

1 **Transient and equilibrium responses of the Atlantic overturning circulation**  
2 **to warming in coupled climate models: the role of temperature and salinity**

3 David B. Bonan,<sup>a</sup> Andrew F. Thompson,<sup>a</sup> Emily R. Newsom,<sup>b</sup> Shantong Sun,<sup>a</sup> Maria  
4 Rugestein<sup>c</sup>

5 <sup>a</sup> *California Institute of Technology, Pasadena, California*

6 <sup>b</sup> *University of Oxford, Oxford, United Kingdom*

7 <sup>c</sup> *Colorado State University, Fort Collins, Colorado*

8 *Corresponding author:* David B. Bonan, [dbonan@caltech.edu](mailto:dbonan@caltech.edu)

9 ABSTRACT: The long-term response of the Atlantic meridional overturning circulation (AMOC)  
10 to climate change remains poorly understood, in part due to the computational expense associated  
11 with running atmosphere-ocean general circulation models (GCMs) to equilibrium. Here, we  
12 use a collection of millennial-length GCM simulations to examine the transient and equilibrium  
13 responses of the AMOC to an abrupt quadrupling of atmospheric carbon-dioxide. We find that  
14 GCMs consistently simulate an AMOC weakening during the first century but exhibit diverse  
15 behaviors over longer time scales, showing different recovery levels. To explain the AMOC behavior  
16 we use a thermal-wind expression, which links the overturning circulation to the meridional density  
17 difference between deep-water formation regions and the Atlantic basin. Using this expression,  
18 we attribute the evolution of the AMOC on different timescales to changes in temperature and  
19 salinity in distinct regions. The initial shoaling and weakening occurs on centennial timescales  
20 and is attributed to a warming of the deep-water formation region. The partial recovery occurs  
21 over the next few centuries, and is linked to a simultaneous warming of the Atlantic basin and a  
22 high-latitude salinity anomaly from the subtropical Atlantic that reduces subsurface stratification  
23 and reinvigorates deep-water formation. GCMs that exhibit prolonged AMOC weakening tend  
24 to have smaller high-latitude salinity anomalies, likely due to increased Arctic sea-ice loss and  
25 other high-latitude freshwater changes. After multiple millennia, the AMOC in some GCMs is  
26 stronger than the initial state due warming of the low-latitude Atlantic. These results highlight the  
27 importance of considering high-latitude freshwater changes when examining the past and future  
28 evolution of the AMOC evolution on long timescales.

29 SIGNIFICANCE STATEMENT: The long-term response of the ocean’s overturning circulation  
30 to warming remains poorly understood largely because it is expensive to run state-of-the-art  
31 climate models for multiple millennia. This study makes use of a unique collection of millennial-  
32 length climate simulations from different climate models to examine the response of the Atlantic  
33 overturning circulation to warming on long timescales. We find that climate models consistently  
34 simulate a weakening of the Atlantic overturning circulation during the first century but disagree on  
35 long-term changes, showing different recovery levels of the Atlantic overturning circulation. Using  
36 a simple expression, which emulates the evolution of the Atlantic overturning circulation in climate  
37 models, we find that climate models with little-to-no recovery tend to have a small North Atlantic  
38 salinity anomaly while climate models with a stronger recovery tend to have a large North Atlantic  
39 salinity anomaly. These results highlight the importance of monitoring high-latitude freshwater  
40 sources throughout the 21st century and considering the relative role of temperature and salinity  
41 changes when examining the future and past evolution of the Atlantic overturning circulation on  
42 long timescales.

## 43 **1. Introduction**

44 The ocean’s global overturning circulation plays a fundamental role in shaping Earth’s climate  
45 through the redistribution of heat and freshwater between low and high latitudes and different ocean  
46 basins. The branch of the overturning circulation that is localized to the Atlantic basin, referred to  
47 as the Atlantic meridional overturning circulation (AMOC), plays a unique role in that it transports  
48 heat northward across all latitudes (Ganachaud and Wunsch 2003) and ventilates the upper ~2000  
49 meters of the ocean (Buckley and Marshall 2016). The AMOC is frequently invoked as a major  
50 player in past climatic changes (Broecker 1997, 1998, 2003; Ferrari et al. 2014) and is thought to  
51 play a large role in setting features of today’s climate, such as shifting the peak of tropical rainfall to  
52 the Northern Hemisphere (Frierson et al. 2013; Marshall et al. 2014; Buckley and Marshall 2016).  
53 Paleoclimate records, based on both  $\delta^{13}\text{C}$  and Cd/Ca ratios in benthic foraminifera, suggest the  
54 AMOC was shallower and potentially weaker at the Last Glacial Maximum (LGM; Curry and Oppo  
55 2005; Lynch-Stieglitz et al. 2007; Burke et al. 2015), although this is still debated as these proxies  
56 may better reflect changes in the distribution of nutrients (Gebbie 2014). Comprehensive general  
57 circulation models (GCMs) also indicate that the AMOC can exist in a range of configurations

58 (Weber et al. 2007; Cheng et al. 2013; Weijer et al. 2020), yet a mechanistic understanding of how  
59 the AMOC transitions between these states is less clear. In particular, the timescales over which  
60 the AMOC might weaken and subsequently recover, and the relative importance of high- versus  
61 low-latitude processes in controlling these responses remain under debate.

62 In response to increased greenhouse-gas concentrations, GCMs consistently predict a weakening  
63 of the AMOC over the 21st century (Gregory et al. 2005; Weaver et al. 2007; Cheng et al. 2013;  
64 Kostov et al. 2014; Liu et al. 2020; Weijer et al. 2020). The weakening of the AMOC is thought to  
65 arise from a combination of increased freshwater fluxes and weakened surface heat loss in the North  
66 Atlantic (Dixon et al. 1999; Schmittner et al. 2005; Gregory et al. 2005; Stouffer et al. 2006; Weaver  
67 et al. 2007; Bakker et al. 2016; Maroon et al. 2018), both of which reduce the formation of North  
68 Atlantic Deep Water (NADW) through positive surface buoyancy flux anomalies. Advection of  
69 heat and salinity anomalies into the North Atlantic can also affect the AMOC by altering conditions  
70 that support deep convection and the depth to which NADW penetrates. For instance, studies have  
71 also attributed a weakening of the AMOC in GCMs to Arctic sea ice loss (Sévellec et al. 2017) and  
72 subsurface warming of the North Atlantic (Haskins et al. 2020; Levang and Schmitt 2020), which  
73 both increase ocean stratification and inhibit deep convection.

74 Beyond the 21st century, the response of the AMOC to rising greenhouse-gas concentrations  
75 is less clear. From a logistical standpoint, the computational expense of running a GCM to  
76 equilibrium, which typically takes multiple millennia, has limited the number of studies that  
77 consider millennial or longer timescales. But the few GCMs that have generated millennial-length  
78 climate simulations have also produced divergent AMOC responses, including both recovered and  
79 collapsed states. For example, Manabe and Stouffer (1994) and Stouffer and Manabe (2003) found  
80 that in response to an abrupt increase in atmospheric carbon-dioxide, the AMOC weakens initially  
81 and then recovers to its initial strength after approximately a thousand years. Other studies, which  
82 have focused only on changes to freshwater forcing, have also found similar recoveries of the  
83 AMOC on millennial timescales (Jackson 2013; Haskins et al. 2019). Yet, in some studies, the  
84 AMOC has remained in a weakened state for prolonged time periods (Stocker and Schmittner 1997;  
85 Liu et al. 2017; He et al. 2019) and taken much longer to fully recover (Stocker and Schmittner  
86 1997; Wiebe and Weaver 1999; Stouffer and Manabe 2003). These studies did not provide a

87 mechanistic explanation for the slower AMOC recovery. There remains a pressing need to identify  
88 the mechanisms that lead to divergent AMOC responses in GCMs on millennial timescales.

89 The complexities of coupled GCMs, however, make it challenging to identify individual mecha-  
90 nisms underpinning changes to the ocean's overturning circulation, which can arise from modifi-  
91 cation to the ocean's surface freshwater, heat and momentum fluxes Building on previous studies  
92 (e.g., Thorpe et al. 2001; De Boer et al. 2010; Butler et al. 2016; Haskins et al. 2019, 2020; Ragen  
93 et al. Submitted), here we use the dynamic link between overturning strength and zonal/meridional  
94 density gradients in the Atlantic basin to attribute circulation anomalies to thermal and haline  
95 anomalies. This approach assumes that the response of the AMOC to forcing is sufficiently slow  
96 that it remains in geostrophic balance and the spatial structure of the overturning is related to the  
97 density distribution in the Atlantic basin through a thermal wind relationship. An advantage of  
98 this approach is that changes in the AMOC can be attributed to specific processes that modify  
99 the density field in various regions. For instance, Haskins et al. (2019) found that in response to  
100 freshwater input from hosing experiments, the AMOC weakens initially, but then recovers after a  
101 few centuries. The weakening is attributed to changes in the density localized to the high-latitude  
102 North Atlantic, while the recovery is attributed to density changes occurring in the low latitudes of  
103 the Atlantic basin. However, assessing changes to the AMOC from meridional density gradients  
104 alone has limitations (e.g., De Boer et al. 2010), and can be subject to misinterpretation when not  
105 accounting for changes to the AMOC depth that may be associated with changes in wind stress or  
106 deep-ocean water masses.

107 Connecting the ocean's density structure to overturning strength has been more prevalent in  
108 idealized, theoretical studies that have considered the response of the AMOC to various external  
109 forcing (Nikurashin and Vallis 2012; Jansen et al. 2018; Jansen and Nadeau 2019; Nadeau and  
110 Jansen 2020). Specifically, Jansen et al. (2018) and Jansen and Nadeau (2019) showed that the  
111 thermal wind relationship in an Atlantic-like basin is able to emulate the response of the ocean's  
112 overturning in an ocean-only simulation and is useful when mechanistically interpreting simulated  
113 overturning changes to atmospheric warming. Yet, these idealized models often neglect more  
114 sophisticated processes of the climate system that may substantially influence the AMOC. For  
115 instance, most conceptual models make assumptions about ocean-atmosphere coupling by either  
116 prescribing the isopycnal overturning circulation in the North Atlantic or neglecting interactions

117 between the surface buoyancy forcing and overturning in the North Atlantic (Nikurashin and  
118 Vallis 2012; Jansen et al. 2018; Jansen and Nadeau 2019; Thomas and Fedorov 2019). Both of  
119 these approaches raise questions about the role of ocean-atmosphere interactions and the role of  
120 local North Atlantic versus remote lower latitude changes in the response of the AMOC to external  
121 forcing. While these idealized models have improved our fundamental understanding of the AMOC  
122 dynamics, the long-term behavior of the AMOC in more complex, coupled GCMs remains unclear.  
123 In particular, there is no current explanation for why GCMs produce both recovered and prolonged  
124 weakened AMOC states, despite being given the same greenhouse-gas forcing.

125 In this study, we introduce a framework to examine processes influencing the transient and  
126 equilibrium responses of the AMOC to warming in coupled GCMs. To do this, we use a collection  
127 of millennial-length climate simulations in which the atmospheric carbon-dioxide was abruptly  
128 quadrupled above pre-industrial levels and held constant for the remainder of the simulation ( $4 \times$   
129  $\text{CO}_2$ ) compiled as part of LongRunMIP (Rugenstein et al. 2019). A key asset of LongRunMIP is  
130 that it contains a diverse set of GCMs that exhibit strikingly different behaviors in response to  $4$   
131  $\times \text{CO}_2$ , specifically in the ocean (Frölicher et al. 2020; Rugenstein et al. 2019, 2020). Thus, this  
132 suite of GCM simulations provide a unique opportunity to identify processes governing the long-  
133 term evolution of the AMOC without the temporal limitations normally imposed by computational  
134 constraints. In what follows, we first show that the time-evolution of the AMOC in GCMs can  
135 be well approximated by a thermal wind expression that relates the overturning circulation in  
136 the North Atlantic to the density difference between the region of deep-water formation and the  
137 Atlantic basin. Using this expression, we then isolate thermal and haline anomalies leading to the  
138 time-dependent stratification and circulation changes, which allows us to explicitly attribute the  
139 responses to contributions from surface forcing and advective processes in different regions of the  
140 Atlantic basin. From this approach, key controls on the level and timescale of AMOC recoveries  
141 are summarized and discussed.

## 142 **2. Materials and methods**

### 143 *a. Model output and diagnostics*

144 LongRunMIP is a model intercomparison project that aims to better understand centennial and  
145 millennial timescale atmosphere-ocean processes in coupled GCMs (Rugenstein et al. 2019). In

146 this study, we limit our analysis to  $4 \times \text{CO}_2$  simulations from seven different GCMs that span  
 147 1,000 years or longer. A summary of the GCMs, including the length of each simulation, the  
 148 resolution of the atmosphere and ocean components, and the relevant reference for further details  
 149 are provided in Table (1). The GCM with the longest simulation is CESM1, which is run for  
 150 5900 years and the GCMs with the shortest simulations are IPSL-CM5A-LR and MPI-ESM1.2,  
 151 which are both run for 1000 years.

152 TABLE 1. Overview of each simulation. The resolution of the atmosphere and ocean is given in number of  
 153 grid points per latitude  $\times$  longitude  $\times$  depth. The last column contains references that describe the GCMs and  
 154 simulations in more detail. Note, some simulations are the extensions of simulations discussed in the references.

Model Name	Length	Atmosphere Resolution	Ocean Resolution	Reference
CESM1	5900 years	$96 \times 144$	$384 \times 320 \times 60$	Rugenstein et al. (2016a)
CNRM-CM6.1	1850 years	$128 \times 256$	$180 \times 360 \times 75$	Saint-Martin et al. (2019)
GISS-E2-R	5000 years	$90 \times 144$	$180 \times 288 \times 32$	Schmidt et al. (2014)
HadGEM2-ES	1295 years	$145 \times 192$	$216 \times 360 \times 40$	Andrews et al. (2015)
IPSL-CM5A-LR	1000 years	$96 \times 96$	$149 \times 182 \times 31$	Dufresne et al. (2013)
MPI-ESM1.1	4458 years	$96 \times 192$	$220 \times 256 \times 40$	Mauritsen et al. (2019)
MPI-ESM1.2	1000 years	$96 \times 192$	$220 \times 256 \times 40$	Mauritsen et al. (2019)

155 We focus on the evolution of the meridional overturning streamfunction in the Atlantic basin,  
 156 which is provided in the LongRunMIP repository (given by the name ‘moc’). The overturning is  
 157 archived as the Eulerian-mean meridional streamfunction, which is calculated as

$$\psi(t, y, z) = - \int_{-H}^z \int_{x_W}^{x_E} v(t, x, y, z) dx dz, \quad (1)$$

158 where  $x$  is longitudinal displacement,  $y$  is latitudinal displacement,  $z$  is depth with  $H$  the depth  
 159 of the ocean bottom,  $v$  is meridional velocity (including the bolus contribution), and  $x_W$  and  $x_E$   
 160 are the western and eastern boundaries of the Atlantic basin. In the analyses below, the AMOC  
 161 strength is defined as the maximum value of the meridional streamfunction in the North Atlantic,  
 162 between  $40^\circ\text{N}$  to  $60^\circ\text{N}$ .

