



The role of anthropogenic aerosol forcing in the 1850–1985 strengthening of the AMOC in CMIP6 historical simulations

Jon Robson,^a Matthew B. Menary,^b Rowan T. Sutton,^a Jenny Mecking,^c Jonathan M. Gregory,^a Colin Jones,^d Bablu Sinha,^c David P. Stevens,^e and Laura J. Wilcox^a

^a *National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK.*

^b *LOCEAN, Sorbonne Université, Paris, France.*

^c *National Oceanography Centre, Southampton, UK.*

^d *National Centre for Atmospheric Science, University of Leeds, Leeds, UK.*

^e *Centre for Ocean and Atmospheric Sciences, School of Mathematics, University of East Anglia, Norwich, UK.*

Corresponding author: Jon Robson, j.i.robson@reading.ac.uk

Early Online Release: This preliminary version has been accepted for publication in *Journal of Climate*, may be fully cited, and has been assigned DOI 10.1175/JCLI-D-22-0124.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT: Previous work has shown that anthropogenic aerosol (AA) forcing drives a strengthening in the Atlantic Meridional Overturning Circulation (AMOC) in CMIP6 historical simulations over 1850–1985, but the mechanisms have not been fully understood. Across CMIP6 models, it is shown that there is a strong correlation between surface heat loss over the subpolar North Atlantic (SPNA) and the forced strengthening of the AMOC. Despite the link to AA forcing, the AMOC response is not strongly related to the contribution of anomalous downwelling surface shortwave radiation to SPNA heat loss. Rather, the spread in AMOC response is primarily due to the spread in turbulent heat loss. We hypothesize that turbulent heat loss is larger in models with strong AA forcing because the air advected over the ocean is colder and drier, in turn because of greater AA forced cooling over the continents upwind, especially North America. The strengthening of the AMOC also feeds back on itself positively in two distinct ways: by raising the sea surface temperature and hence further increasing turbulent heat loss in the SPNA, and by increasing the sea surface density across the SPNA due to increased northward transport of saline water. A comparison of key indices suggests that the AMOC response in models with strong AA forcing is not likely to be consistent with observations.

1. Introduction

The Atlantic Meridional Overturning Circulation (AMOC) is an important component of the Earth's climate. It plays a key role in the poleward transport of heat within the Northern Hemisphere (Trenberth et al. 2019) and across the Equator (Marshall et al. 2014), and is correlated with the sequestration of heat and carbon in the deep ocean (Sarmiento and Le Quéré 1996; Kostov et al. 2014). Moreover, a range of observational, paleo, and modelling work has linked changes in the strength of the AMOC, and related ocean heat transports, to a long list of regional and global climate impacts including impacts on sea surface temperature (SST), rainfall, hurricanes, and sea level (Yin et al. 2009; Zhang et al. 2019; Bellomo et al. 2021). Therefore, it is important to improve our understanding of how the AMOC can shape regional and global climate in order to reduce uncertainty in climate predictions.

The AMOC is known to vary on a range of timescales and due to a range of processes. For example, changes in the strength of the wind stress are known to be important on daily to decadal timescales through its impact on Ekman currents and Ekman upwelling/downwelling (Polo et al. 2014; Blaker et al. 2015; Buckley and Marshall 2016; Jackson et al. 2022). AMOC variability on decadal and longer timescales has often been linked to changes in surface buoyancy fluxes and their impact on dense water formation in the North Atlantic (Robson et al. 2012; Grist et al. 2014; Gregory et al. 2016; Delworth and Zeng 2016; Xu et al. 2019; Yeager et al. 2021; Megann et al. 2021; Jackson et al. 2022). Climate models indicate that internal variability is an important source of AMOC variability on daily to centennial timescales (Buckley and Marshall 2016; Jackson et al. 2022). However, changes in natural and anthropogenic external forcings can also drive AMOC variability. For example, changes in solar radiation and volcanic emissions have been shown to induce decadal-to-multidecadal variability in the AMOC (Otterå et al. 2010; Menary and Scaife 2014; Swingedouw et al. 2015). In the future, the AMOC is expected to weaken significantly as an increasingly warmer climate weakens the surface buoyancy forcing of the AMOC (Weijer et al. 2020; Couldrey et al. 2021).

A recent focus has been on understanding the relative role of different anthropogenic forcings in explaining past AMOC change. In particular, over the historical period there has been a competition between greenhouse gas (GHG) forcing and anthropogenic aerosol (AA) forcing. Many studies have highlighted that AA forcing can have a large impact on the North Atlantic, and AA forcing has

been shown to strengthen the AMOC in climate models (Cai et al. 2006; Delworth and Dixon 2006; Menary et al. 2013; Undorf et al. 2018; Andrews et al. 2020; Menary et al. 2020; Robson et al. 2020; Hassan et al. 2021; Fiedler and Putrasahan 2021). Previously these competing anthropogenic effects were thought to lead to relatively little externally-forced change over the historical period (Delworth and Dixon 2006; Cheng et al. 2013). However, some models with strong AA forcing were known to have an increase in AMOC (Menary et al. 2013). More recently, Menary et al. (2020) showed that the multi-model mean AMOC increased significantly over 1850–1985 in historical simulations made for the sixth coupled modelling inter-comparison project (CMIP6). Furthermore, Menary et al. (2020) attributed the AMOC increase to stronger AA forcing in CMIP6 compared to CMIP5, primarily due to the inclusion of aerosol-cloud interactions in more models. However, the increase in the historically simulated AMOC in CMIP6 is in stark contrast to observation- or paleo-based proxies of the AMOC that indicate it has already declined significantly (Thornalley et al. 2018; Caesar et al. 2018, 2021). As one interpretation of this mismatch is that models do not adequately represent the sensitivity of AMOC to historical anthropogenic forcing, there is a pressing need to further understand the simulation of historical AMOC.

Unfortunately, there are few studies that have explored *how* AA affects the AMOC and, furthermore, those studies highlight different mechanisms. For example, Delworth and Dixon (2006) and Menary et al. (2013) both focus on an analysis of upper ocean density anomalies and argue AA forcing drives AMOC through changes in upper ocean salinity. However, Delworth and Dixon (2006) attribute salinity changes to decreased high-latitude precipitation, whereas Menary et al. (2013) attributes it to increased evaporation driven by atmospheric circulation changes. Furthermore, neither study explores the role of surface heat fluxes (SHF). In contrast, Hassan et al. (2021) took a surface density flux approach and attributed the multi-model mean AMOC increase in CMIP6 to changes in SHF. In particular, they argued that the AMOC strengthening was driven primarily by reduced downwelling shortwave radiation and wind-driven changes in turbulent heat fluxes. Menary et al. (2020) have also highlighted that the AMOC changes were consistent with inter-hemispheric energy constraints, but they performed no evaluation relative to other mechanisms.

