S_0 c\right)} \tag{18}$$

Hence the threshold τ_H for the self-sustained oscillation decreases as c increases (thick solid line in Figures 4a and 4b), consistent with the numerical solution (Figure 4b). Mathematically, there are additional high-frequency solutions $\tau_H\omega = 2n\pi + \frac{\pi}{2}$ (with $n = 1, 2, 3, \ldots$) for Equations 14 and 15 and the corresponding threshold is $\tau_H = \frac{2n\pi + \pi/2}{\lambda}$, that is, λ would be unrealistically larger for a reasonable τ_H threshold and vice versa, which are unrealistic and not considered further.

If $\omega = 0$, Equation 12 becomes

$$\sigma = -\lambda e^{-\tau_H\sigma} < 0 \tag{19}$$

hence there is only exponential decay and no oscillation if real solutions $\sigma < 0$ can be found for Equation 19. By equaling the derivatives of σ at both sides of Equation 19, we get the threshold separating the pure exponential decay from the damped oscillation, that is, $1 = \tau_H\lambda e^{-\tau_H\sigma} = -\tau_H\sigma$, hence $\tau_H = \frac{1}{e\lambda}$ (thick dashed line in Figures 4a and 4b). In summary, with $\tau_H \geq \frac{\pi}{2\lambda}$, there is self-sustained oscillation; with $\frac{\pi}{2\lambda} > \tau_H > \frac{1}{e\lambda}$, there is damped oscillation and no real solution for Equation 19; with $\tau_H \leq \frac{1}{e\lambda}$, negative real solutions for Equation 19 exist, that is, there is only exponential decay and no oscillation.

The analytical threshold τ_H for the self-sustained oscillation matches well with the numerical solution when c is large, but is overestimated by a few years when c is small (Figure 4b). This is mainly because Equation 9 neglects the high-latitude temperature contribution ($-k\rho_0\beta_T T'_H(t - \tau_H)$) to the AMOC anomaly q', and thus Equations 10 and 11 underestimate the net delayed negative feedback strength λ, resulting in a larger analytical threshold τ_H. An example of improving the analytical solution by considering the T'_H contribution is shown in Text S2 in Supporting Information S1 and Figure 4c. When c is large enough, the threshold is also strongly affected by the coupled freshwater feedback (Equation 18) and thus insensitive to neglecting the T'_H contribution (Figure 4b).

4. Conclusion and Discussion

In summary, our simple conceptual model, a revised Stommel's Two-Box Model, provides a novel theoretical framework/perspective to understand the mechanisms of reconstructed and simulated multidecadal AMOC variability, its two-way interactions with the Arctic salinity variability, and the different variability periods simulated in climate models. When the advective time delay and coupled freshwater feedback are considered, self-sustained multidecadal AMOC oscillations are possible solutions in the revised Stommel's Two-Box Model calibrated by the OSNAP observations under modern climate. The key factors affecting the self-sustained AMOC delayed oscillator are the advective time scale for the Arctic density/salinity anomaly to reach the subpolar North Atlantic and the coupled freshwater feedback strength. Observational-based estimates are needed to reduce uncertainties/biases of these key factors in climate models in future studies.

The revised Stommel's Two-Box Model is too simple to include many complicated processes. In reality other factors including atmosphere/ocean stochastic noise could also contribute to multidecadal AMOC variability. Nevertheless, we keep the conceptual model as simple as possible and it suggests that the Arctic salinity anomaly and associated delayed negative freshwater feedback play an important role in the multidecadal AMOC variability. This is the first time that self-sustained multidecadal AMOC oscillations are obtained in a simple revised Stommel's Two-Box Model. Previous studies identified the tropical Pacific oceanic wave propagation as a key delay mechanism for the inter-annual ENSO delayed oscillator (e.g., Battisti & Hirst, 1989; Suarez & Schopf, 1988). Here for the multidecadal AMOC delayed oscillator, it is the oceanic advection of the salinity anomaly from the Arctic to the subpolar North Atlantic that provides a key delay mechanism.