163 To provide a more mechanistic interpretation of the AMOC responses in each GCM, we relate  
 164 the overturning circulation in the North Atlantic to the density difference between the the region  
 165 of deep-water formation  $\rho_n$  and the Atlantic basin  $\rho_b$ . Density gradients have been widely used to

166 understand controls on overturning strength in a number of theoretical studies (e.g., Stommel 1961;  
 167 Welander 1971; Gnanadesikan 1999; Wolfe and Cessi 2010; Nikurashin and Vallis 2012; Jansen  
 168 et al. 2018; Jansen and Nadeau 2019), and they also have been used to interpret the behavior of  
 169 coupled GCMs (see references in Section 1). We use potential density  $\rho$  referenced to the surface,  
 170 calculated from the model-archived conservative temperature (‘thetao’), and absolute salinity (‘so’).  
 171 To attribute variations in the density field to thermal and haline changes, we use a linear equation  
 172 of state with time-varying coefficients of thermal expansion  $\alpha$  and haline contraction  $\beta$ , which can  
 173 be expressed as,

$$\rho(t, z) = \rho_0 \left( \underbrace{1 - \bar{\alpha}(t, z) (T(t, z) - T_0)}_{\text{thermal}} + \underbrace{\bar{\beta}(t, z) (S(t, z) - S_0)}_{\text{haline}} \right), \quad (2)$$

174 where  $\rho_0$  is a reference potential density,  $T_0$  is a reference temperature and  $S_0$  is a reference salinity.  
 175 The coefficients,  $\bar{\alpha}$  and  $\bar{\beta}$ , are computed individually for each GCM and at each timestep as a  
 176 function of depth throughout the entire Atlantic basin (30°S to 60°N) using TEOS-10 (McDougall  
 177 and Barker 2011). The reference values,  $\rho_0$ ,  $T_0$ , and  $S_0$ , are area-weighted and depth-averaged  
 178 values for the Atlantic basin, calculated separately for each GCM, during years 1 – 10. The potential  
 179 density of the North Atlantic  $\rho_n$  is computed using conservative temperature and absolute salinity  
 180 averaged in zonal and meridional directions at each depth between 40°N and 60°N. The potential  
 181 density of the Atlantic basin  $\rho_b$  is computed using conservative temperature and absolute salinity  
 182 averaged in zonal and meridional directions at each depth in the Atlantic basin between 30°S and  
 183 60°N. The sensitivity of the results to these exact domains are discussed in Section 3b. A linear  
 184 equation of state is chosen for its ease in partitioning density changes into thermal and haline  
 185 contributions. Below we show that density changes based on the linear equation of state recover  
 186 the diagnosed overturning changes from each GCM.

187 We also use the net surface heat flux (including those from ice formation and melting processes),  
 188 freshwater flux (including precipitation, evaporation, runoff, melt from sea ice, frazil ice processes,  
 189 salt flux from ocean-ice processes), sea surface salinity (‘sos’) and sea surface temperature (‘tos’) to  
 190 calculate surface-forced water mass transformation (Walín 1982) for those GCMs that provided the  
 191 necessary output (only CESM1 and CNRM-CM6.1). Finally, we use sea ice concentration (‘sic’)  
 192 to compute Arctic sea ice area. In the analyses below, all variables are annual-means computed

193 from monthly output and smoothed with a Savitzky-Golay filter with a window size of 11 years.  
194 All anomalies are calculated relative to an average of years 1 – 10.

195 *b. Interpretation of the simulated Atlantic overturning circulation*

196 Following Nikurashin and Vallis (2012) and Jansen et al. (2018), we relate the meridional  
197 overturning circulation  $\psi$  in the northern part of the Atlantic basin to the density difference  
198 between the deep-water formation region and the larger Atlantic basin as

$$\frac{\partial^2 \psi}{\partial z^2} = -\frac{g}{\rho_0 f} (\rho_n - \rho_b), \quad (3)$$

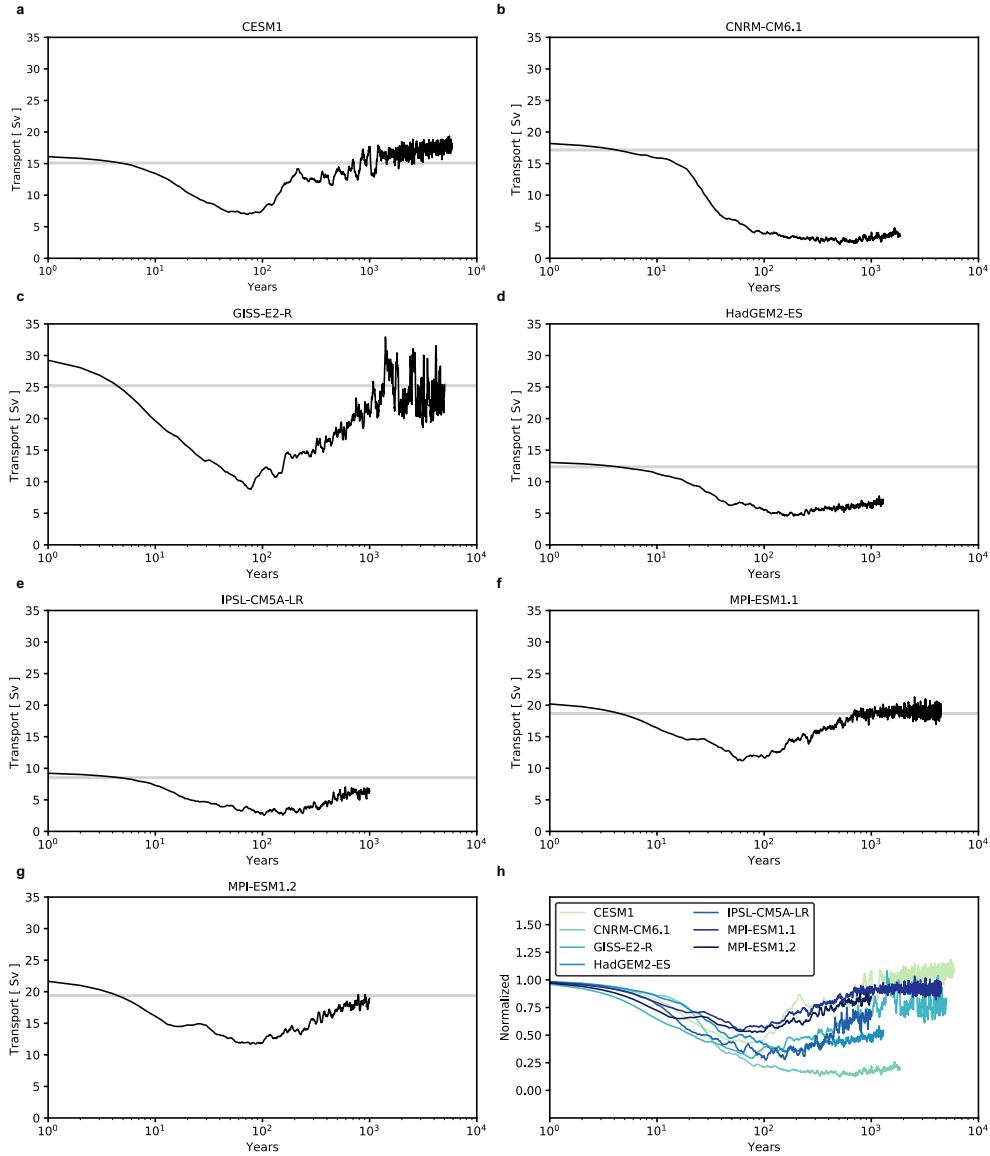
199 where  $g$  is the gravitational acceleration ( $9.81 \text{ m s}^{-2}$ ),  $\rho_0$  is a reference density ( $1025 \text{ kg m}^{-3}$ ),  
200  $f$  is the Coriolis parameter ( $f = 1.2 \times 10^{-4} \text{ s}^{-1}$ ), and  $\psi$  is the meridional overturning transport  
201 streamfunction ( $\text{m}^3 \text{ s}^{-1}$ ). Throughout this study, Eq. (3), which we will subsequently refer to as  
202 the TWR (thermal wind relationship), is solved with the boundary conditions,  $\psi = 0$  at the surface  
203 and bottom of the ocean.

204 The TWR, as expressed in Eq. (3), explicitly relates the zonal transport to a meridional density  
205 difference. A key assumption of this expression is that in geostrophic balance, the meridional  
206 overturning transport and the zonal overturning transport are equivalent through continuity. In  
207 other words, a vertically sheared zonal flow at midlatitudes is connected to the depth-latitude  
208 overturning in the North Atlantic, a relationship that was originally proposed by Nikurashin and  
209 Vallis (2012). To our knowledge, this expression has not been tested in realistic GCM simulations  
210 and it is not obvious that this expression will emulate the GCMs discussed in this study. The link  
211 between AMOC and the TWR is tied to the fact that isopycnals associated with the maximum of  
212 the upper-overturning cell tilt up and outcrop toward the north and west. Differences between the  
213 streamfunction diagnosed from GCMs and the streamfunction diagnosed from TWR result from  
214 the choice of the averaging region for  $\rho_n$ , which is discussed in more detail in Section 3b.

215 **3. Response of the Atlantic overturning circulation to abrupt forcing**

216 In response to  $4 \times \text{CO}_2$ , the LongRunMIP GCMs show a relatively uniform response on centennial  
217 and shorter timescales, encompassing the weakening of the AMOC, but exhibit more diverse  
218 responses on long timescales during which the AMOC recovers at various rates. In the following

219 subsections, we summarize key features of the long-term AMOC evolution, and then demonstrate  
 220 that the TWR presented in Eq. (3) provides a suitable framework for understanding these AMOC  
 221 changes.

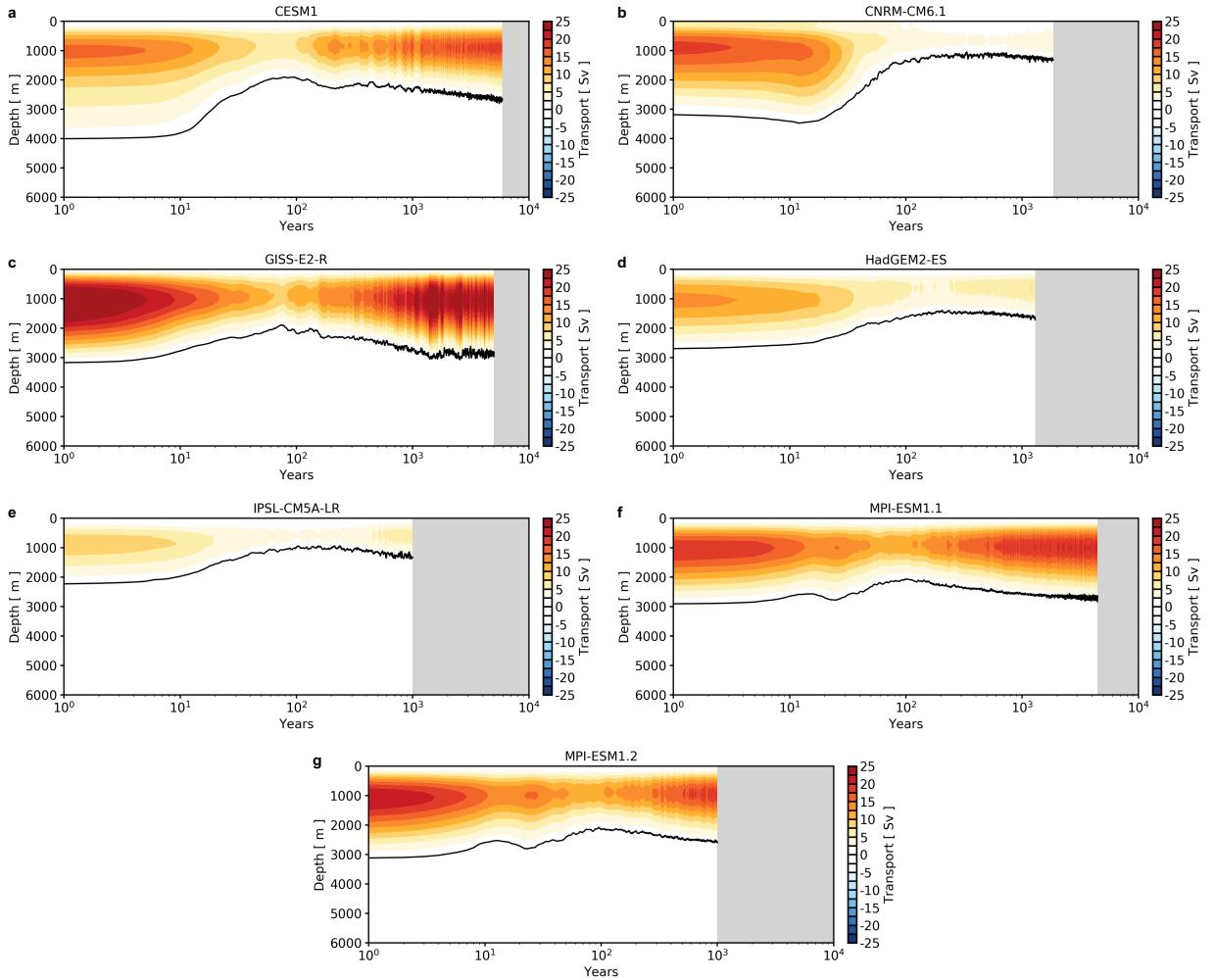


222 FIG. 1. Maximum value of the meridional streamfunction between  $40^{\circ}\text{N}$  and  $60^{\circ}\text{N}$  in the Atlantic basin from  $4 \times \text{CO}_2$  simulations in (a) CESM1, (b) CNRM-CM6.1, (c) GISS-E2-R, (d) HadGEM2-ES, (e) IPSL-CM5A-LR,  
 223 (f) MPI-ESM1.1, and (g) MPI-ESM1.2. (h) AMOC strength normalized by the initial strength (years 1 – 10) for  
 224 each GCM. A Savitzky-Golay filter with a window size of 11 years was applied to each timeseries. In panels  
 225 (a-g) the grey line denotes the initial strength (years 1 – 10). The streamfunction used here is defined in Eq. (1).  
 226