All these studies also explore the mechanism in only one model (e.g., Delworth and Dixon (2006) and Menary et al. (2013)) or in the multi-model mean (e.g., Menary et al. (2020) and Hassan et al. (2021)). However, the magnitude of AA forcing is highly uncertain in terms of magnitude and both

spatial and temporal variability (Zelinka et al. 2014; Smith et al. 2021). Furthermore, due to the contemporaneous and competing external forcings it is not clear whether the dominant mechanisms are the same for all time-periods (Menary et al. 2020; Hassan et al. 2021; Kang et al. 2021). Finally, models are also known to suffer from a range of biases in surface climate and AMOC that may affect the response (Wang et al. 2014; Menary et al. 2015; Jackson et al. 2020). Therefore, a range of questions remain about how AA may have shaped the evolution of the AMOC in the real world.

The goal of this study is, therefore, to better understand how AA forcing drives the AMOC across CMIP6 models. We do this by exploring the range of simulated AMOC responses by contrasting the difference between models with strong or weak sensitivity to AA forcing. In particular, we focus on the period 1850–1985 as this is the time-period over which Menary et al. (2020) attributed AA forcing as the primary driver of externally-forced AMOC change in CMIP6. After this period, the role of AA forcing becomes less clear, likely, in part, due to the increasingly important role of GHG forcing (Menary et al. 2020) and the more complex changes in AA forcing (i.e. large regional changes but small global changes (Dittus et al. 2021; Kang et al. 2021)). As in Hassan et al. (2021), we focus on understanding the forced response in a surface density flux framework (i.e., the role of surface heat and freshwater fluxes in driving AMOC) and focus exclusively on CMIP6 simulations.

The paper is structured as follows. An overview of data and analysis methods is included in section 2. Analyses of surface density fluxes and surface heat fluxes are presented in section 3 and section 4, respectively. In section 5 we discuss the importance of continental regions for mediating the response of the AMOC to AA forcing, and section 6 discusses evidence of important feedbacks for the AMOC response. A discussion of the results, including a comparison with observations, is presented in section 7. We conclude in section 8 with a summary of the key results.

2. Data and methods

a. CMIP6 historical simulations

We explore the externally-forced changes in AMOC by using the historical simulations made for CMIP6 (Eyring et al. 2016). These simulations include time-varying changes in external forcings due to both natural (e.g., solar and volcanic aerosols) and anthropogenic sources (e.g., well mixed greenhouse gases, anthropogenic aerosol and their precursors, and land use change). For further

details of the historical simulations we direct the reader to Eyring et al. (2016). We include 17 CMIP6 models in total (see supplementary material for model details and references). These models were chosen as they include most of the variables that we analyse for at least 3 members. To ensure models have similar numbers of ensemble members, and to simplify data analysis, we also use only the first 9 members of any individual model. Overall we use 113 individual realisations (see table S1).

b. Diagnostics and observational data

The model data used in this study was retrieved from the CMIP6 archive. Before analysis, all model data was regridded to a $2.5^\circ \times 2.5^\circ$ latitude-longitude grid in order to aid comparison. As we focus on the externally-forced changes, we analyse ensemble means of individual models or multi-model means (except in the analysis presented in figure 12), and the time-series are smoothed with a 10-year running mean.

As in Menary et al. (2020), we focus on the AMOC at 35°N computed in z -coordinates, but this latitude is representative of the wider Atlantic (see Menary et al. (2020) for details). To ensure consistency across models, all AMOC data was generated directly from ocean velocities, except for NorESM2-LM, which uses potential density as the vertical co-ordinate for the ocean and, hence, CMIP6 variable *msfyz* (ocean mass stream function in Y) was used instead. As shown in Menary et al. (2020), the AMOC indices computed from ocean velocities capture the annual AMOC variability when compared to *msfyz* and the subsequent results are not sensitive to this choice (see Menary et al. (2020) for further details of the evaluation of AMOC indices).

We use the CMIP6 variable *hfds* for net surface heat fluxes into the ocean (which we refer to as SHF). Note that *hfds* was not available in some models (namely BCC-CSM2-MR, INM-CM5-0, and MIROC6). However, for these models, anomalies have been reconstructed using heat flux components from the atmospheric grid (e.g., turbulent and radiative heat fluxes). Although we do not include all terms, such as those associated with sea ice, we find that correlations of the decadal-smoothed reconstructed SHF and *hfds* averaged over the subpolar North Atlantic for models with both available are >0.95 (not shown) implying we are capturing the majority of heat flux variability. Averages in variables over regions of the ocean were computed after a land-sea mask was applied.

Surface density fluxes, and specifically the thermal and haline components, are computed here using the standard equations (e.g. following Speer and Tziperman (1992)). Due to the spread in model climatologies of sea surface temperature and salinity (not shown) we use fixed constants for the thermal and haline expansion coefficients which allows us to quickly compare the magnitude of surface density flux anomalies. Specifically, we use $\alpha = -1.4 \times 10^{-3} \text{ K}^{-1}$ and $\beta = 0.8 \times 10^{-3} \text{ ppt}^{-1}$ to compute subpolar North Atlantic surface density flux based on figure 1 from Griffies et al. (2014). Note that broad conclusions on the relative importance of thermal compared to haline components presented in section 3 are not sensitive to the exact values of α and β . The salinity and temperature components of surface density anomalies are computed by holding the other variable at a monthly varying climatology (Delworth et al. 1993). All climatologies, unless otherwise stated, are for the period 1850–1879.