Climate models often underestimate the amplitude of multidecadal AMOC variability (e.g., Kim et al., 2018; Yan et al., 2018), which might be related to the underestimated Arctic salinity anomalies in climate models (Rosenblum et al., 2021). Recent observations show significant negative salinity anomalies in the Arctic (e.g., Rosenblum et al., 2021; J. Zhang et al., 2021). Monitoring the potential downstream propagations of Arctic salinity anomalies would be valuable for predicting the timing and amplitude of future AMOC changes.

Data Availability Statement

The data used to calibrate the revised Stommel's Two-Box Model under modern climate is from the OSNAP observations (F. Li, Lozier, Bacon, et al., 2021; F. Li, Lozier, Holliday, 2021, https://www.o-snap.org/data-access/). The numerical solutions of the revised Stommel's Two-Box Model Equations 1–7 are solved with the RK-4 scheme using the widely available software Matlab.

Acknowledgments
Mitch Bushuk and Brandon Reichl are acknowledged for their comments and suggestions on an internal review of the manuscript. Robert Hallberg and Ching-Yao Lai are acknowledged for helpful comments on this work. We thank one anonymous reviewer, KK Tung, and Jack Whitehead for their constructive feedback on the original submission. X. Wei is supported by Princeton University Graduate School and CIMES Task II funds for graduate research under award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. R. Zhang is supported by GFDL base funding. We acknowledge GFDL resources made available for this research. R. Zhang conceived the study, designed the revised Stommel's Two-Box Model for the AMOC delayed oscillator mechanism, and directed main (numerical and analytical) approaches. X. Wei conducted the numerical modeling/experiments and analyses of the revised Stommel's Two-Box Model. The authors discussed the results/interpretations and wrote the paper together.

References

Battisti, D. S., & Hirst, A. C. (1989). Interannual variability in a tropical atmosphere–ocean model: Influence of the basic state, ocean geometry and nonlinearity. *Journal of the Atmospheric Sciences*, *46*(12), 1687–1712. https://doi.org/10.1175/1520-0469(1989)046<1687:iviata>2.0.co;2

Belkin, I. M., Levitus, S., Antonov, J., & Malmberg, S. A. (1998). "Great salinity anomalies" in the North Atlantic. *Progress in Oceanography*, *41*(1), 1–68. https://doi.org/10.1016/s0079-6611(98)00015-9

Broecker, W. S., Peteet, D. M., & Rind, D. (1985). Does the ocean–atmosphere system have more than one stable mode of operation? *Nature*, *315*(6014), 21–26. https://doi.org/10.1038/315021a0

Carmack, E. C., Yamamoto-Kawai, M., Haine, T. W., Bacon, S., Bluhm, B. A., Lique, C., et al. (2016). Freshwater and its role in the Arctic Marine System: Sources, disposition, storage, export, and physical and biogeochemical consequences in the Arctic and global oceans. *Journal of Geophysical Research: Biogeosciences*, *121*(3), 675–717. https://doi.org/10.1002/2015jg003140

Chen, L., & Young, W. R. (1995). Density compensated thermohaline gradients and diapycnal fluxes in the mixed layer. *Journal of Physical Oceanography*, *25*(12), 3064–3075. https://doi.org/10.1175/1520-0485(1995)025<3064:dctgad>2.0.co;2

Chen, X., & Tung, K. K. (2018). Global surface warming enhanced by weak Atlantic overturning circulation. *Nature*, *559*(7714), 387–391. https://doi.org/10.1038/s41586-018-0320-y

Colin de Verdière, A. (2007). A simple model of millennial oscillations of the thermohaline circulation. *Journal of Physical Oceanography*, *37*(5), 1142–1155. https://doi.org/10.1175/jpo3056.1

Colin de Verdière, A., Ben Jelloul, M., & Sévellec, F. (2006). Bifurcation structure of thermohaline millennial oscillations. *Journal of Climate*, *19*(22), 5777–5795. https://doi.org/10.1175/jcli3950.1

Delworth, T., Manabe, S., & Stouffer, R. J. (1993). Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *Journal of Climate*, *6*(11), 1993–2011. https://doi.org/10.1175/1520-0442(1993)006<1993:ivottc>2.0.co;2