227 *a. Model response*

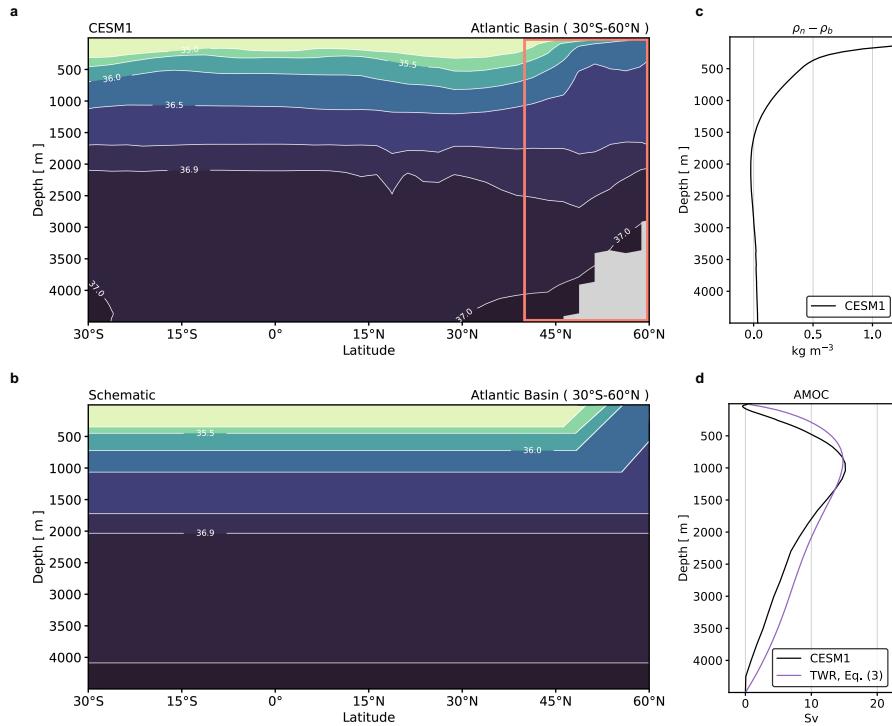
228 The response of the AMOC to  $4 \times \text{CO}_2$  in each GCM is first examined using the maximum  
229 value of the meridional streamfunction in the North Atlantic, defined between  $40^\circ\text{N}$  to  $60^\circ\text{N}$ . All  
230 of the GCMs exhibit a weakening of the AMOC over centennial timescales (Fig. 1). GCMs with  
231 a stronger initial AMOC strength typically exhibit a larger AMOC weakening, defined as a change  
232 in volume transport, with changes spanning 12 and 20 Sv ( $1 \text{ Sv} \equiv 1 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ ; Fig. 1c).  
233 GCMs with a weaker initial AMOC strength still experience substantial changes, but only weaken  
234 between 5 and 7 Sv (Fig. 1e). After the initial weakening of the AMOC, nearly all the GCMs  
235 (except for CNRM-CM6.1) exhibit a sustained strengthening of the AMOC, although the rate of  
236 this strengthening varies across the GCMs (Fig. 1h). In these GCMs, at least a partial recovery  
237 of the AMOC occurs, and in some cases, the AMOC strength surpasses its initial magnitude,  
238 though this occurs over different timescales (Fig. 1h). In CESM1, for instance, the AMOC reaches  
239 75% of its original strength around 300 years after the abrupt forcing, while in HadGEM2-ES, the  
240 AMOC only recovers to 25% of its initial strength after approximately 1300 years (Fig. 1h). In  
241 CNRM-CM6.1, the AMOC weakens to approximately 2 Sv and remains weakened until around  
242 1000 years, where it begins a slow, gradual recovery (Fig. 1b). Notably, approximately 3000 years  
243 after the abrupt forcing, the AMOC strength in CESM1, GISS-E2-R, and MPI-ESM1.1 is slightly  
244 stronger than their initial states by approximately 1 – 3 Sv (Fig. 1h). At the end of each simulation,  
245 some GCMs have reached a new steady state, whereas others are still evolving. None of the GCMs  
246 exhibit a longer-term AMOC weakening.

252 In addition to change in AMOC strength, the vertical structure of the AMOC also evolves in  
253 response to  $4 \times \text{CO}_2$ . We examine the vertical structure of the AMOC in the North Atlantic as a  
254 function of time by calculating the maximum value of the meridional streamfunction at each depth  
255 between  $40^\circ\text{N}$  to  $60^\circ\text{N}$ . All GCMs exhibit a shoaling of the AMOC depth, defined as the lower  
256 boundary of the overturning cell where  $\psi = 0$  (see black line in Fig. 2), during the period of AMOC  
257 weakening on centennial timescales (Fig. 2). In contrast, the depth of the maximum value of the  
258 AMOC is relatively unchanged in response to increased greenhouse-gas concentrations. However,  
259 the magnitude of shoaling in the AMOC depth varies substantially across the GCMs (black line in  
260 Fig. 2). In CNRM-CM6.1 the AMOC shoals by approximately 2000 m (Fig. 2b), while in MPI-  
261 ESM1.1 the AMOC shoals by approximately 500 m (Fig. 2f). In all GCMs except CNRM-CM6.1,



247 FIG. 2. Hovmöller plot of the maximum value of the meridional streamfunction between  $40^{\circ}\text{N}$  and  $60^{\circ}\text{N}$   
 248 in the Atlantic basin as a function of depth from  $4 \times \text{CO}_2$  simulations in (a) CESM1, (b) CNRM-CM6.1, (c)  
 249 GISS-E2-R, (d) HadGEM2-ES, (e) IPSL-CM5A-LR, (f) MPI-ESM1.1, and (g) MPI-ESM1.2. A Savitzky-Golay  
 250 filter with a window size of 11 years was applied to each timeseries. The black line denotes where  $\psi = 0$ . The  
 251 streamfunction used here is defined in Eq. (1).

262 the AMOC recovers in strength and deepens over the next millennia. In CESM1, GISS-E2-R,  
 263 MPI-ESM1.1, and MPI-ESM1.2, the AMOC is eventually stronger than its initial state between  
 264 500 – 1000 m, but remains weakened below 1000 m (Fig. 2). After approximately 3000 years, the  
 265 AMOC deepens back to its initial depth in GISS-E2-R, MPI-ESM1.1, and MPI-ESM1.2, but not in  
 266 CESM1, HadGEM2-ES, or IPSL-CM5A-LR. In CNRM-CM6.1, the AMOC stays weakened and  
 267 remains shoaled by approximately 2000 m, even 1000 years after the  $4 \times \text{CO}_2$  (Fig. 2b).



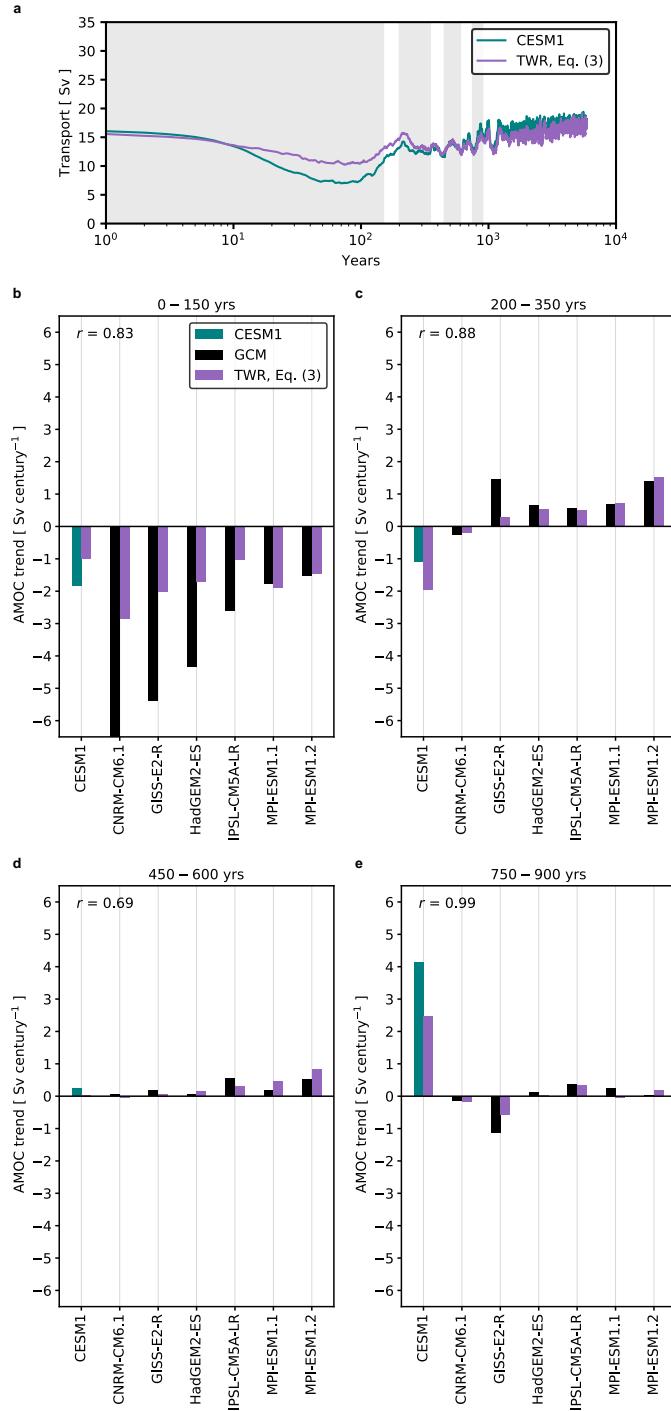
269 FIG. 3. Schematic showing how the thermal wind expression, Eq. (3), is applied to the LongRunMIP  
 270 simulations. (a) Meridional section of isopycnals averaged zonally across the Atlantic basin during years 1 – 10  
 271 for CESM1. The red box denotes the averaging region for  $\rho_n$ , which is the same for all GCMs. (b) Schematic of  
 272 the isopycnals in the Atlantic basin which remain relatively flat throughout most of the basin and outcrop in the  
 273 North Atlantic. (c) Density difference ( $\rho_n - \rho_b$ ) between the density profile in the North Atlantic ( $40^\circ\text{N} - 60^\circ\text{N}$ )  
 274  $\rho_n$  and in the Atlantic basin ( $30^\circ\text{S} - 60^\circ\text{N}$ )  $\rho_b$  in CESM1 during years 1 – 10. (d) The meridional streamfunction  
 275 for (black) CESM1 and (purple) using the thermal wind relationship, Eq. (3), with the meridional density  
 276 difference in (c).

277 To better understand how the AMOC responses described in the previous section arise, we first  
 278 examine the ability of the TWR in Eq. (3) to emulate the AMOC responses in each GCM across a  
 279 range of timescales. We compute spatially-averaged potential density profiles in the North Atlantic  
 280 and the entire Atlantic basin for each GCM (see Section 2), and then apply the TWR in Eq. (3)  
 281 to recreate the time-dependent overturning circulation in the North Atlantic. The results presented  
 282 below are insensitive to the exact averaging choice of Atlantic basin domain, so long as estimates  
 283 of  $\rho_b$  include the Atlantic basin northward of  $30^\circ\text{S}$  and estimates of  $\rho_n$  incorporates deep water

284 formation regions, which also correspond to the location where the mid-depth isopycnals outcrop  
285 (Fig. 3a-b).

286 The meridional density difference,  $\rho_n - \rho_b$ , is largest near the surface and decreases to approxi-  
287 mately zero around 1500 – 2000 m in most GCMs (Fig. 3c). The vertical structure and magnitude  
288 of the streamfunction computed from the TWR largely reproduces the streamfunction diagnosed  
289 from CESM1 (Fig. 3d). Note, the streamfunction below approximately 2000 m is overestimated  
290 by the TWR in CESM1 and in most GCMs (not shown). More importantly, the maximum value  
291 of the overturning derived from the TWR approximates the time-evolution of the AMOC for all  
292 GCMs. We provide one example of this agreement for CESM1 (Fig. 4a); the other GCMs are  
293 similar. The TWR captures the initial weakening, partial recovery, and gradual re-strengthening  
294 of the AMOC, although the TWR tends to underestimate the weakening of the AMOC in most  
295 GCMs.

302 To more quantitatively assess the skill of the TWR in reproducing the temporal evolution of  
303 the AMOC in each GCM, we compute a linear trend in the maximum streamfunction across four  
304 different 150-year time periods (years 0 – 150, years 200 – 350, years 450 – 600, and years 750 – 900)  
305 that span common years of each simulation. During the initial weakening (years 0 – 150), the TWR  
306 captures the simulated robust negative AMOC trends, but underestimates the magnitude, which  
307 spans  $-7$  to  $-1$  Sv century<sup>-1</sup> (Fig. 4a). This underestimation may come about because of the  
308 non-trivial relationship between surface-forced water mass transformation, interior stratification,  
309 and zonal and meridional overturning on centennial timescales (see Section 5a). Nevertheless, the  
310 TWR accounts for approximately 50% of the intermodel variance. A comparison of the linear  
311 AMOC trend diagnosed from each GCM and the TWR across three time periods after the initial  
312 weakening (years 200 – 350, years 450 – 600, and years 750 – 900) shows that the TWR emulates  
313 the AMOC evolution during other time periods (Figs 4b-d). For each period, the TWR accounts  
314 for 50 – 90% of the intermodel variance and also captures the broad tendency of a slow AMOC  
315 recovery in each GCM. The TWR also captures the shoaling of the AMOC during its weakening  
316 phase and the gradual deepening of the AMOC (not shown). The most significant deviation  
317 between overturning rates derived from Eq. (1) and the TWR occurs during the initial weakening  
318 phase of the AMOC. In the following, we use the TWR to examine drivers of the AMOC responses



296 FIG. 4. Comparison of the meridional streamfunction calculated from GCMs and the thermal wind relationship  
 297 in Eq. (3). (a) Maximum value of the meridional streamfunction between 40°N and 60°N from (teal) Eq. (1)  
 298 and (purple) Eq. (3) for CESM1. Linear trend of the maximum value of the meridional streamfunction between  
 299 40°N and 60°N in the Atlantic basin from (b) 0 – 150 years, (c) 200 – 350 years, (d) 450 – 600 years, (e) 750 – 900  
 300 years for (black) GCMs and (purple) Eq. (3). The grey shaded regions in (a) denote the periods for the linear  
 301 trends.