We also compare some of the key variables with observations. SST data is taken from ERSST-v5 (Huang et al. 2017). Subpolar salinity data is taken from the ‘Annually binned Sea Surface Salinity’ data set of Reverdin et al. (2019). Note that to define the observed subpolar North Atlantic (SPNA) sea surface salinity (SSS) time-series we simply take the unweighted average of anomalies over all grid boxes north of 45°N in the observational salinity dataset, due to the use of an irregular grid in this data set (This is not the case for other data sets, where we use a average over the regions specified weighted by grid area). Surface air temperature data is taken from Berkeley Earth surface temperature data set (BEST, Rohde and Hausfather (2020)).

c. Stratification of models into sub-ensembles based on inter-hemispheric energy imbalance

Menary et al. (2020) showed that the strength of the externally-forced AMOC change over 1850–1985 in CMIP5 and CMIP6 historical simulations was largely proportional to the inter-hemispheric difference in the absorbed (or net) solar radiation at the top of the atmosphere (which they called ASR_HD). In particular, they defined the ASR_HD index as the Southern Hemisphere mean net shortwave *minus* the Northern Hemisphere mean net shortwave, so that *positive* values indicate that *less* solar radiation is absorbed in the Northern Hemisphere (as incoming radiation is the same in the two hemispheres). Furthermore, they showed that an increase in the ASR_HD index over 1850–1985 was associated with an increase in AMOC, whereas a decrease over 1985–2014 was associated with a weakening AMOC. Using single forcing simulations made for the Detection and

Attribution Model Intercomparison project, (DAMIP, Gillett et al. (2016)), Menary et al. (2020) also showed that for ~1850–1985, the changes in the ASR_{HD} index were primarily controlled by changes in the AA forcing (see their figure 2). However, after 1985 the relationship became more complicated with both AA and GHG forcing contributing.

We take advantage of the relationship between ASR_{HD} and AMOC changes in order to stratify models into sub-ensembles which have ‘strong’ and ‘weak’ responses to aerosol forcing over the period 1850–1985. Specifically, we split models into those that simulate a linear change in ASR_{HD} over 1850–1985 that is larger than 1.5 Wm^{-2} based on a linear regression (which we name ‘strong’) and those which have a smaller change (which we name ‘weak’). We do this by fitting a linear trend to the ensemble-mean time-series and multiplying by the number of years between 1850–1985 (i.e. 136). Note that 1.5 Wm^{-2} is an arbitrary choice in order to have comparable sized sub-ensembles (9 models and 55 members for ‘strong’ and 8 models and 58 members for ‘weak’, respectively) but the results are not sensitive to this exact threshold.

Figure 1 shows the resulting time series of ASR_{HD} and the AMOC at 35°N as well as the relationship in their linear trends over the period 1850–1985. For the multi-model mean (MMM) there is an increase in both ASR_{HD} and the AMOC from 1850–1985 with the fastest increase over ~1940–1985 (see fig 1a and b, respectively). Focusing on the 1965–1985 time period, which corresponds to peak values of both ASR_{HD} and AMOC (and also to peak sulphur dioxide emissions from the US and Europe in CMIP6 (Sutton et al. 2018; Dittus et al. 2021)), we find MMM time-mean anomalies are 1.4 Wm^{-2} and 1.0 Sv, respectively. However, when considering the ‘strong’ and ‘weak’ sub-ensembles separately we find a large spread in the overall response. For example, ‘strong’ models have a larger increase in both ASR_{HD} and AMOC, with 1965–1985-mean anomalies of 2.1 Wm^{-2} and 1.7 Sv, respectively, whereas ‘weak’ models have small, but still statistically significant, anomalies of 0.5 Wm^{-2} and 0.2 Sv, respectively. Consistent with Menary et al. (2020), we also find that this relationship applies generally across each model with individual model ensemble-mean trends of ASR_{HD} and AMOC correlated with a value of 0.92 (see fig. 1c).

Although the peak values in ASR_{HD} and AMOC occur in the late 20th Century, we also note that both ASR_{HD} and AMOC indices increase in the later 19th Century in the MMM and ‘strong’ and ‘weak’ sub-ensembles giving rise to a smaller peak in AMOC at ~1920. This earlier increase

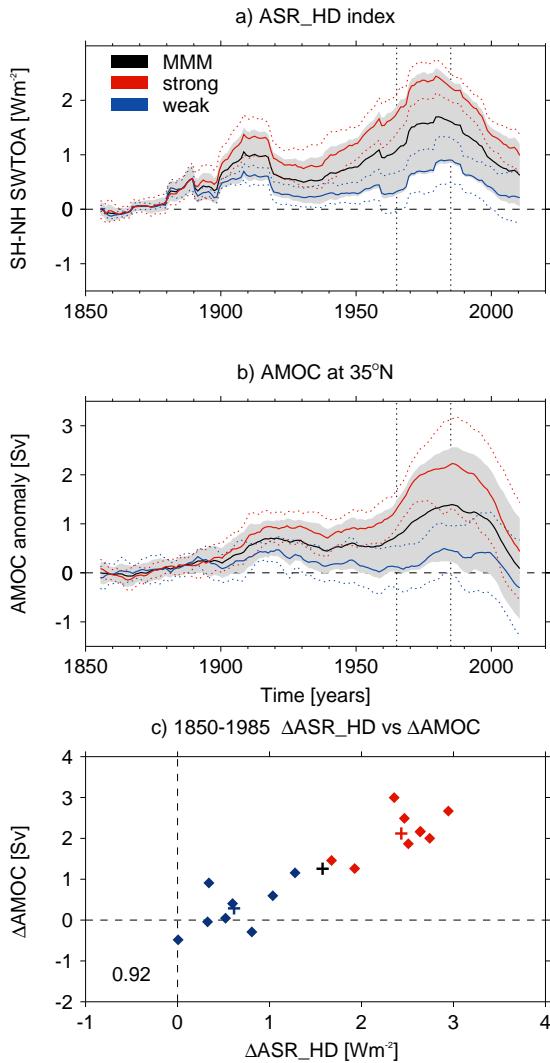


FIG. 1. Simulated inter-hemispheric imbalance in absorbed solar radiation (ASR_HD) and AMOC. a) shows anomalies of ASR_HD index for the multi-model mean (MMM, black), and for the sub-ensembles for ‘strong’ and ‘weak’ aerosol forcing (red and blue, respectively). Thick lines show the ensemble mean and grey shading or dotted lines indicate the 1σ spread. Dashed vertical lines indicate the 1965–1985 time-period. Anomalies are relative to an 1850–1879 climatology. b) shows the same but for the AMOC at 35°N. c) diamonds show a scatter plot of changes in ASR_HD vs AMOC over 1850–1985 for each models ensemble mean. Changes are computed by fitting a linear trend to individual model ensemble-mean time series to compute change per year and then multiplying by the number of years (e.g. 136). Red or blue indicate ‘strong’ and ‘weak’ models. Crosses indicate MMM (black) and ‘strong’ (red) and ‘weak’ (blue) sub-ensemble means, respectively.

is likely to be, in part, related to changes in inter-hemispheric volcanic forcing (as shown by the spikes in ASR_{HD} index in Menary et al. (2020)). However, it is also consistent with the increase in sulphur dioxide emissions over North America and Europe in CMIP6 (see Sutton et al. (2018) figure 1c) and previous work highlighting the role of AA forcing at this time (Wilcox et al. 2013). As both ASR_{HD} and AMOC indices appear to diverge between ‘strong’ and ‘weak’ models in ~1900, we infer that AA forcing plays an important role in driving AMOC over the whole historical period.