Delworth, T. L., Manabe, S., & Stouffer, R. J. (1997). Multidecadal climate variability in the Greenland Sea and surrounding regions: A coupled model simulation. *Geophysical Research Letters*, *24*(3), 257–260. https://doi.org/10.1029/96gl03927

Dickson, R. R., Meincke, J., Malmberg, S. A., & Lee, A. J. (1988). The "great salinity anomaly" in the northern North Atlantic 1968–1982. *Progress in Oceanography*, *20*(2), 103–151. https://doi.org/10.1016/0079-6611(88)90049-3

Fraser, N. J., & Cunningham, S. A. (2021). 120 Years of AMOC variability reconstructed from observations using the Bernoulli inverse. *Geophysical Research Letters*, *48*(18), e2021GL093893. https://doi.org/10.1029/2021gl093893

Griffies, S. M., & Tziperman, E. (1995). A linear thermohaline oscillator driven by stochastic atmospheric forcing. *Journal of Climate*, *8*(10), 2440–2453. https://doi.org/10.1175/1520-0442(1995)008<2440:altodb>2.0.co;2

Hawkins, E., & Sutton, R. (2007). Variability of the Atlantic thermohaline circulation described by three-dimensional empirical orthogonal functions. *Climate Dynamics*, *29*(7), 745–762. https://doi.org/10.1007/s00382-007-0354-6

Jackson, L., & Vellinga, M. (2013). Multidecadal to centennial variability of the AMOC: HadCM3 and a perturbed physics ensemble. *Journal of Climate*, *26*(7), 2390–2407. https://doi.org/10.1175/jcli-d-11-00601.1

Jackson, L. C., Peterson, K. A., Roberts, C. D., & Wood, R. A. (2016). Recent slowing of Atlantic overturning circulation as a recovery from earlier strengthening. *Nature Geoscience*, *9*(7), 518–522. https://doi.org/10.1038/ngeo2715

Jiang, W., Gastineau, G., & Codron, F. (2021). Multicentennial variability driven by salinity exchanges between the Atlantic and the Arctic Ocean in a coupled climate model. *Journal of Advances in Modeling Earth Systems*, *13*(3), e2020MS002366. https://doi.org/10.1029/2020ms002366

Jungclaus, J. H., Haak, H., Latif, M., & Mikolajewicz, U. (2005). Arctic–North Atlantic interactions and multidecadal variability of the meridional overturning circulation. *Journal of Climate*, *18*(19), 4013–4031. https://doi.org/10.1175/jcli3462.1

Keenlyside, N. S., Ba, J., Mecking, J., Omrani, N. E., Latif, M., Zhang, R., & Msadek, R. (2016). North Atlantic multi-decadal variability—Mechanisms and predictability. In *Climate change: Multidecadal and beyond* (pp. 141–157).

Kim, W. M., Yeager, S., Chang, P., & Danabasoglu, G. (2018). Low-frequency North Atlantic climate variability in the Community Earth System Model large ensemble. *Journal of Climate*, *31*(2), 787–813. https://doi.org/10.1175/jcli-d-17-0193.1

Li, D., Zhang, R., & Knutson, T. (2018). Comparison of Mechanisms for low-frequency variability of summer Arctic sea ice in three coupled models. *Journal of Climate*, *31*(3), 1205–1226. https://doi.org/10.1175/jcli-d-16-0617.1

Li, F., Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., de Jong, M. F., et al. (2021). Subpolar North Atlantic western boundary density anomalies and the Meridional Overturning Circulation. *Nature Communications*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-23350-2

Li, F., Lozier, M. S., Holliday, N. P., Johns, W. E., Le Bras, I. A., Moat, B. I., et al. (2021). Observation-based estimates of heat and freshwater exchanges from the subtropical North Atlantic to the Arctic. *Progress in Oceanography*, *197*, 102640. https://doi.org/10.1016/j.pocean.2021.102640