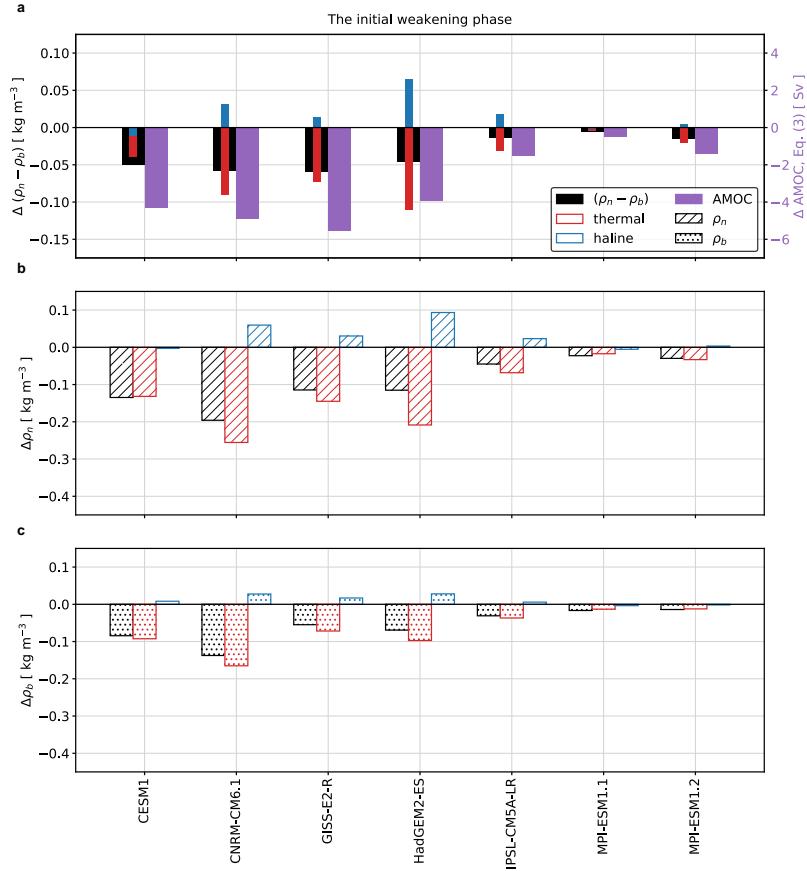
319 in each GCM, which can be attributed to density changes in the high-latitude North Atlantic ( $\rho_n$ )  
 320 or throughout the Atlantic basin ( $\rho_b$ ).

#### 321 4. Thermal and haline controls on changes to the Atlantic overturning circulation

322 TABLE 2. Timescale of the initial AMOC weakening in each GCM. The timescale is estimated as the first local  
 323 minimum of the timeseries in Figure 1.

Model Name	Timescale
CESM1	48 years
CNRM-CM6.1	45 years
GISS-E2-R	29 years
HadGEM2-ES	40 years
IPSL-CM5A-LR	24 years
MPI-ESM1.1	21 years
MPI-ESM1.2	22 years

324 We next focus on changes to  $\rho_n$ ,  $\rho_b$ , and  $\rho_n - \rho_b$  since the Atlantic density distribution is coupled  
 325 to the strength and structure of the AMOC. In the analyses below, the symbol  $\Delta$  indicates a  
 326 change with respect to the initial state. Density changes are partitioned into thermal and haline  
 327 contributions for three time periods that highlight different states of the AMOC evolution: the  
 328 initial weakening phase, the partial recovery phase, and the re-strengthening phase. The initial  
 329 weakening phase is identified as a 30-year average centered on the first local minimum of the  
 330 AMOC magnitude for each model (see Figure 1). The initial weakening phase spans slightly  
 331 different years in the various GCMs, but it generally occurs within 50 years of the abrupt forcing  
 332 (see Table 2). The partial recovery phase is chosen to be years 470 – 500, a period when the  
 333 AMOC is increasing in most GCMs. Finally, the re-strengthening phase is chosen to cover years  
 334 970 – 1000, a period of common years across all GCMs that occurs long after the abrupt forcing.  
 335 Because some GCMs provide longer simulations than others, we also consider the new equilibrium  
 336 state after multiple millennia (years 3500 – 3530) in the CESM1, GISS-E2-R, and MPI-ESM1.1  
 337 simulations.



338 FIG. 5. (a) Bar plot showing (black) the average density change between 250 – 2000 m for  $(\rho_n - \rho_b)$  and  
 339 (purple) the maximal change in the meridional streamfunction using Eq. (3) for each GCM during the initial  
 340 weakening phase (see Table 1) relative to years 1 – 10. The red and blue bars represent the thermal and haline  
 341 contributions to  $\Delta(\rho_n - \rho_b)$ . (b) Bar plot showing the thermal and haline contributions to  $\Delta\rho_n$  during the initial  
 342 weakening phase relative to years 1 – 10. (c) Bar plot showing the thermal and haline contributions to  $\Delta\rho_b$  during  
 343 the initial weakening phase relative to years 1 – 10.

### 344 a. The initial weakening phase

345 During the initial weakening phase of the AMOC,  $\Delta(\rho_n - \rho_b) < 0$  for all of the GCMs, indicating  
 346 a weakening of the meridional density difference (Fig. 5a). The density difference decreases by  
 347  $0.01 - 0.06 \text{ kg m}^{-3}$  between 250 and 2000 m (Fig. 5a). The decrease in the basin-wide density  
 348 contrast for each GCM is almost entirely thermally driven (compare red bar with blue bar), with  
 349 the haline contribution actually increasing the basin-wide density gradient in most GCMs (Fig.

350 5a). The weakening of  $\rho_n - \rho_b$  across all GCMs results in a universal weakening and shoaling of  
351 the AMOC of approximately 1 – 6 Sv. GCMs with larger negative values of  $\Delta(\rho_n - \rho_b)$  also show  
352 a greater weakening of the AMOC streamfunction (Fig. 5a).

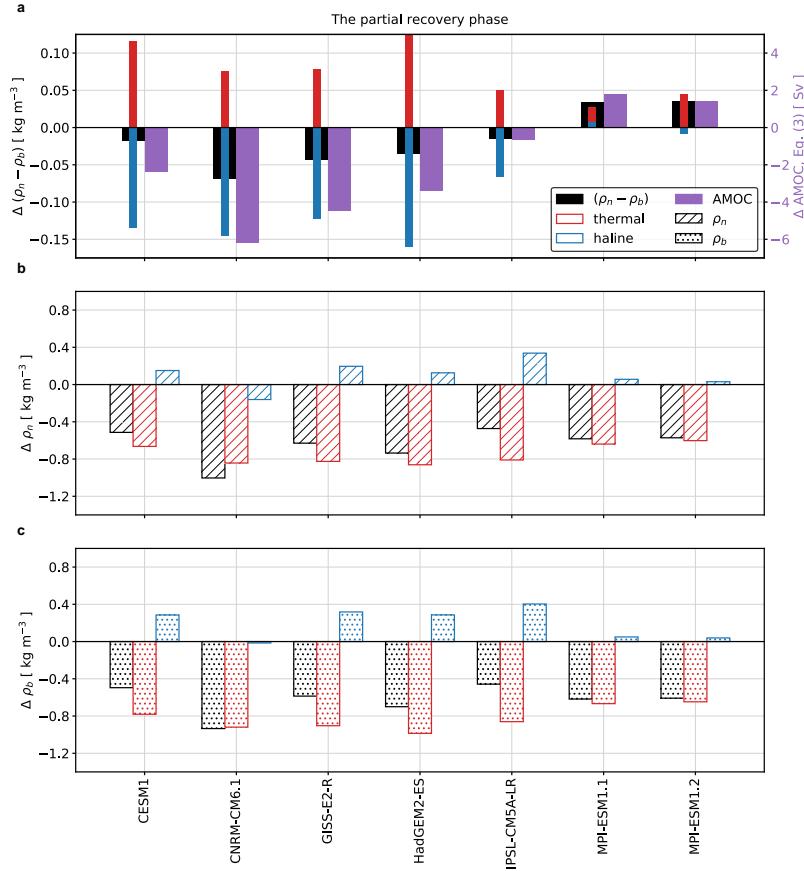
353 Both  $\Delta\rho_n$  and  $\Delta\rho_b$  are negative during the initial weakening phase, indicating large-scale lighten-  
354 ing of density throughout the Atlantic basin (Fig. 5b-c). In both regions this change is dominated by  
355 warming (the thermal contribution) with magnitudes that are roughly twice the size of  $\Delta(\rho_n - \rho_b)$ .  
356 However, changes to  $\rho_n$  exceed the changes to  $\rho_b$  due to warming during this initial weakening  
357 phase (i.e.,  $|\Delta\rho_n| > |\Delta\rho_b|$ ), which gives rise to the reduced meridional density difference. Salinity  
358 changes are weaker than and oppose the thermal changes. GCMs with larger AMOC reductions  
359 (CNRM-CM6.1) tend to warm more in the North Atlantic, while GCMs with smaller AMOC  
360 reductions (IPSL-CM5A-LR and MPI-ESM1.1) tend to warm less in the North Atlantic.

### 367 *b. The partial recovery phase*

368 After the the initial weakening, GCMs exhibit a much broader range of intermodel variability.  
369 Five hundred years after the abrupt forcing, the AMOC in CESM1 recovers to approximately 90%  
370 of its initial strength, while the AMOC in CNRM-CM6.1 remains weakened to 15% of its initial  
371 strength (Fig. 1h).

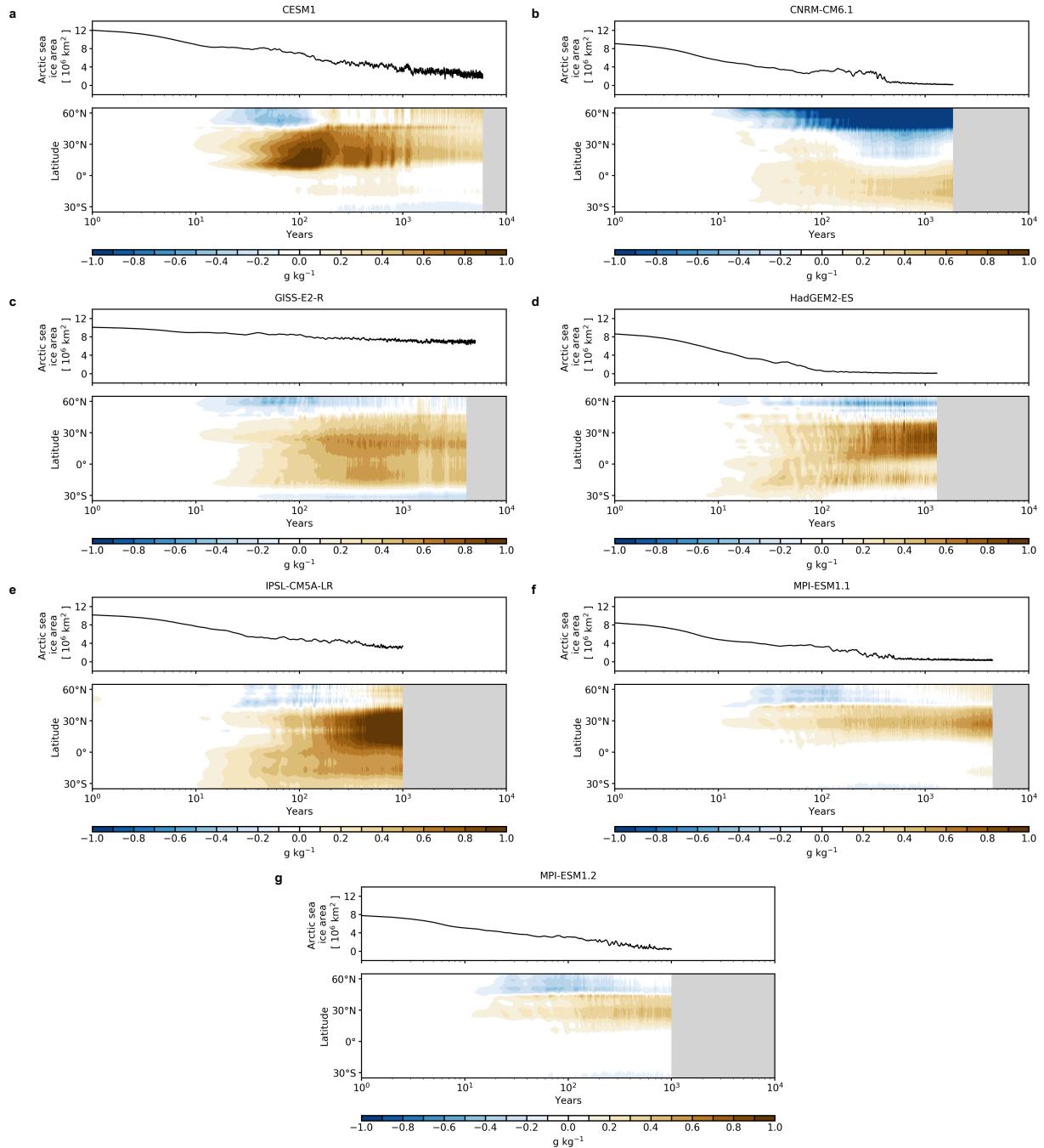
372 During years 470 – 500, a majority of the GCMs still show that  $\Delta(\rho_n - \rho_b) < 0$ , but the values  
373 are less negative when compared to the initial weakening phase, indicating a slight increase in  
374 the meridional density difference (Fig. 6a). In MPI-ESM1.1 and MPI-ESM1.2,  $\Delta(\rho_n - \rho_b) > 0$ ,  
375 indicating a stronger meridional density difference and a stronger AMOC when compared to the  
376 initial state (Fig. 6a). In most GCMs, the AMOC partially recovers by 1 – 3 Sv. In CNRM-CM6.1,  
377 however,  $\rho_n - \rho_b$  further decreases and the anomaly becomes more negative when compared to the  
378 initial weakening phase.