3. Surface density drivers of AMOC

We now turn our attention to changes in surface density flux over the North Atlantic and explore its relationship to AMOC.

Figure 2 shows that for the 1965–1985 time-mean (i.e., immediately prior to the peak AMOC anomalies) there are significant changes in MMM annual-mean surface heat fluxes (SHF) and Precipitation minus Evaporation (PmE) fluxes (see fig. 2a and e). Note that we focus on PmE due to its importance in previous studies (e.g., Delworth and Dixon (2006) and Menary et al. (2013)), and as not all freshwater flux terms were available in all models (not shown). Increased heat loss (negative anomalies) is simulated across much of the higher latitude North Atlantic, with significant anomalies along the Gulf Stream Extension, in the subpolar North Atlantic (SPNA). Significant changes in PmE are also seen with a reduction over the subpolar North Atlantic and Tropical North Atlantic.

However, there are substantial differences between the ‘strong’ and ‘weak’ models, especially over the SPNA (compare fig. 2b and f with fig. 2c and g). The difference between the ‘strong’ and ‘weak’ ensembles (i.e., ‘strong’ *minus* ‘weak’ and henceforth the ‘the difference’) are shown in fig. 2d and fig. 2h for heat and PmE fluxes, respectively. The comparison shows that the SHF anomalies in the MMM are dominated by the ‘strong’ models. In contrast, there is relatively little change in SHF in the ‘weak’ models over much of the North Atlantic, including over the SPNA. For PmE, both ‘strong’ and ‘weak’ models show significant changes over many regions of the North Atlantic. In particular, there is a reduction in PmE (negative anomalies) in the subtropics in both ‘strong’ and ‘weak’ models which is associated with increased evaporation (not shown).

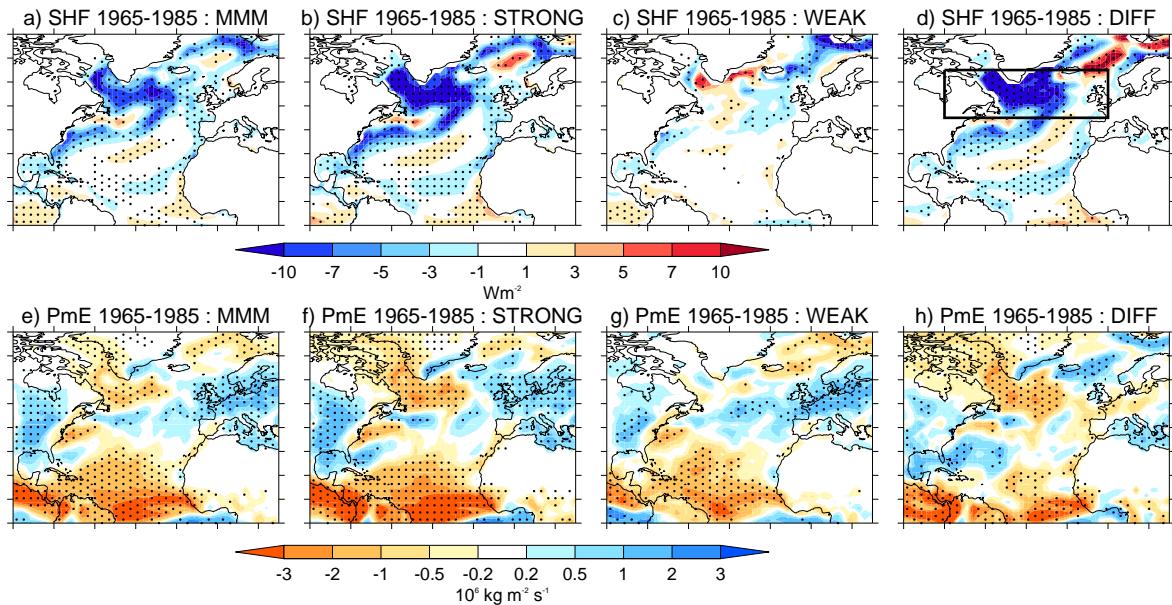


FIG. 2. Comparison of annual-mean surface fluxes anomalies in CMIP6 Historical simulations over the period 1965–1985 relative to 1850–1979. a) shows the MMM anomalies of total surface heat flux (SHF, negative anomalies is cooling). b) and c) show the same but for the ‘strong’ and ‘weak’ models, respectively. d) shows the difference between ‘strong’ and ‘weak’ ensembles (i.e. ‘strong’ – ‘weak’). e)–h) shows the same, but now for precipitation minus evaporation (PmE, negative is reduced freshwater). For panels a)–c) and e)–g) stippling shows where the anomalies are significantly different to climatology at the $p \leq 0.05$ level based on a Student’s t-test. For panels d) and h) (i.e., the difference) stippling shows where the anomalies of ‘strong’ and ‘weak’ are significantly different to each other at the $p \leq 0.05$ level based on a Student’s t-test. Black box on panel d) indicates the region defined as the SPNA ($45\text{--}65^\circ\text{N}$, $80\text{--}0^\circ\text{W}$).

Nevertheless, the difference in PmE anomalies over the SPNA are large, with a significant reduction in PmE only seen in the ‘strong’ models.

As the differences in surface fluxes are dominated by changes over the SPNA, and as we expect changes in surface fluxes over the SPNA to drive AMOC (Delworth and Zeng 2016; Xu et al. 2019), we focus on this region for the remainder of this paper. Figure 3 summarises the time series of changes in SHF (fig. 3a) and PmE (fig 3c) averaged over the SPNA region. Fig 3a shows negative (i.e. increased heat loss) from ~ 1900 , before a larger decrease from $\sim 1940\text{--}\sim 1985$. However, the MMM SHFs are dominated by changes in the ‘strong’ models over this time period. In contrast, there is a small initial cooling in ‘weak’ models in ~ 1900 , but the SHF anomalies

diverge between the ‘strong’ and ‘weak’ models after this point. PmE also decreases (largely due to increased evaporation) in the MMM after ~1900. However, in contrast to surface heat fluxes, there is little divergence between ‘strong’ and ‘weak’ models until after ~1940. Note that the temporal correspondence between SHF and PmE, especially after 1950, suggests that anomalous evaporative cooling is contributing to the SHF anomalies (compare fig. 3a and c). We will address the overall contributions to SHF in section 4.