Lonita, M., Scholz, P., Lohmann, G., Dima, M., & Prange, M. (2016). Linkages between atmospheric blocking, sea ice export through Fram Strait and the Atlantic Meridional Overturning Circulation. *Scientific Reports*, *6*(1), 1–10. https://doi.org/10.1038/srep32881

Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., De Jong, M. F., De Steur, L., et al. (2017). Overturning in the Subpolar North Atlantic Program: A new international ocean observing system. *Bulletin of the American Meteorological Society*, *98*(4), 737–752. https://doi.org/10.1175/bams-d-16-0057.1

Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., et al. (2019). A sea change in our view of overturning in the subpolar North Atlantic. *Science*, *363*(6426), 516–521. https://doi.org/10.1126/science.aau6592

Manabe, S., & Stouffer, R. J. (1997). Coupled ocean-atmosphere model response to freshwater input: Comparison to Younger Dryas event. *Paleoceanography*, *12*(2), 321–336. https://doi.org/10.1029/96pa03932

Mjell, T. L., Ninnemann, U. S., Kleiven, H. F., & Hall, I. R. (2016). Multidecadal changes in Iceland Scotland Overflow Water vigor over the last 600 years and its relationship to climate. *Geophysical Research Letters*, *43*(5), 2111–2117. https://doi.org/10.1002/2016gl068227

Peings, Y., & Magnusdottir, G. (2014). Forcing of the wintertime atmospheric circulation by the multidecadal fluctuations of the North Atlantic Ocean. *Environmental Research Letters*, *9*(3), 034018. https://doi.org/10.1088/1748-9326/9/3/034018

Polyakov, I. V., Alexeev, V. A., Belchansky, G. I., Dmitrenko, I. A., Ivanov, V. V., Kirillov, S. A., et al. (2008). Arctic Ocean freshwater changes over the past 100 years and their causes. *Journal of Climate*, *21*(2), 364–384. https://doi.org/10.1175/2007jcli1748.1

Rahmstorf, S. (2002). Ocean circulation and climate during the past 120, 000 years. *Nature*, *419*(6903), 207–214. https://doi.org/10.1038/nature01090

Rosenblum, E., Fajber, R., Stroeve, J. C., Gille, S. T., Tremblay, L. B., & Carmack, E. C. (2021). Surface salinity under transitioning ice cover in the Canada Basin: Climate model biases linked to vertical distribution of fresh water. *Geophysical Research Letters*, *48*(21), e2021GL094739. https://doi.org/10.1029/2021gl094739

Rossby, T., Chafik, L., & Houpert, L. (2020). What can hydrography tell us about the strength of the Nordic Seas MOC over the last 70 to 100 years? *Geophysical Research Letters*, *47*(12), e2020GL087456. https://doi.org/10.1029/2020gl087456

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., et al. (2018). The North Atlantic Ocean is in a state of reduced overturning. *Geophysical Research Letters*, *45*(3), 1527–1533. https://doi.org/10.1002/2017gl076350

Stommel, H. (1961). Thermohaline convection with two stable regimes of flow. *Tellus*, *13*(2), 224–230. https://doi.org/10.3402/tellusb.v13i2.12985

Suarez, M. J., & Schopf, P. S. (1988). A delayed action oscillator for ENSO. *Journal of the Atmospheric Sciences*, *45*(21), 3283–3287. https://doi.org/10.1175/1520-0469(1988)045<3283:adaofe>2.0.co;2

Sutton, R. T., & Allen, M. R. (1997). Decadal predictability of North Atlantic sea surface temperature and climate. *Nature*, *388*(6642), 563–567. https://doi.org/10.1038/41523

Thomas, M. D., & Zhang, R. (2022). Two sources of deep decadal variability in the central Labrador Sea open-ocean convection region. *Geophysical Research Letters*, *49*(11), e2022GL098825. https://doi.org/10.1029/2022gl098825

Welander, P. (1982). A simple heat–salt oscillator. *Dynamics of Atmospheres and Oceans*, *6*(4), 233–242. https://doi.org/10.1016/0377-0265(82)90030-6