384 During the partial recovery phase, density throughout the Atlantic basin continues to get lighter.  
385  $\Delta\rho_n$  and  $\Delta\rho_b$  are more negative than in the weakening phase, but the haline contribution to density  
386 changes is now of comparable size and modulates the response across the different GCMs (Fig.  
387 6). The partial AMOC recovery is associated with a simultaneous warming of the Atlantic basin  
388 relative to the high-latitudes (compare red-dashed to red-dotted), and a positive salinity anomaly  
389 in the North Atlantic (Fig. 6b-c). A positive salinity anomaly develops in the subtropical Atlantic



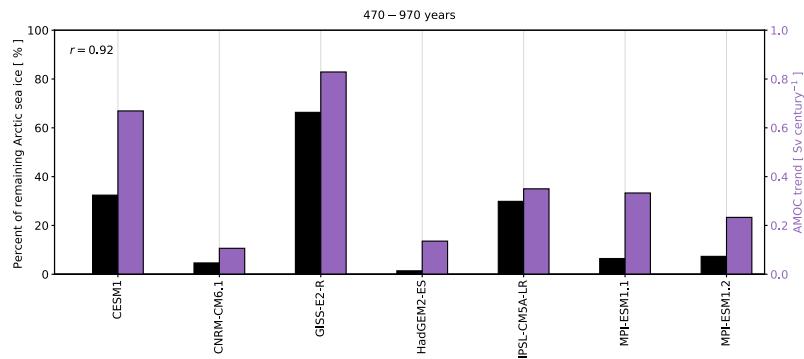
361 FIG. 6. (a) Bar plot showing (black) the average density change between 250 – 2000 m for  $(\rho_n - \rho_b)$  and  
 362 (purple) the maximal change in the meridional streamfunction using Eq. (3) for each GCM during the partial  
 363 recovery phase (years 470 – 500) relative to years 1 – 10. The red and blue bars represent the thermal and haline  
 364 contributions to  $\Delta(\rho_n - \rho_b)$ . (b) Bar plot showing the thermal and haline contributions to  $\Delta\rho_n$  during the partial  
 365 recovery phase relative to years 1 – 10. (c) Bar plot showing the thermal and haline contributions to  $\Delta\rho_b$  during  
 366 the partial recovery phase relative to years 1 – 10.

390 basin and is advected into the North Atlantic between 250 – 1000 m after the AMOC weakening.  
 391 This northward salinity transport causes  $\rho_n$  to become denser, while the warming at lower latitudes  
 392 cause  $\rho_b$  to become lighter. Together, these processes increase the magnitude of  $\rho_n - \rho_b$  and  
 393 cause  $\Delta(\rho_n - \rho_b)$  to become less negative (Fig. 6b-c). In CNRM-CM6.1, which undergoes the  
 394 slowest AMOC recovery, warming occurs throughout the Atlantic basin. However, instead of a  
 395 positive salinity anomaly in the high latitudes, the haline contribution indicates a freshening, or



379 FIG. 7. Hovmöller plot of the anomalous zonally averaged salinity in the Atlantic basin from 250–2000 m  
 380 for (a) CESM1, (b) CNRM-CM6.1, (c) GISS-E2-R, (d) HadGEM2-ES, (e) IPSL-CM5A-LR, (f) MPI-ESM1.1,  
 381 and (g) MPI-ESM1.2. A Savitzky-Golay filter with a window size of 11 years was applied to all timeseries. All  
 382 anomalies are calculated relative to an average of years 1–10. The upper plot shows annual Arctic sea-ice area  
 383 for each GCM.

396  $\Delta\rho_n < 0$  (blue bar, Fig. 6b). The high-latitude freshening further decreases the meridional density  
 397 difference and suppresses the AMOC recovery in this GCM.



398 FIG. 8. Bar plot showing (black) the percent of remaining Arctic sea ice relative to the initial state (years  
 399 1 – 10) and (purple) the corresponding AMOC trend for years 470 – 970 in each GCM. The Pearson correlation  
 400 coefficient between the two is shown in the upper left hand corner.

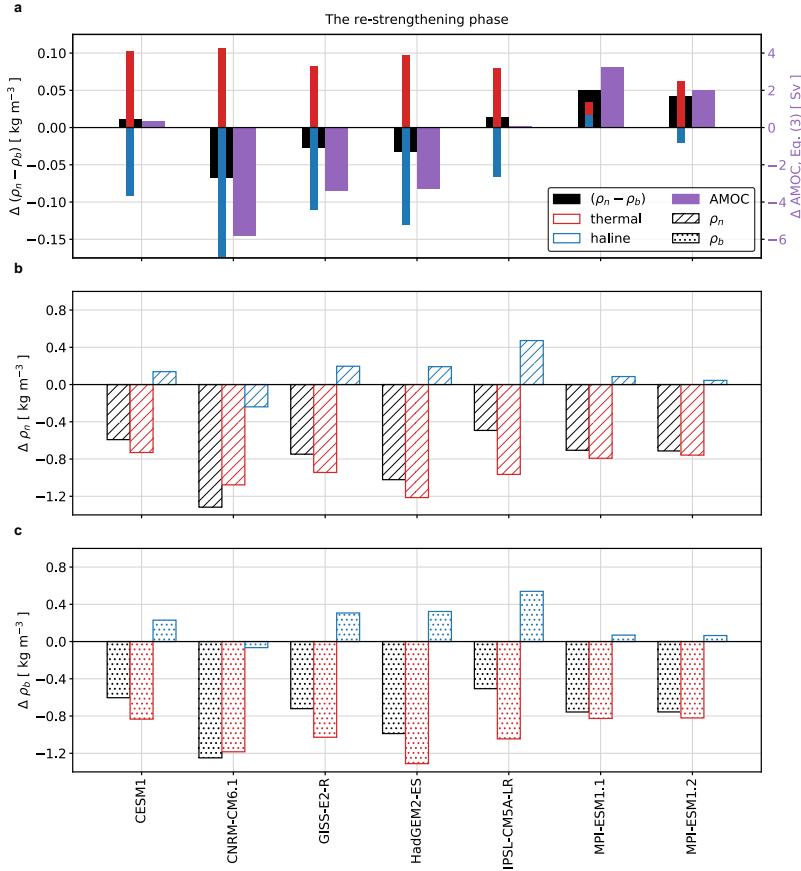
401 The primary difference between CNRM-CM6.1, which undergoes the slowest AMOC recovery,  
 402 and the other GCMs is the high-latitude freshwater anomaly. During the first 100 to 150 years  
 403 there is a freshening of the high latitude Atlantic basin in CNRM-CM6.1 due to reduced salinity  
 404 transport by the AMOC weakening, increased sea ice melt and an intensified hydrological cycle  
 405 (Fig. 7). There is also a convergence of salinity in the subtropical Atlantic basin (30°S–30°N) that  
 406 is nearly universal across GCMs, though it is weaker in CNRM-CM6.1 (Fig. 7). After 100 to 150  
 407 years the subtropical salinity anomaly is transported northward from the subtropical gyre region  
 408 to the subpolar gyre region and mixes with the fresh high-latitude waters. This salinity anomaly  
 409 increases the density of the high-latitude subsurface waters and helps to erode the high-latitude  
 410 vertical stratification, which enables a reinvigoration of convection. In CNRM-CM6.1, which  
 411 shows little-to-no AMOC recovery, the local freshening signal overwhelms the salinity anomaly  
 412 coming from the subtropics (Fig. 7b). Similarly, HadGEM2-ES, which also shows a slower  
 413 AMOC recovery when compared to other GCMs, experiences larger high latitude freshening when  
 414 compared to other GCMs (Fig. 7d). Both of these GCMs experience greater Arctic sea ice loss  
 415 when compared to other GCMs (Fig. 7). In fact, the intermodel spread of trends in AMOC strength  
 416 between years 470 and 970 is strongly correlated with the percent of remaining Arctic sea-ice area  
 417 during this period (Fig. 8). GCMs with less Arctic sea ice loss tend to show more positive trends

418 in AMOC strength or stronger AMOC recoveries, whereas GCMs with more Arctic sea ice loss  
419 tend to show less positive trends in AMOC strength or weaker AMOC recoveries (Fig. 8). At the  
420 end of the simulation in CNRM-CM6.1, the AMOC has a slight positive trend and the freshening  
421 signal begins to erode away (Fig. 7b), suggesting that during a longer simulation a recovery  
422 may ultimately occur after a sufficiently large build up of salinity at the low latitudes. Another  
423 possible explanation for the large freshening of the subpolar gyre region is the AMOC weakening  
424 itself, which can reduce salinity transport to the North Atlantic. However, there is no significant  
425 relationship between the magnitude of AMOC weakening and freshening of the subpolar gyre  
426 region across the various GCMs, indicating that other sources of freshening are at play. Note,  
427 CNRM-CM6.1 does warm more than other GCMs suggesting that changes to precipitation minus  
428 evaporation ( $P - E$ ) may be larger in the high-latitudes following the “wet gets wetter, dry gets  
429 drier” paradigm (Held and Soden 2006).

### 430 *c. The re-strengthening phase and equilibrium response*

431 During the re-strengthening phase of the AMOC (years 970 – 1000),  $\Delta(\rho_n - \rho_b) > 0$  for more  
432 than half of the GCMs (Fig. 9a). That is, the meridional density difference becomes stronger  
433 when compared to the initial state. However, for CNRM-CM6.1 and HadGEM2-ES, which both  
434 experience a prolonged AMOC weakening and weak recovery,  $\Delta(\rho_n - \rho_b)$  remains negative.  
435 Note,  $\Delta(\rho_n - \rho_b)$  also remains negative for GISS-E2-R, but it becomes slightly less negative when  
436 compared to the partial recovery phase. In the GCMs that show positive values of  $\Delta(\rho_n - \rho_b)$ ,  
437 the implied overturning from the TWR is approximately 2 – 6 Sv stronger when compared to the  
438 initial strength (Fig. 9a). For MPI-ESM1.1 and MPI-ESM1.2, the overturning is also deeper than  
439 the initial state (not shown).

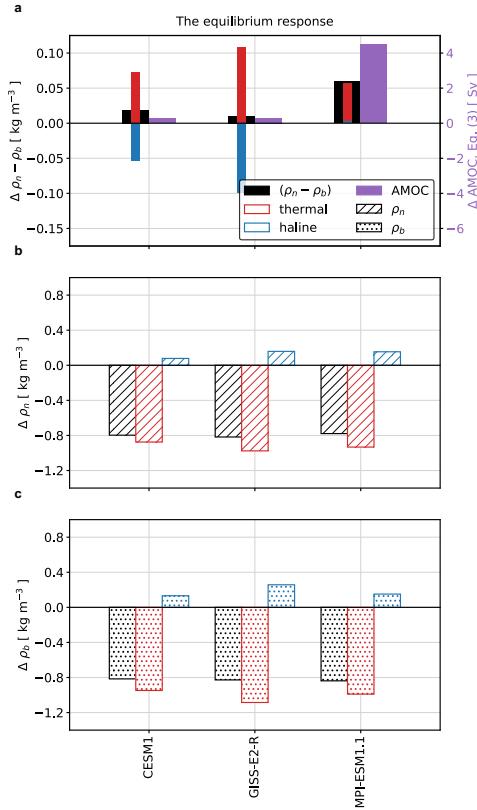
446 A decomposition of the thermal and haline contributions to  $\Delta\rho_n$  and  $\Delta\rho_b$  shows that the positive  
447 values of  $\Delta(\rho_n - \rho_b)$  are driven mainly by warming throughout the Atlantic basin (Fig. 9a-c). In  
448 contrast to the weakening phase, here  $\Delta\rho_b$  becomes less dense due to more warming as compared  
449 to  $\Delta\rho_n$  (Fig. 9b-c). The increase in the basin-wide density contrast supports an overturning that  
450 is stronger than the initial state. For most GCMs salinity acts to reduce the basin-wide density  
451 contrast. However, GCMs with a slow AMOC recovery (CNRM-CM6.1 and HadGEM2-ES)



440 FIG. 9. (a) Bar plot showing (black) the average density change between 250 – 2000 m for  $(\rho_n - \rho_b)$  and  
 441 (purple) the maximal change in the meridional streamfunction using Eq. (3) for each GCM during the re-  
 442 strengthening phase (years 970 – 1000) relative to years 1 – 10. The red and blue bars represent the thermal and  
 443 haline contributions to  $\Delta(\rho_n - \rho_b)$ . (b) Bar plot showing the thermal and haline contributions to  $\Delta\rho_n$  during the  
 444 re-strengthening phase relative to years 1 – 10. (c) Bar plot showing the thermal and haline contributions to  $\Delta\rho_b$   
 445 during the re-strengthening phase relative to years 1 – 10.

452 exhibit larger negative values of  $\Delta(\rho_n - \rho_b)$ , largely due to the haline component of  $\Delta(\rho_n - \rho_b)$   
 453 being strongly negative.

460 For all of the GCMs that simulate multiple millennia,  $\Delta(\rho_n - \rho_b) > 0$  meaning  $\rho_n - \rho_b$  exceeds  
 461 its initial state and supports a stronger overturning than the initial state (Fig. 10a). During the  
 462 equilibrium phase, the increase in  $\rho_n - \rho_b$  is primarily due to warming of the low-latitude Atlantic  
 463 basin, which contributes a  $0.05 - 0.10 \text{ kg m}^{-3}$  increase in the meridional density difference. The  
 464 salinity anomaly becomes weaker for  $\Delta\rho_b$  and  $\Delta\rho_n$ , but modulates the magnitude of the basin-



454 FIG. 10. (a) Bar plot showing (black) the average density change between 250–2000 m for  $(\rho_n - \rho_b)$  and  
 455 (purple) the maximal change in the meridional streamfunction using Eq. (3) for each GCM during the equilibrium  
 456 response (years 3500–3530) relative to years 1–10. The red and blue bars represent the thermal and haline  
 457 contributions to  $\Delta(\rho_n - \rho_b)$ . (b) Bar plot showing the thermal and haline contributions to  $\Delta \rho_n$  during the  
 458 equilibrium response relative to years 1–10. (c) Bar plot showing the thermal and haline contributions to  $\Delta \rho_b$   
 459 during the equilibrium response relative to years 1–10.