Converting the SPNA surface fluxes into surface density flux anomalies shows that SHF dominates the overall surface density flux anomalies, consistent with Hassan et al. (2021) for the MMM (note difference in y-axis values between fig. 3b and fig. 3d). The evolution of the thermal surface density flux is also consistent with the evolution of the AMOC over the historical period for the MMM and ‘strong’ and ‘weak’ sub-ensembles with the surface heat flux leading AMOC at 35°N by a few years (compare with figure 1b). Indeed, similar to Hassan et al. (2021), we find a correlation of 0.80 when thermal surface density flux leads AMOC by 4 years (not shown). However, the correlation is 0.92 for the ‘strong’ models with a lead of 2 years (not shown). Therefore, we conclude that the historical MMM AMOC anomalies in CMIP6 are consistent with the evolution of SPNA surface heat fluxes, and their impact of surface density fluxes, driving the AMOC in the ‘strong’ models.

4. Decomposing the subpolar North Atlantic surface heat fluxes

The analysis in the previous section showed that SPNA SHF anomalies appear to dominate the evolution of AMOC in the historical simulations, and the differences between ‘strong’ and ‘weak’ models. Therefore, in this section we break down the SHF to understand the processes at play.

Figure 4 shows the breakdown of the total annual-mean SHF into the parts associated with turbulent (turHF), surface net shortwave (sNetSW), and surface net longwave (sNetLW) on the top row. When using the MMM, we find that the changes in the sNetSW dominate changes in SHF; the sNetSW is -3.2 W m^{-2} for the 1965–1985 mean, and accounts for the majority of the -4.2 W m^{-2} total SHF. In contrast, turHF only accounts for -0.9 W m^{-2} , and sNetLW changes account for a small increase. Hence, the analysis of the MMM is consistent with the hypothesis that AA forcing mainly drives the AMOC through its direct impact on the SPNA SHF through changes in net solar, as discussed in Hassan et al. (2021).

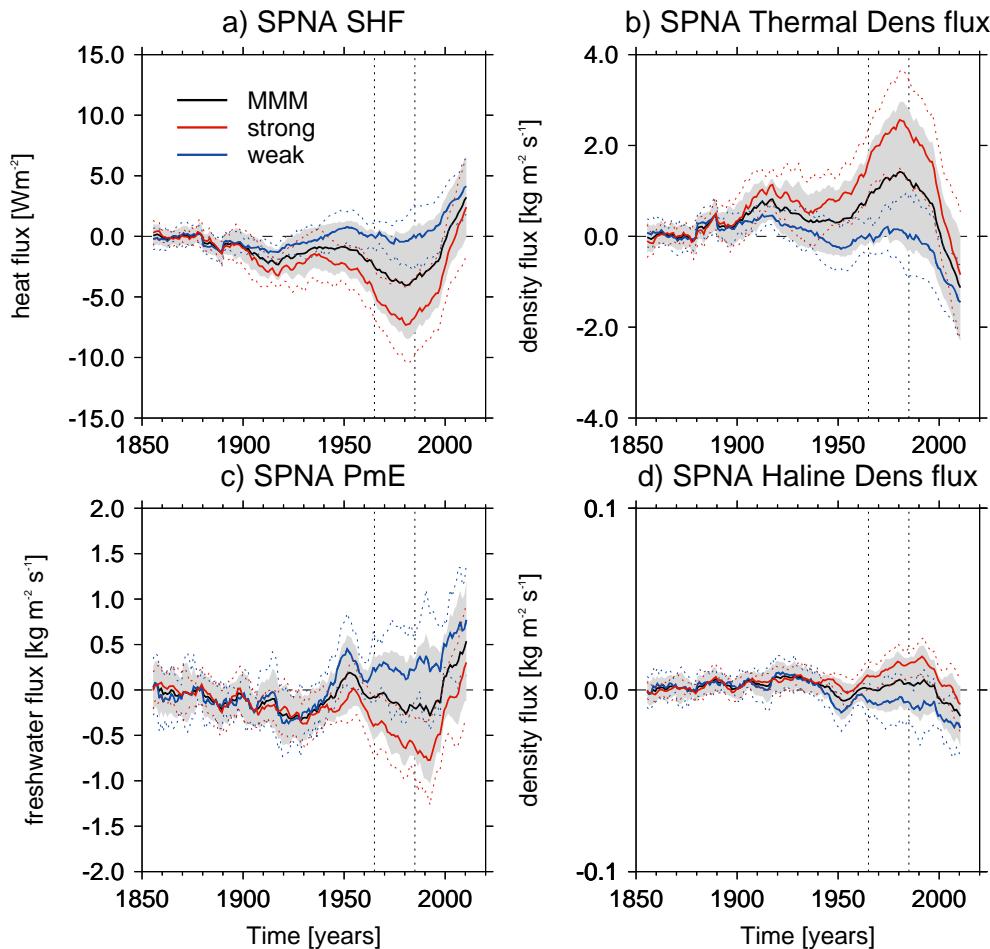


FIG. 3. Comparisons of annual-mean surface heat and freshwater fluxes and equivalent surface density fluxes averaged over the SPNA (45-65°N, 80-0°W). a) shows the total surface heat flux (SHF) and b) shows the implied density flux due to changes in heat fluxes (in $10^{-6} \text{ kg m}^{-2} \text{ s}^{-1}$) - see section 2b for details. c) and d) shows the same but for the precipitation minus evaporation (PmE). Thick lines show the multi-model mean (MMM, black) and ‘strong’ (red) and ‘weak’ (blue) sub-ensembles, respectively. The 1σ spread of the means is shown in grey shading for the MMM, and with thin dashed red and blue lines for the strong and weak models, respectively. Dashed vertical lines indicate the 1965–1985 time-period. All time-series are smoothed with a 10-year running mean.