Whitehead, J. A. (2018). Transitions and oscillations in a thermohaline deep ocean circulation model. *Geophysical & Astrophysical Fluid Dynamics*, *112*(5), 321–344. https://doi.org/10.1080/03091929.2018.1530351

Whitehead, J. A., Te Raa, L., Tozuka, T., Keller, J. B., & Bradley, K. (2005). Laboratory observations and simple models of slow oscillations in cooled salt-stratified bodies. *Tellus A: Dynamic Meteorology and Oceanography*, *57*(5), 798–809. https://doi.org/10.1111/j.1600-0870.2005.00150.x

Winton, M., & Sarachik, E. S. (1993). Thermohaline oscillations induced by strong steady salinity forcing of ocean general circulation models. *Journal of Physical Oceanography*, *23*(7), 1389–1410. https://doi.org/10.1175/1520-0485(1993)023<1389:toibss>2.0.co;2

Yan, X., Zhang, R., & Knutson, T. R. (2017). The role of Atlantic overturning circulation in the recent decline of Atlantic major hurricane frequency. *Nature Communications*, *8*(1), 1–8. https://doi.org/10.1038/s41467-017-01377-8

Yan, X., Zhang, R., & Knutson, T. R. (2018). Underestimated AMOC variability and implications for AMV and predictability in CMIP models. *Geophysical Research Letters*, *45*(9), 4319–4328. https://doi.org/10.1029/2018gl077378

Zhang, J., Weijer, W., Steele, M., Cheng, W., Verma, T., & Veneziani, M. (2021). Labrador Sea freshening linked to Beaufort Gyre freshwater release. *Nature Communications*, *12*(1), 1–8. https://doi.org/10.1038/s41467-021-21470-3

Zhang, R. (2007). Anticorrelated multidecadal variations between surface and subsurface tropical North Atlantic. *Geophysical Research Letters*, *34*(12), L12713. https://doi.org/10.1029/2007gl030225

Zhang, R. (2008). Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *35*(20), L20705. https://doi.org/10.1029/2008gl035463

Zhang, R. (2015). Mechanisms for low-frequency variability of summer Arctic sea ice extent. *Proceedings of the National Academy of Sciences*, *112*(15), 4570–4575. https://doi.org/10.1073/pnas.1422296112

Zhang, R., Follows, M., & Marshall, J. (2002). Mechanisms of thermohaline mode switching with application to warm equable climates. *Journal of Climate*, *15*(15), 2056–2072. https://doi.org/10.1175/1520-0442(2002)015<2056:motmsw>2.0.co;2

Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y. O., Marsh, R., Yeager, S. G., et al. (2019). A review of the role of the Atlantic meridional overturning circulation in Atlantic multidecadal variability and associated climate impacts. *Reviews of Geophysics*, *57*(2), 316–375. https://doi.org/10.1029/2019rg000644

Zhang, R., & Thomas, M. (2021). Horizontal circulation across density surfaces contributes substantially to the long-term mean northern Atlantic Meridional Overturning Circulation. *Communications Earth & Environment*, *2*(1), 1–12. https://doi.org/10.1038/s43247-021-00182-y

Zhang, R., & Vallis, G. K. (2006). Impact of great salinity anomalies on the low-frequency variability of the North Atlantic climate. *Journal of Climate*, *19*(3), 470–482. https://doi.org/10.1175/jcli3623.1

References From the Supporting Information

Born, A., Mignot, J., & Stocker, T. F. (2015). Multiple equilibria as a possible mechanism for decadal variability in the North Atlantic Ocean. *Journal of Climate*, *28*(22), 8907–8922. https://doi.org/10.1175/jcli-d-14-00813.1

Huang, R. X., Luyten, J. R., & Stommel, H. M. (1992). Multiple equilibrium states in combined thermal and saline circulation. *Journal of Physical Oceanography*, *22*(3), 231–246. https://doi.org/10.1175/1520-0485(1992)022<0231:mesict>2.0.co;2

Straneo, F. (2006). On the connection between dense water formation, overturning, and poleward heat transport in a convective basin. *Journal of Physical Oceanography*, *36*(9), 1822–1840. https://doi.org/10.1175/jpo2932.1