465 wide density difference (Fig. 10b-c). Previous studies have attributed a low-latitude warming in  
 466 the Atlantic basin to reduced upwelling in the tropical region (e.g., Stouffer and Manabe 2003).  
 467 Another possible explanation for the low-latitude warming is a deepening of isopycnals that is  
 468 related to AMOC weakening (Sun et al. 2020) or Southern Ocean warming.

469 **5. Discussion**

470 *a. Surface-forced water mass transformation*

471 Across all of the GCMs, the interior density distribution and its evolution in response to 4  
 472  $\times \text{CO}_2$  provides a framework for understanding transitions in the strength and structure of the  
 473 AMOC. Ultimately, however, changes in meridional overturning must be balanced by density-  
 474 layer volume modifications, diapycnal mixing, and surface buoyancy-forced transformation. An  
 475 alternative approach to understanding controls on the strength of the AMOC acknowledges that the  
 476 overturning deliver waters to regions of water mass transformation (Newsom and Thompson 2018)  
 477 and its structure in density space can be derived from the surface-forced water mass transformation  
 478 (e.g., Walin 1982; Speer and Tziperman 1992; Groeskamp et al. 2019).

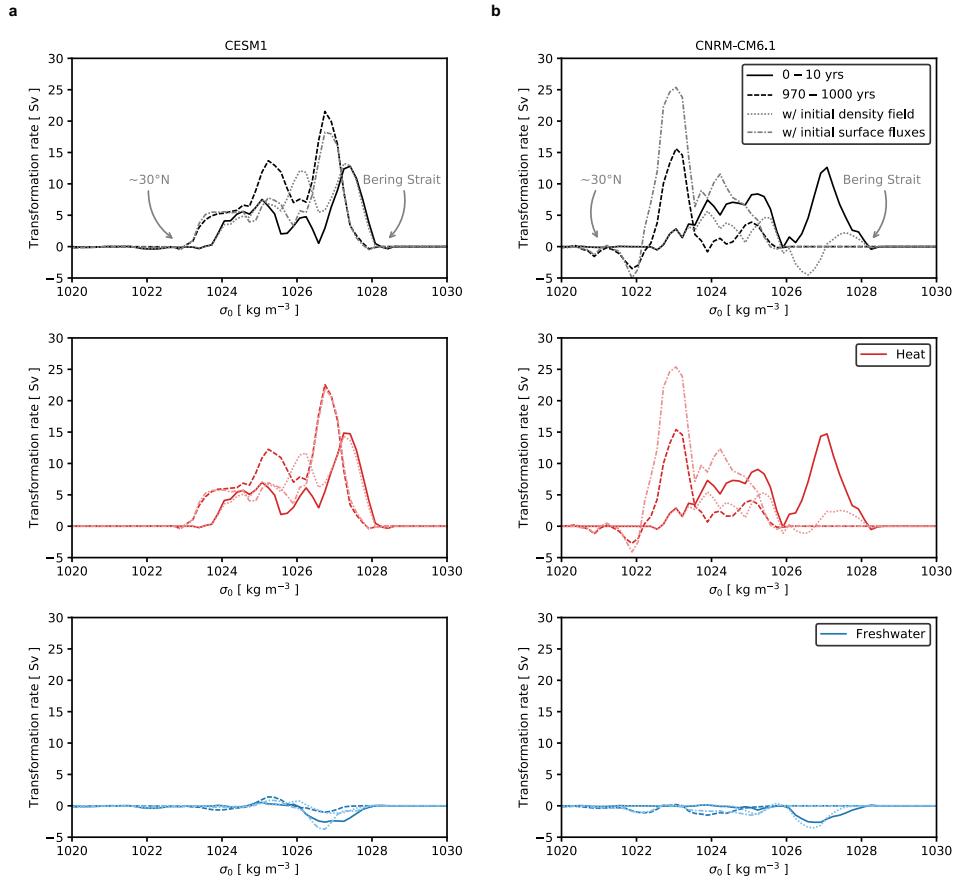
488 Surface-forced water mass transformation specifically quantifies the circulation across a given  
 489 density class sustained by surface buoyancy fluxes from heat and freshwater forcing components.  
 490 To illustrate the connection between surface and interior circulation changes, we calculate the  
 491 surface-forced water mass transformation for two GCMs (CESM1 and CNRM-CM6.1) that provide  
 492 the output needed to diagnose these processes. While this analysis cannot be performed on all  
 493 GCMs due to a most GCMs missing various surface fluxes, CESM1 and CNRM-CM6.1 fortunately  
 494 represent the end members of the strong and weak AMOC recoveries. Following Walin (1982),  
 495 the circulation across a given density class can be quantified as

$$F(y, \sigma) = \frac{\partial}{\partial \sigma} \int_{A[\sigma' > \sigma]} D(x, y, t) \mathcal{H}(\sigma'(x) - \sigma_{\min}(y)) dA, \quad (4)$$

496 where

$$D(x, y, t) = -\frac{\alpha(x, y, t)}{c_p} Q_H(x, y, t) - \beta(x, y, t) S(x, y, t) Q_F(x, y, t) \quad (5)$$

497 is the surface density flux,  $\alpha$  and  $\beta$  are the thermal expansion and haline contraction coefficients,  
 498  $Q_H$  and  $Q_F$  are the surface heat and freshwater fluxes ( $\text{kg m}^{-2} \text{s}^{-1}$ ), respectively,  $S$  is the surface  
 499 absolute salinity, and  $c_p$  is the specific heat capacity of seawater ( $4186 \text{ J kg}^{-1} \text{ K}^{-1}$ ). Here, the  
 500 surface heat and freshwater fluxes are defined to be positive into the ocean.  $F$  has units of  $\text{m}^3$   
 501  $\text{s}^{-1}$  and represents the formation of water masses as a function of density. In Eq. (4),  $\sigma_{\min}(y)$  is  
 502 the minimum density at latitude  $y$ , and  $A$  is the surface outcrop area for all densities greater than



479 FIG. 11. Surface-forced water mass transformation, Eq. (4), in the North Atlantic (north of  $30^\circ\text{N}$ ) in (a)  
 480 CESM1 and (b) CNRM-CM6.1. CESM1 simulates a strong AMOC recovery whereas the AMOC in CNRM-  
 481 CM6.1 essentially has no recovery over the period simulated. Surface transformation is calculated as a function of  
 482 potential density referenced to the surface ( $\sigma_0$ ). The total transformation from all diabatic processes is provided  
 483 in black (top), which sums contributions from heat (middle) and freshwater (bottom). The solid lines indicates  
 484 the initial state (years 1 – 10) and the dashed line indicates the years 970 – 1000. The light-colored lines denote  
 485 the transformation rates for the years 970 – 1000 but with either the (dotted) initial density field or (dash-dotted)  
 486 initial surface flux field. For visualization, approximate geographical boundaries are labeled. Here, positive  
 487 transformation represents a volume flux towards denser classes.

503 a given density,  $\sigma$ . Eq. (4) can be decomposed into contributions to the buoyancy flux from heat  
 504 and freshwater.

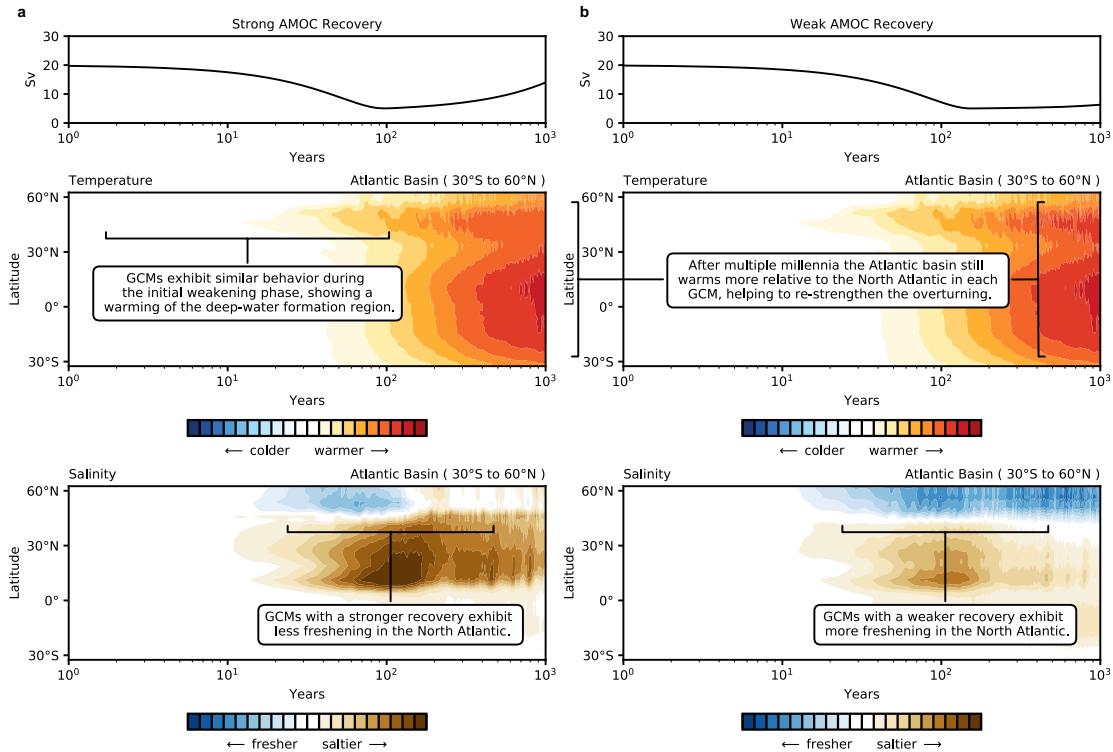
505 Applying Eq. (4) to CESM1 and CNRM-CM6.1 indicates that the formation of NADW is largely  
 506 heat-driven and that changes to the formation rate of NADW are also largely heat-driven (Fig. 11).

507 During the initial state, CESM1 and CNRM-CM6.1 both have two peaks of transformation rates  
508 around  $\sigma_0 = 1024$  and  $1027 \text{ kg m}^{-3}$ . Approximately 1000 years after the  $4 \times \text{CO}_2$  forcing, surface-  
509 forced water mass transformation occurs in lighter density classes for each GCM, primarily because  
510 the ocean warms. During the re-strengthening phase (years 970 – 1000), when CESM1 shows a  
511 recovered AMOC and CNRM-CM6.1 shows a weakened AMOC, the peak transformation rate  
512 around  $1027 \text{ kg m}^{-3}$  is notably absent in CNRM-CM6.1. To understand causes of these changes,  
513 we compute the surface-forced water mass transformation using either the initial density field and  
514 the surface flux field from years 970 – 1000 or the initial surface flux field and the density field from  
515 years 970 – 1000. This enables us to identify mechanisms of NADW changes. A decomposition of  
516 these changes due to either changes in the surface density field (dotted line) and the surface fluxes  
517 (dash-dotted) shows that this decrease in the transformation rate is largely the result of changes in  
518 the surface density field. In CNRM-CM6.1, the sustained freshening limits the ability of convection  
519 associated with heat loss to penetrate deep into the water column. However, in CESM1 changes to  
520 the density field are much smaller, enabling the heat loss to penetrate to greater depth.

521 The LongRunMIP repository lacks sufficient output to apply Eq. (4) to all GCMs. Still, these  
522 surface-forced water mass transformation results support the idea of large freshwater anomalies  
523 in the subsurface ocean inhibiting deep convection and preventing an AMOC recovery. These  
524 transformation rates show that the competition between a high-latitude freshwater anomaly and  
525 low-latitude salinity anomaly can determine the timing and magnitude of an AMOC recovery.  
526 Further work, however, is required to understand how the formation of surface waters interacts  
527 with local and nonlocal processes to suppress deep-water formation.

### 528 *b. Mechanisms for a partial recovery of the Atlantic overturning circulation*

542 Our results indicate that the recovery of the AMOC is associated with two simultaneous processes:  
543 (1) a warming of the Atlantic basin relative to the North Atlantic and (2) a subtropical salinity  
544 anomaly that moves poleward. The former enhances the Atlantic meridional density difference,  
545 while the latter erodes the near-surface stratification and reinvigorates deep convection. Figure 12  
546 shows a schematic of the processes that contribute to the time-evolution of the AMOC in response  
547 to  $4 \times \text{CO}_2$  in coupled GCMs. During the initial weakening, all GCMs show a warming of the  
548 North Atlantic water column, though some GCMs warm more (and the AMOC weakens more)



529 FIG. 12. Schematic depicting a GCM with a (a) strong AMOC recovery and (b) weak AMOC recovery or a  
 530 prolonged weakening AMOC phase. The middle panel denotes the temperature changes in the Atlantic basin and  
 531 the bottom panel denotes the salinity changes in the Atlantic basin. There are three phases: a weakening phase,  
 532 which is characterized by a warming of the high latitude North Atlantic; a partial recovery phase or prolonged  
 533 weakening phase, which is characterized as a warming of the Atlantic basin and an increase or decrease in  
 534 salinity of the high latitude North Atlantic; and a re-strengthening phase, which is characterized by a warming of  
 535 the low-latitude Atlantic basin. Warming of the low-latitude Atlantic basin is similar across all GCMs.  
 536 The major feature that distinguishes GCMs with faster and slower AMOC recoveries is the salinity anomalies in  
 537 the high-latitude North Atlantic. A subtropical salinity anomaly forms when the AMOC is weakened and then  
 538 is advected to the high-latitudes by the subtropical gyre. GCMs with slower AMOC recoveries instead show  
 539 a freshening of the high-latitude North Atlantic, which is linked to large differences in sea ice melt across the  
 540 GCMs. This suggests that the representation of sea ice and other high-latitude freshwater sources in coupled  
 541 GCMs is critical for representing the future evolution of the AMOC.

549 than others. After the initial weakening, GCMs with a strong AMOC recovery (e.g., CESM1) have  
 550 positive salinity anomaly in the North Atlantic, whereas GCMs with a weak AMOC recovery (e.g.,

551 CNRM-CM6.1 and HadGEM2-ES) have either a weak positive salinity anomaly or a negative  
552 anomaly indicative of high latitude freshening (Fig. 12). All GCMs show greater warming of the  
553 Atlantic basin relative to the North Atlantic, but the degree of high-latitude freshening correlates  
554 strongly with the level of the AMOC recovery. This highlights the role that high-latitude freshwater  
555 forcing plays in generating differences in the magnitude and rate of AMOC recovery across GCMs  
556 in response to the same external forcing.