However, figure 4 shows that the divergence in SHF between ‘strong’ and ‘weak’ ensembles is *not* driven by changes in sNetSW. Instead, changes in turHF dominate the spread. In terms of the sub-ensemble means, we find ‘strong’ models have turHF cooling of -2.5 W m^{-2} by the 1965–1985

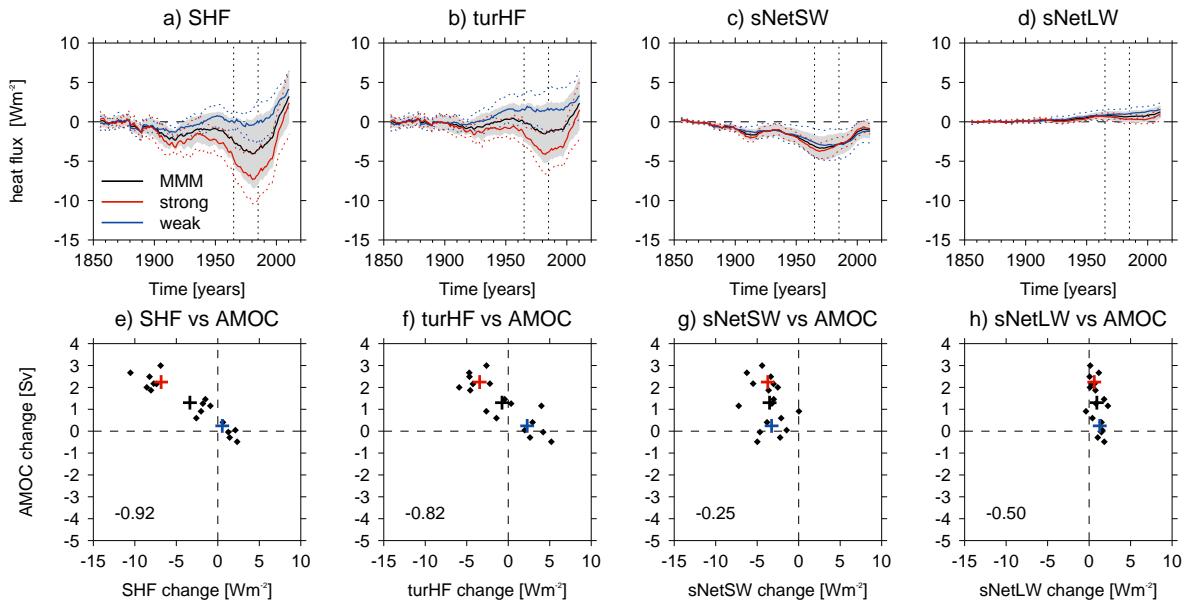


FIG. 4. Time series and trends of annual-mean surface flux anomalies averaged over the SPNA. a) shows the net surface heat flux anomalies (SHF, W m^{-2}) for the multi-model mean (black), and the ‘strong’ and ‘weak’ sub-ensembles (red and blue, respectively). Gray shading and coloured dotted lines shows 1σ ensemble spread. Vertical dotted lines highlight 1965–1985 - note this is the same index shown in fig. 3a. b) - d) show the same, but for the turbulent heat fluxes (turHF), the net surface shortwave (sNetSW) and the net surface longwave (sNetLW). All anomalies are made relative to the 1850–1879 mean. f) shows a scatter plot of linear trends in total heat flux (SHF) compared to trends in AMOC. All trends are computed over 1850–1985, and value shows the total change over that period computed from the linear fit. All time series are smoothed with a 10-year running mean.

mean, whereas ‘weak’ models have a net warming of 1.5 W m^{-2} (i.e. a difference of -4 W m^{-2}). In contrast, the difference between the sNetSW and sNetLW between ‘strong’ and ‘weak’ models for the 1965–1985 mean is -0.7 W m^{-2} and -0.4 W m^{-2} , respectively. The importance of turHF is also evident when we explore linear trends for individual models; figure 4 bottom row shows that there is a strong relationship between trends in SHF and AMOC in individual models over 1850–1985 due to the trends in turHF (correlation of -0.92 and -0.82 , respectively). In contrast, the relationship between sNetSW and AMOC is substantially weaker at -0.25 . Note, we also find a correlation of -0.50 between sNetLW and AMOC, but the magnitudes of the sNetLW anomalies are small in comparison. Although we focus on annual-mean SHFs for brevity, we also note that differences in SPNA turHF anomalies are largely dominated by increased winter heat loss due to

both sensible and latent heat fluxes, although summer does also play a significant role (see figure S1 in supplementary materials for the seasonal heat flux breakdown).

Alongside the changes in surface heat fluxes, we also find changes in atmospheric circulation. Indeed, Hassan et al. (2021) attributed changes in MMM turbulent heat flux anomalies to changes in atmospheric circulation and, specifically, surface wind speed. These changes in annual-mean atmospheric circulation and surface wind (sfcWind) for the MMM ensemble are summarised in figure 5a -5h for the 1965–1985 time-period. Note that seasonal changes in atmospheric circulation in summer and winter are similar (not shown). In all seasons, we see an increase in the meridional pressure gradient in the North Atlantic region and, hence, a strengthening of the westerly winds. Again we find that the magnitude of the atmospheric response over the North Atlantic is dominated by the ‘strong’ models. Therefore, one interpretation is that the larger atmospheric circulation response in ‘strong’ models is driving the larger response of the AMOC.

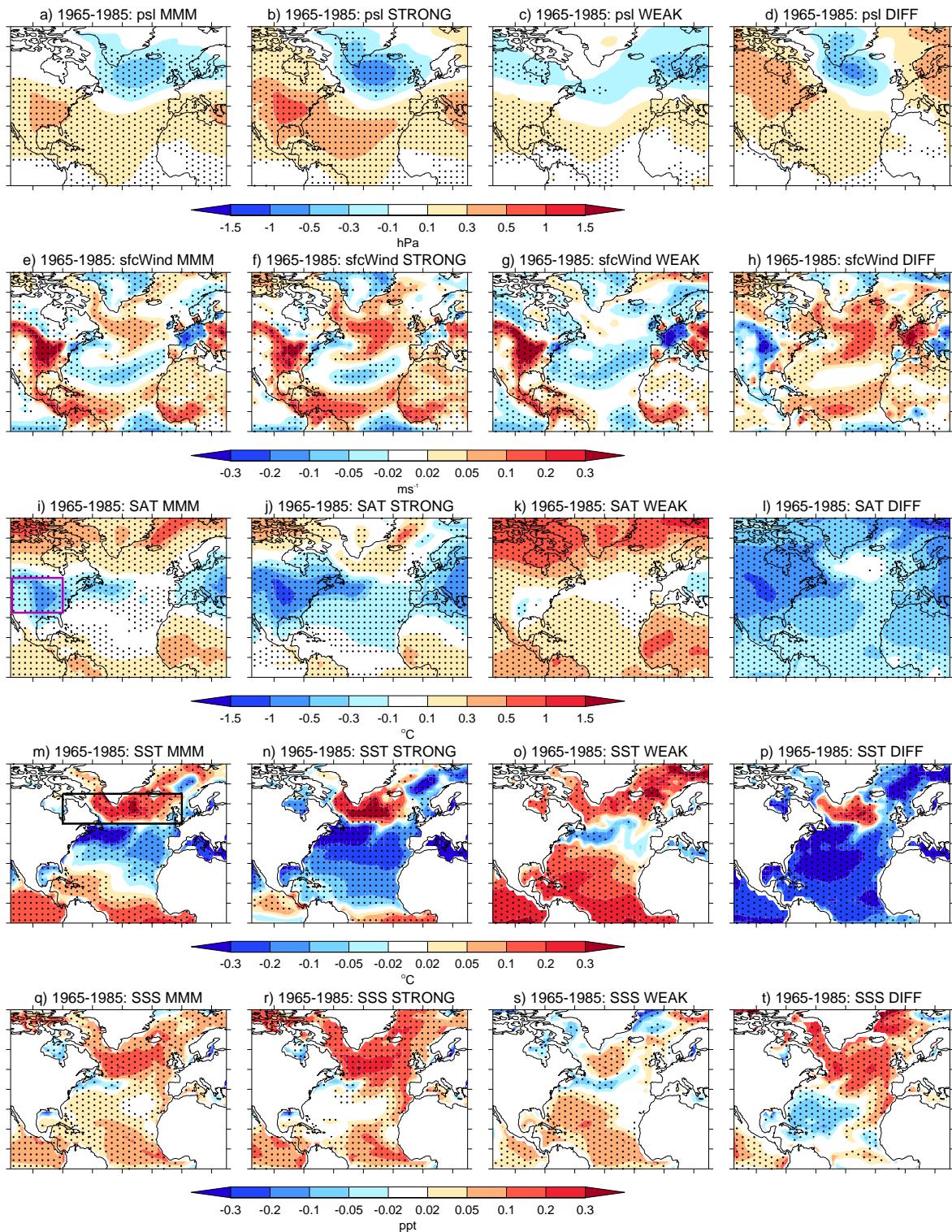


FIG. 5. same as figure 2 but now for surface pressure (CMIP6 variable psl , a-d), near surface wind speed (CMIP6 variable $sfcWind$, e-f), surface air temperature (SAT (CMIP6 variable tas), i-j) sea surface temperature (SST (CMIP6 variable tos), m-p) and sea surface salinity (SSS (CMIP6 variable sos), q-t).

However, although they are significant, it is important to note that the changes in atmospheric circulation reported here (and in Hassan et al. (2021)) are relatively small compared to the total variance. For example, the difference in the total change in North Atlantic pressure gradients over 1850–1985 (computed by scaling the linear fit by the total number of years) is only $\sim 0.5\text{hPa}$ between ‘strong’ and ‘weak’ for winter means (i.e., where turbulent heat flux anomalies dominate, not shown). Furthermore, turHFs are not just driven by changes in wind speed; contrasts in temperature and humidity at the air-sea interface (i.e. the ‘thermodynamic’ terms) are key drivers of turHF alongside wind speed (see Large and Yeager (2009)). We denote these terms ΔT and ΔH , where ΔT is computed as the difference in surface air temperature and sea surface temperature (e.g., SAT *minus* SST) and ΔH is computed as the difference in surface specific humidity (e.g., CMIP6 variable *huss*, but henceforth *huss*) and saturated specific humidity computed from SST (e.g., *huss minus satH*), respectively. Exploring these terms further, we also find that there are concurrent differences in anomalies of ΔT and ΔH that are consistent with the differences in turHF anomalies seen between ‘strong’ and ‘weak’ models (not shown, but see figure S1 in supplementary material).

To explore the drivers further, we perform a linear decomposition of the turHF and find that the thermodynamic changes are driving both the anomalous turbulent heat fluxes and the difference between ‘strong’ and ‘weak’ models. Figure 6 shows a simplified linear decomposition of the annual-mean turHF. Specifically, it shows the multiplication of the air-sea contrasts (e.g., ΔT and ΔH) by *sfcWind* in order to generate ΔT *sfcWind* and the ΔH *sfcWind* terms and the subsequent linear decomposition terms. In other words, we compute $AB = A'[B] + [A]B' + A'B'$, where A is either ΔT or ΔH , B is *sfcWind*, ' indicates deviation from the climatology, and [] indicates the 1850–1879 climatological mean. These terms are computed for each grid-point from monthly-mean data before making the annual and spatial averages. The turbulent heat fluxes are proportional to these terms (e.g., see Large and Yeager (2009)), and we find that the SPNA-mean time-series correlate highly with the sensible and latent heat flux terms on decadal time-scales (see numbers in fig. 6a and d). Figure 6 shows that the evolution of these terms are dominated by the anomalies in the ‘thermodynamic’ air-sea contrasts (e.g., $\Delta T'$ [*sfcWind*] and $\Delta H'$ [*sfcWind*]; see fig. 6b and e). In contrast, the role of the anomalous wind anomalies (e.g., [ΔT] *sfcWind'* and [ΔH] *sfcWind'*; see fig. 6c and f) is small throughout most of the time period. Note that we do not show the covariance terms (e.g., $A'B'$) as they are small throughout.