557 Our results differ from similar studies that have examined the transient and equilibrium responses  
558 of the AMOC to surface warming in ocean-only models (Jansen et al. 2018; Jansen and Nadeau  
559 2019) and coupled GCMs (Stouffer and Manabe 2003). For instance, the recovery of the AMOC in  
560 ocean-only models is attributed to a density reduction in the lower thermocline of the high-latitudes  
561 (Jansen et al. 2018; Jansen and Nadeau 2019), but it is unclear whether this modification results  
562 from temperature or salinity changes as these models resolve buoyancy explicitly. Here we find  
563 that salinity anomalies reduce high-latitude stratification and reinvigorate convection, which is  
564 consistent with other coupled GCM studies (Vellinga et al. 2002; Yin and Stouffer 2007; Krebs  
565 and Timmermann 2007; Wu et al. 2011; Sigmond et al. 2020; Zhu and Liu 2020; Ackermann  
566 et al. 2020). However, the role of salinity is more nuanced than previously suggested. Our results  
567 indicate that the interplay of temperature changes, which warm the Atlantic basin, and salinity  
568 anomalies, which densify the high-latitude ocean, set the magnitude of the AMOC recovery. This  
569 stands in contrast to Wu et al. (2011) and Sigmond et al. (2020), which argue that salinity is the  
570 primary driver of the AMOC recovery. Given that a convergence of salinity in the Atlantic basin  
571 is apparent in observations (e.g., Zhu and Liu 2020), it is possible the AMOC will experience a  
572 recovery beyond the 21st century.

573 The equilibrium response of the AMOC in ocean-only models has been attributed to a diffusive  
574 adjustment of abyssal density, in response to a modification in bottom waters exported from  
575 the Southern Ocean (Jansen et al. 2018). Given that the LongRunMIP simulations are at most  
576 5900 years long, and some have still not reached a new steady-state, it is difficult to attribute  
577 the equilibrium response to precise mechanisms, like localized basin dynamics versus inter-basin  
578 transport or Southern Ocean processes. The re-strengthening phase of the AMOC is indeed linked  
579 to a warming of the low-latitude Atlantic basin based on the TWR, but this does not necessarily  
580 imply that the AMOC evolution results from processes occurring in the Atlantic basin alone.

581 Recently, Newsom and Thompson (2018) and Holmes et al. (2019) have revisited the ocean's  
582 buoyancy and heat budget, respectively, and found that much of the heat loss in the North Atlantic  
583 is balanced by heat uptake in the low-latitude Indo-Pacific, rather than the low-latitude Atlantic.  
584 This balance is enabled by the much larger areal extent of the low-latitude Indo-Pacific, and  
585 heat is subsequently transferred to the Atlantic through the upper ocean's inter-basin overturning  
586 circulation (Gordon 1986; Newsom et al. 2021). The dynamics of inter-basin exchange, mediated  
587 by both the Southern Ocean, and low-latitude pathways such as the Indonesian Throughflow and  
588 the Aghulas Retroflexion, over millennial timescales remains largely unstudied.

### 589 *c. Utility and limitations of LongRunMIP*

590 The LongRunMIP simulations provide the first model inter-comparison of the long-term evolu-  
591 tion of the AMOC in response to abrupt forcing. Even so, the longest simulation in this study is  
592 5900 years. Recent work has shown that the AMOC continues to change, following warming, on  
593 timescales out to 10,000 – 15,000 years (Rugenstein et al. 2016b; Jansen et al. 2018; Jansen and  
594 Nadeau 2019). Thus, even the LongRunMIP simulations are too short to capture all of the pro-  
595 cesses that influence the AMOC evolution, especially long-term mechanisms such as the diffusive  
596 adjustment of the deep ocean (Jansen et al. 2018). It may be interesting to investigate changes to  
597 the AMOC in coupled GCMs subject to abrupt forcing beyond 10,000 years, as this may enable a  
598 better theoretical understanding of past and future AMOC changes.

599 Because LongRunMIP is a MIP of opportunity, without an agreed-upon protocol, model output  
600 is not uniformly reported making it hard to precisely attribute the intermodel spread in the AMOC  
601 recovery to a specific process. For instance, it would be helpful to have ocean velocity fields as  
602 the relative role of changes in surface forcing versus changes in salinity and temperature due to  
603 ocean circulation changes could be quantified (e.g., Garuba and Rasch 2020). Similarly, the lack  
604 of uniform output makes it difficult to partition freshwater changes into  $P - E$ , runoff, and ice  
605 fluxes. Indeed, GCMs with a slower AMOC recovery tend to have greater Arctic sea ice loss and  
606 high-latitude freshwater input, which likely limits the ability of the subtropical salinity anomaly  
607 to reduce high-latitude stratification and reinvigorate convection. However, additional work is  
608 required to assess the gradual recovery of the AMOC and the role of Arctic sea ice loss, or other  
609 freshwater sources such as changes to the hydrologic cycle or ice sheet processes. CNRM-CM6.1,

610 for instance, which experiences a prolonged AMOC weakening, tends to warm much more than  
611 other GCMs (Rugenstein et al. 2019), implying that changes to  $P - E$  may be larger in the high-  
612 latitudes following the “wet gets wetter, dry gets drier” paradigm (Held and Soden 2006). It may  
613 be useful to investigate the relative role of temperature and salinity dynamics in models similar to  
614 the toy models of Jansen and Nadeau (2019) and Nadeau and Jansen (2020), as these models can  
615 be used to explore parameter space that can impact long-term ocean circulation changes. Such  
616 work may also permit a better theoretical understanding of how salinity and temperature interact  
617 and influence the response of the AMOC to external forcing.

618 Finally, an important caveat of this work is that the GCMs do not resolve ocean mesoscale  
619 eddies or dynamics of narrow, boundary currents in the North Atlantic. In response to warming  
620 and freshwater perturbations, higher resolution GCMs exhibit markedly different changes to ocean  
621 circulation (Mecking et al. 2016; Newsom et al. 2016; Gent 2018; Jackson et al. 2020; Hirschi et al.  
622 2020), and often experience less weakening of the AMOC when compared to low-resolution GCMs  
623 (e.g., Li et al. 2021; Jackson et al. 2020). Further investigation of the transient and equilibrium  
624 responses of the AMOC to increased greenhouse-gas concentrations with high-resolution eddy-  
625 resolving GCMs may modify the relative importance of haline and thermal contributions to AMOC  
626 changes at different timescales. For example, higher resolution ocean models have been shown to  
627 change the magnitude of the subtropical gyre transport (e.g., Lévy et al. 2010), which could impact  
628 the northward advection of salinity that helps to erode near-surface stratification and reinvigorate  
629 deep convection.

## 630 **6. Conclusions**

631 This study examines the transient and equilibrium responses of the AMOC to warming in an  
632 unprecedented collection of millennial-length climate simulations from atmosphere-ocean GCMs.  
633 We introduce a framework for understanding the various AMOC responses in each GCM using  
634 a simple thermal wind expression, which relates the overturning circulation in the North Atlantic  
635 to the density difference between the region of deep-water formation and the Atlantic basin. We  
636 find that this expression captures the initial weakening of the AMOC on centennial timescales and  
637 various levels of AMOC recoveries as simulated by GCMs in response to  $4 \times \text{CO}_2$ .

638 Using this expression, we attribute the evolution of the AMOC on different timescales to changes  
639 in temperature and salinity in distinct regions. The universal weakening of the AMOC across  
640 GCMs is dominated by warming of the northern convection region. The partial recovery of the  
641 AMOC, which occurs in most GCMs, results from both a simultaneous warming of the Atlantic  
642 basin and an increase in salinity in the North Atlantic. The positive salinity anomaly forms in the  
643 subtropical Atlantic basin when the AMOC is weakened and is then advected northward, which  
644 helps to erode the stratification of the North Atlantic and reinvigorate deep convection. We find  
645 that the timing and magnitude of the AMOC recovery, at least in the LongRunMIP simulations,  
646 is related to the amount freshwater input in the North Atlantic: GCMs with a slower AMOC  
647 recovery tend to show increased Arctic sea ice loss compared to GCMs with a faster AMOC  
648 recovery. This suggests it will be crucial to monitor Arctic sea ice loss and other high-latitude  
649 freshwater sources throughout and beyond the 21st century, as it may ultimately determine whether  
650 the AMOC will recover. For example, Arctic sea ice has been shown to be more sensitive to  
651 warming in observations when compared to most GCMs (Winton 2011), implying that there  
652 may be more freshwater input from sea ice melt during the coming century than GCMs currently  
653 suggest. Finally, the eventual strengthening of the AMOC, which takes multiple millennia, is linked  
654 to low-latitude warming of the Atlantic basin. For the simulations that span multiple millennia,  
655 the AMOC becomes stronger than the initial state despite having vastly different distributions of  
656 temperature and salinity. This work provides a framework for assessing the relative importance of  
657 temperature and salinity dynamics in the evolution of the AMOC. A similar approach may help  
658 identify whether the AMOC is likely to experience a fast or slow recovery beyond the 21st century,  
659 or if the AMOC has experienced a slow recovery during past climate changes.

660 *Acknowledgments.* The research would not have been possible without the efforts of the contrib-  
661 utors to the LongRunMIP project, which is freely available at <http://www.longrunmip.org/>. The  
662 authors thank Sarah Ragen for helpful comments on an earlier version of this paper. D.B.B. was  
663 supported by an American Meteorological Society (AMS) Graduate Fellowship and the National  
664 Science Foundation Graduate Research Fellowship Program (NSF Grant DGE-1745301). A.F.T.  
665 and S.S. were supported by NSF Grant OCE-1756956 and OCE-2023259. E.R.N was supported by  
666 NERC Project NE/P019218/1 on transient tracer-based Investigation of Circulation and Thermal  
667 Ocean Change (TICTOC).

668 *Data availability statement.* The data for this study will be made available upon acceptance at  
669 <https://github.com/dbonan>.

## 670 **References**

- 671 Ackermann, L., C. Danek, P. Gierz, and G. Lohmann, 2020: AMOC Recovery in a Multicentennial  
672 Scenario Using a Coupled Atmosphere-Ocean-Ice Sheet Model. *Geophysical Research Letters*,  
673 **47 (16)**, e2019GL086810.
- 674 Andrews, T., J. M. Gregory, and M. J. Webb, 2015: The dependence of radiative forcing and  
675 feedback on evolving patterns of surface temperature change in climate models. *Journal of*  
676 *Climate*, **28 (4)**, 1630–1648.
- 677 Bakker, P., and Coauthors, 2016: Fate of the Atlantic Meridional Overturning Circulation: Strong  
678 decline under continued warming and Greenland melting. *Geophysical Research Letters*, **43 (23)**,  
679 12–252.
- 680 Broecker, W. S., 1997: Thermohaline circulation, the Achilles heel of our climate system: Will  
681 man-made CO<sub>2</sub> upset the current balance? *Science*, **278 (5343)**, 1582–1588.
- 682 Broecker, W. S., 1998: Paleocean circulation during the last deglaciation: a bipolar seesaw?  
683 *Paleoceanography*, **13 (2)**, 119–121.
- 684 Broecker, W. S., 2003: Does the trigger for abrupt climate change reside in the ocean or in the  
685 atmosphere? *Science*, **300 (5625)**, 1519–1522.
- 686 Buckley, M. W., and J. Marshall, 2016: Observations, inferences, and mechanisms of the Atlantic  
687 Meridional Overturning Circulation: A review. *Reviews of Geophysics*, **54 (1)**, 5–63.
- 688 Burke, A., A. L. Stewart, J. F. Adkins, R. Ferrari, M. F. Jansen, and A. F. Thompson, 2015:  
689 The glacial mid-depth radiocarbon bulge and its implications for the overturning circulation.  
690 *Paleoceanography*, **30 (7)**, 1021–1039.
- 691 Butler, E., K. Oliver, J.-M. Hirschi, and J. Mecking, 2016: Reconstructing global overturning from  
692 meridional density gradients. *Climate Dynamics*, **46 (7-8)**, 2593–2610.
- 693 Cheng, W., J. C. Chiang, and D. Zhang, 2013: Atlantic meridional overturning circulation (AMOC)  
694 in CMIP5 models: RCP and historical simulations. *Journal of Climate*, **26 (18)**, 7187–7197.

- 695 Curry, W. B., and D. W. Oppo, 2005: Glacial water mass geometry and the distribution of  $\delta^{13}\text{C}$   
696 of  $\Sigma\text{CO}_2$  in the western Atlantic Ocean. *Paleoceanography*, **20** (1).
- 697 De Boer, A. M., A. Gnanadesikan, N. R. Edwards, and A. J. Watson, 2010: Meridional density  
698 gradients do not control the Atlantic overturning circulation. *Journal of Physical Oceanography*,  
699 **40** (2), 368–380.
- 700 Dixon, K. W., T. L. Delworth, M. J. Spelman, and R. J. Stouffer, 1999: The influence of transient  
701 surface fluxes on North Atlantic overturning in a coupled GCM climate change experiment.  
702 *Geophysical Research Letters*, **26** (17), 2749–2752.
- 703 Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth  
704 System Model: from CMIP3 to CMIP5. *Climate dynamics*, **40** (9), 2123–2165.
- 705 Ferrari, R., M. F. Jansen, J. F. Adkins, A. Burke, A. L. Stewart, and A. F. Thompson, 2014:  
706 Antarctic sea ice control on ocean circulation in present and glacial climates. *Proceedings of the*  
707 *National Academy of Sciences*, **111** (24), 8753–8758.
- 708 Frierson, D. M., and Coauthors, 2013: Contribution of ocean overturning circulation to tropical  
709 rainfall peak in the Northern Hemisphere. *Nature Geoscience*, **6** (11), 940–944.
- 710 Frölicher, T., M. Aschwanden, N. Gruber, S. Jaccard, J. Dunne, and D. Paynter, 2020: Contrasting  
711 upper and deep ocean oxygen response to protracted global warming. *Global Biogeochemical*  
712 *Cycles*, **34** (8), e2020GB006601.
- 713 Ganachaud, A., and C. Wunsch, 2003: Large-scale ocean heat and freshwater transports during  
714 the World Ocean Circulation Experiment. *Journal of Climate*, **16** (4), 696–705.
- 715 Garuba, O. A., and P. J. Rasch, 2020: A partial coupling method to isolate the roles of the  
716 atmosphere and ocean in coupled climate simulations. *Journal of Advances in Modeling Earth*  
717 *Systems*, **12** (9), e2019MS002016.
- 718 Gebbie, G., 2014: How much did glacial North Atlantic water shoal? *Paleoceanography*, **29** (3),  
719 190–209.
- 720 Gent, P. R., 2018: A commentary on the Atlantic meridional overturning circulation stability in  
721 climate models. *Ocean Modelling*, **122**, 57–66.

- 722 Gnanadesikan, A., 1999: A simple predictive model for the structure of the oceanic pycnocline.  
723 *Science*, **283 (5410)**, 2077–2079.
- 724 Gordon, A. L., 1986: Interocean exchange of thermocline water. *Journal of Geophysical Research:*  
725 *Oceans*, **91 (C4)**, 5037–5046.
- 726 Gregory, J., and Coauthors, 2005: A model intercomparison of changes in the Atlantic thermohaline  
727 circulation in response to increasing atmospheric CO<sub>2</sub> concentration. *Geophysical Research*  
728 *Letters*, **32 (12)**.
- 729 Groeskamp, S., S. M. Griffies, D. Iudicone, R. Marsh, A. G. Nurser, and J. D. Zika, 2019: The  
730 water mass transformation framework for ocean physics and biogeochemistry. *Annual review of*  
731 *marine science*, **11**, 271–305.
- 732 Haskins, R. K., K. I. Oliver, L. C. Jackson, S. S. Drijfhout, and R. A. Wood, 2019: Explaining  
733 asymmetry between weakening and recovery of the AMOC in a coupled climate model. *Climate*  
734 *dynamics*, **53 (1)**, 67–79.
- 735 Haskins, R. K., K. I. Oliver, L. C. Jackson, R. A. Wood, and S. S. Drijfhout, 2020: Temperature  
736 domination of AMOC weakening due to freshwater hosing in two GCMs. *Climate Dynamics*,  
737 **54 (1)**, 273–286.
- 738 He, C., Z. Liu, and A. Hu, 2019: The transient response of atmospheric and oceanic heat transports  
739 to anthropogenic warming. *Nature Climate Change*, **9 (3)**, 222–226.
- 740 Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming.  
741 *Journal of climate*, **19 (21)**, 5686–5699.
- 742 Hirschi, J. J.-M., and Coauthors, 2020: The Atlantic meridional overturning circulation in high-  
743 resolution models. *Journal of Geophysical Research: Oceans*, **125 (4)**, e2019JC015 522.
- 744 Holmes, R. M., J. D. Zika, R. Ferrari, A. F. Thompson, E. R. Newsom, and M. H. England, 2019:  
745 Atlantic Ocean Heat Transport Enabled by Indo-Pacific Heat Uptake and Mixing. *Geophysical*  
746 *Research Letters*, **46 (23)**, 13 939–13 949.
- 747 Jackson, L., 2013: Shutdown and recovery of the AMOC in a coupled global climate model: the  
748 role of the advective feedback. *Geophysical Research Letters*, **40 (6)**, 1182–1188.

- 749 Jackson, L. C., and Coauthors, 2020: Impact of ocean resolution and mean state on the rate of  
750 AMOC weakening. *Climate Dynamics*, **55** (7), 1711–1732.
- 751 Jansen, M. F., and L.-P. Nadeau, 2019: A toy model for the response of the residual overturning  
752 circulation to surface warming. *Journal of Physical Oceanography*, **49** (5), 1249–1268.
- 753 Jansen, M. F., L.-P. Nadeau, and T. M. Merlis, 2018: Transient versus equilibrium response of the  
754 ocean’s overturning circulation to warming. *Journal of Climate*, **31** (13), 5147–5163.
- 755 Kostov, Y., K. C. Armour, and J. Marshall, 2014: Impact of the Atlantic meridional overturning  
756 circulation on ocean heat storage and transient climate change. *Geophysical Research Letters*,  
757 **41** (6), 2108–2116.
- 758 Krebs, U., and A. Timmermann, 2007: Tropical air–sea interactions accelerate the recovery of the  
759 Atlantic meridional overturning circulation after a major shutdown. *Journal of Climate*, **20** (19),  
760 4940–4956.
- 761 Levang, S. J., and R. W. Schmitt, 2020: What Causes the AMOC to Weaken in CMIP5? *Journal*  
762 *of Climate*, **33** (4), 1535–1545.
- 763 Lévy, M., P. Klein, A.-M. Tréguier, D. Iovino, G. Madec, S. Masson, and K. Takahashi, 2010:  
764 Modifications of gyre circulation by sub-mesoscale physics. *Ocean Modelling*, **34** (1-2), 1–15.
- 765 Li, H., A. Fedorov, and W. Liu, 2021: AMOC stability and diverging response to Arctic sea ice  
766 decline in two climate models. *Journal of Climate*, 1–47.
- 767 Liu, W., A. V. Fedorov, S.-P. Xie, and S. Hu, 2020: Climate impacts of a weakened Atlantic  
768 Meridional Overturning Circulation in a warming climate. *Science Advances*, **6** (26), eaaz4876.
- 769 Liu, W., S.-P. Xie, Z. Liu, and J. Zhu, 2017: Overlooked possibility of a collapsed Atlantic  
770 Meridional Overturning Circulation in warming climate. *Science Advances*, **3** (1), e1601 666.
- 771 Lynch-Stieglitz, J., and Coauthors, 2007: Atlantic meridional overturning circulation during the  
772 Last Glacial Maximum. *science*, **316** (5821), 66–69.
- 773 Manabe, S., and R. J. Stouffer, 1994: Multiple-century response of a coupled ocean-atmosphere  
774 model to an increase of atmospheric carbon dioxide. *Journal of climate*, **7** (1), 5–23.

775 Maroon, E. A., J. E. Kay, and K. B. Karnauskas, 2018: Influence of the Atlantic meridional  
776 overturning circulation on the Northern Hemisphere surface temperature response to radiative  
777 forcing. *Journal of Climate*, **31** (22), 9207–9224.

778 Marshall, J., A. Donohoe, D. Ferreira, and D. McGee, 2014: The ocean’s role in setting the mean  
779 position of the Inter-Tropical Convergence Zone. *Climate Dynamics*, **42** (7-8), 1967–1979.

780 Mauritsen, T., and Coauthors, 2019: Developments in the MPI-M Earth System Model version  
781 1.2 (MPI-ESM1.2) and its response to increasing CO<sub>2</sub>. *Journal of Advances in Modeling Earth*  
782 *Systems*, **11** (4), 998–1038.

783 McDougall, T. J., and P. M. Barker, 2011: *Getting started with TEOS-10 and the Gibbs Seawater*  
784 *(GSW) oceanographic toolbox*, Vol. 37.

785 Mecking, J., S. S. Drijfhout, L. C. Jackson, and T. Graham, 2016: Stable AMOC off state in an  
786 eddy-permitting coupled climate model. *Climate Dynamics*, **47** (7), 2455–2470.

787 Nadeau, L.-P., and M. F. Jansen, 2020: Overturning Circulation Pathways in a Two-Basin Ocean  
788 Model. *Journal of Physical Oceanography*, **50** (8), 2105–2122.

789 Newsom, E. R., C. M. Bitz, F. O. Bryan, R. Abernathey, and P. R. Gent, 2016: Southern Ocean  
790 deep circulation and heat uptake in a high-resolution climate model. *Journal of Climate*, **29** (7),  
791 2597–2619.

792 Newsom, E. R., and A. F. Thompson, 2018: Reassessing the role of the Indo-Pacific in the ocean’s  
793 global overturning circulation. *Geophysical Research Letters*, **45** (22), 12–422.

794 Newsom, E. R., A. F. Thompson, J. F. Adkins, and E. D. Galbraith, 2021: A hemispheric  
795 asymmetry in poleward ocean heat transport across climates: Implications for overturning and  
796 polar warming. *Earth and Planetary Science Letters*, **568**, 117 033.

797 Nikurashin, M., and G. Vallis, 2012: A theory of the interhemispheric meridional overturning  
798 circulation and associated stratification. *Journal of Physical Oceanography*, **42** (10), 1652–  
799 1667.

800 Ragen, S., K. Armour, L. Thompson, A. Shao, and D. Darr, Submitted: The role of ocean basin  
801 geometry in Atlantic meridional overturning circulation. *Journal of Climate*.

- 802 Rugenstein, M., and Coauthors, 2019: LongRunMIP: Motivation and design for a large collection  
803 of millennial-length AOGCM simulations. *Bulletin of the American Meteorological Society*,  
804 **100 (12)**, 2551–2570.
- 805 Rugenstein, M., and Coauthors, 2020: Equilibrium climate sensitivity estimated by equilibrating  
806 climate models. *Geophysical Research Letters*, **47 (4)**, e2019GL083 898.
- 807 Rugenstein, M. A., J. M. Gregory, N. Schaller, J. Sedláček, and R. Knutti, 2016a: Multiannual  
808 ocean–atmosphere adjustments to radiative forcing. *Journal of Climate*, **29 (15)**, 5643–5659.
- 809 Rugenstein, M. A., J. Sedláček, and R. Knutti, 2016b: Nonlinearities in patterns of long-term  
810 ocean warming. *Geophysical Research Letters*, **43 (7)**, 3380–3388.
- 811 Saint-Martin, D., and Coauthors, 2019: Fast-Forward to Perturbed Equilibrium Climate. *Geophys-  
812 ical Research Letters*, **46 (15)**, 8969–8975.
- 813 Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2  
814 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, **6 (1)**,  
815 141–184.
- 816 Schmittner, A., M. Latif, and B. Schneider, 2005: Model projections of the North Atlantic thermo-  
817 haline circulation for the 21st century assessed by observations. *Geophysical research letters*,  
818 **32 (23)**.
- 819 Sévellec, F., A. V. Fedorov, and W. Liu, 2017: Arctic sea-ice decline weakens the Atlantic  
820 meridional overturning circulation. *Nature Climate Change*, **7 (8)**, 604–610.
- 821 Sigmond, M., J. C. Fyfe, O. A. Saenko, and N. C. Swart, 2020: Ongoing AMOC and related  
822 sea-level and temperature changes after achieving the Paris targets. *Nature Climate Change*,  
823 **10 (7)**, 672–677.
- 824 Speer, K., and E. Tziperman, 1992: Rates of water mass formation in the North Atlantic Ocean.  
825 *Journal of Physical Oceanography*, **22 (1)**, 93–104.
- 826 Stocker, T. F., and A. Schmittner, 1997: Influence of CO<sub>2</sub> emission rates on the stability of the  
827 thermohaline circulation. *Nature*, **388 (6645)**, 862–865.

828 Stommel, H., 1961: Thermohaline convection with two stable regimes of flow. *Tellus*, **13** (2),  
829 224–230.

830 Stouffer, R., and S. Manabe, 2003: Equilibrium response of thermohaline circulation to large  
831 changes in atmospheric CO<sub>2</sub> concentration. *Climate Dynamics*, **20** (7-8), 759–773.

832 Stouffer, R. J., and Coauthors, 2006: Investigating the causes of the response of the thermohaline  
833 circulation to past and future climate changes. *Journal of climate*, **19** (8), 1365–1387.

834 Sun, S., A. F. Thompson, and I. Eisenman, 2020: Transient overturning compensation between  
835 Atlantic and Indo-Pacific basins. *Journal of Physical Oceanography*, **50** (8), 2151–2172.

836 Thomas, M. D., and A. V. Fedorov, 2019: Mechanisms and impacts of a partial AMOC recovery  
837 under enhanced freshwater forcing. *Geophysical Research Letters*, **46** (6), 3308–3316.

838 Thorpe, R., J. M. Gregory, T. Johns, R. Wood, and J. Mitchell, 2001: Mechanisms determining  
839 the Atlantic thermohaline circulation response to greenhouse gas forcing in a non-flux-adjusted  
840 coupled climate model. *Journal of Climate*, **14** (14), 3102–3116.

841 Vellinga, M., R. A. Wood, and J. M. Gregory, 2002: Processes governing the recovery of a  
842 perturbed thermohaline circulation in HadCM3. *Journal of Climate*, **15** (7), 764–780.

843 Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the  
844 ocean. *Tellus*, **34** (2), 187–195.

845 Weaver, A. J., M. Eby, M. Kienast, and O. A. Saenko, 2007: Response of the Atlantic merid-  
846 ional overturning circulation to increasing atmospheric CO<sub>2</sub>: Sensitivity to mean climate state.  
847 *Geophysical Research Letters*, **34** (5).

848 Weber, S., and Coauthors, 2007: The modern and glacial overturning circulation in the Atlantic  
849 ocean in PMIP coupled model simulations. *Climate of the Past*, **3** (1), 51–64.

850 Weijer, W., W. Cheng, O. A. Garuba, A. Hu, and B. Nadiga, 2020: CMIP6 models predict significant  
851 21st century decline of the Atlantic Meridional Overturning Circulation. *Geophysical Research*  
852 *Letters*, **47** (12), e2019GL086075.

- 853 Welander, P., 1971: A discussion on ocean currents and their dynamics-The thermocline problem.  
854 *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical*  
855 *Sciences*, **270 (1206)**, 415–421.
- 856 Wiebe, E., and A. Weaver, 1999: On the sensitivity of global warming experiments to the parametri-  
857 sation of sub-grid scale ocean mixing. *Climate Dynamics*, **15 (12)**, 875–893.
- 858 Winton, M., 2011: Do climate models underestimate the sensitivity of Northern Hemisphere sea  
859 ice cover? *Journal of Climate*, **24 (15)**, 3924–3934.
- 860 Wolfe, C. L., and P. Cessi, 2010: What sets the strength of the middepth stratification and  
861 overturning circulation in eddying ocean models? *Journal of Physical Oceanography*, **40 (7)**,  
862 1520–1538.
- 863 Wu, P., L. Jackson, A. Pardaens, and N. Schaller, 2011: Extended warming of the northern high  
864 latitudes due to an overshoot of the Atlantic meridional overturning circulation. *Geophysical*  
865 *research letters*, **38 (24)**.
- 866 Yin, J., and R. J. Stouffer, 2007: Comparison of the stability of the Atlantic thermohaline circulation  
867 in two coupled atmosphere–ocean general circulation models. *Journal of Climate*, **20 (17)**, 4293–  
868 4315.
- 869 Zhu, C., and Z. Liu, 2020: Weakening Atlantic overturning circulation causes South Atlantic  
870 salinity pile-up. *Nature Climate Change*, **10 (11)**, 998–1003.