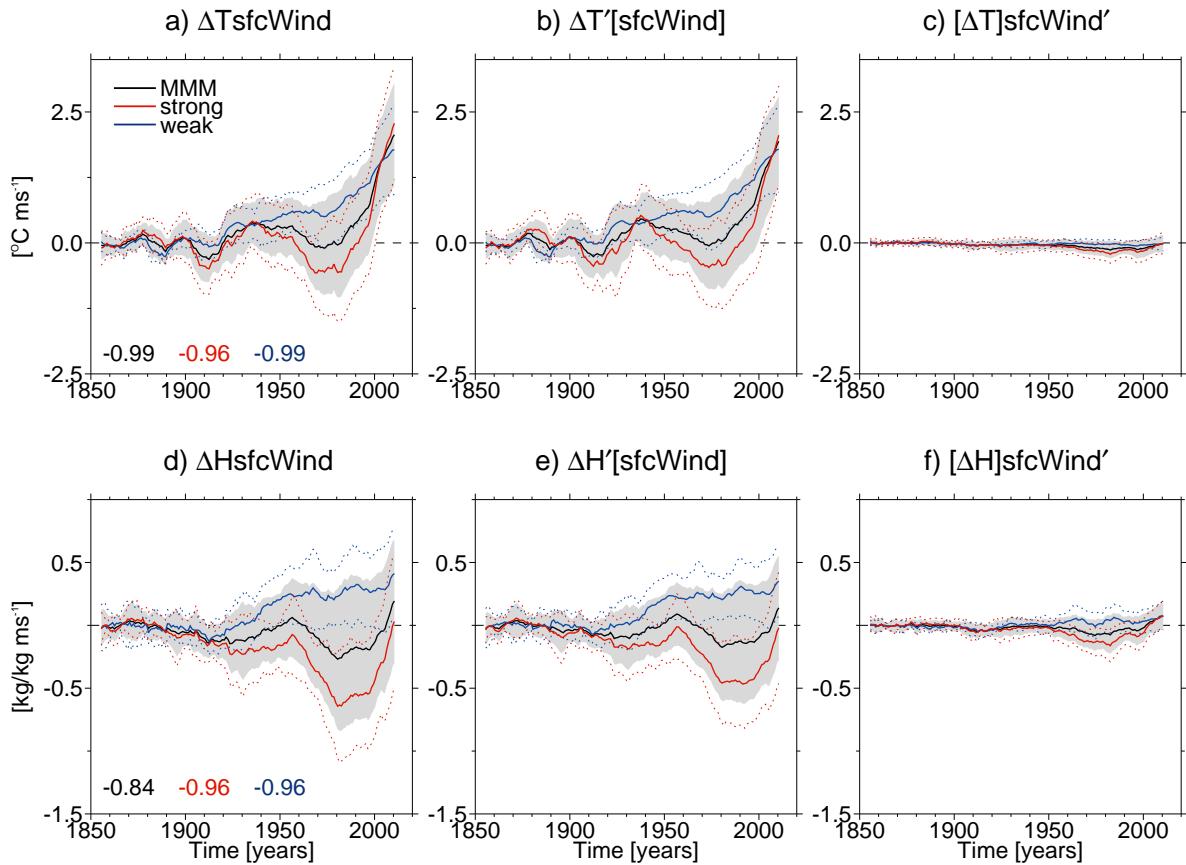


FIG. 6. Shows the linear decomposition of the turbulent heat flux drivers. a) shows anomalies in the air-sea temperature contrast multiplied by surface windspeed anomalies (i.e., $\Delta T_{\text{sfcWind}}$). Numbers inset show the correlation of this term with the sensible heat flux (hfss). b) and c) shows anomalies due to air-sea temperature contrasts ($\Delta T'[\text{sfcWind}]$) and anomalous wind speed ($[\Delta T]\text{sfcWind}'$) respectively. Here the ' indicates anomalies relative to the time-mean climatology and square-brackets indicate the time-mean climatology. All climatologies are relative to 1850–1879 and time-series are smoothed with a 10-year running mean. d) to f) shows the same, but for the ΔH terms and correlations are with the latent heat terms (hfls).

5. Continental origin of thermodynamic anomalies?

In the previous section we concluded that increased turbulent heat fluxes in the 'strong' models arises primarily due to increased thermodynamic contrast between the atmosphere and the SPNA. In other words, relative to the sea surface, the atmosphere is colder and drier in models with stronger AA forcing. Figure 7 shows that this cooling and drying of the atmosphere over the SPNA is consistent with the broad impact of external forcing on the inter-hemispheric difference

in surface air temperature and specific humidity (SAT_HD and huss_HD respectively; note the HD indicates the hemispheric difference, which is the Southern Hemisphere-mean minus the Northern Hemisphere-mean to be consistent with the ASR_HD index). The evolution of the SAT_HD and huss_HD time-series are consistent with the evolution of ASR_HD index and the AMOC in the MMM, with both indices peaking in the 1965–1985 time period (i.e., see fig. 1). However, as with the ASR_HD index and the AMOC, MMM anomalies in both SAT_HD and huss_HD time-series are dominated by the ‘strong’ models and, in contrast, there is relatively little change in the ‘weak’ ensemble-mean post ~1920. Therefore, the AMOC response appears broadly consistent with the inter-hemispheric imprint of AA forcing on surface temperature and humidity.

Nonetheless, it is well known that AA forcing is heterogeneous in both time and space (Wang et al. 2015; Lund et al. 2019). Consequently, there are significant regional differences in the pattern of surface forced changes between the ‘strong’ and ‘weak’ models. For example, figure 5i) shows that the largest SAT anomalies in the broader North Atlantic region are located over the North American continent for the 1965–1985 mean. Such a cooling of the continental North American region, compared to the ocean, is consistent with the larger changes in net shortwave (see figure S2 in supplementary material for global comparison of shortwave and SAT anomalies) as well as the lower heat capacity of the land and other processes that lead to land-sea contrasts in temperatures (Sutton et al. 2007; Joshi et al. 2008).

Focusing on the average of the North American continental region (which we call SAT_NAm), figure 7c shows that SAT_NAm cools in the MMM by approximately $-0.44\text{ }^{\circ}\text{C}$ by 1965–1985. However, this cooling is, again, dominated by the ‘strong’ models which have a 1965–1985 anomaly of $-0.89\text{ }^{\circ}\text{C}$, whereas ‘weak’ models have a small warming of $0.08\text{ }^{\circ}\text{C}$. This relationship is also clear for individual models, with a correlation of -0.92 between 1850–1985 trends in ASR_HD and SAT_NAm (not shown). The cooling is also associated with a drying of the atmosphere in the ‘strong’ relative to ‘weak’ models (i.e., see figure 7d). Note, the largest cooling of the ‘strong’ models occurs in ~1970, and is broadly consistent with the peak in North American sulphur dioxide emissions that occurred in the 1960s–1970s (Dittus et al. 2021).

Therefore, we hypothesise that the regional cooling and drying over the North American continent (which is upwind of the North Atlantic) played an important role in shaping the subsequent evolution of the externally-forced AMOC through its impact on the air-sea contrast in temperatures

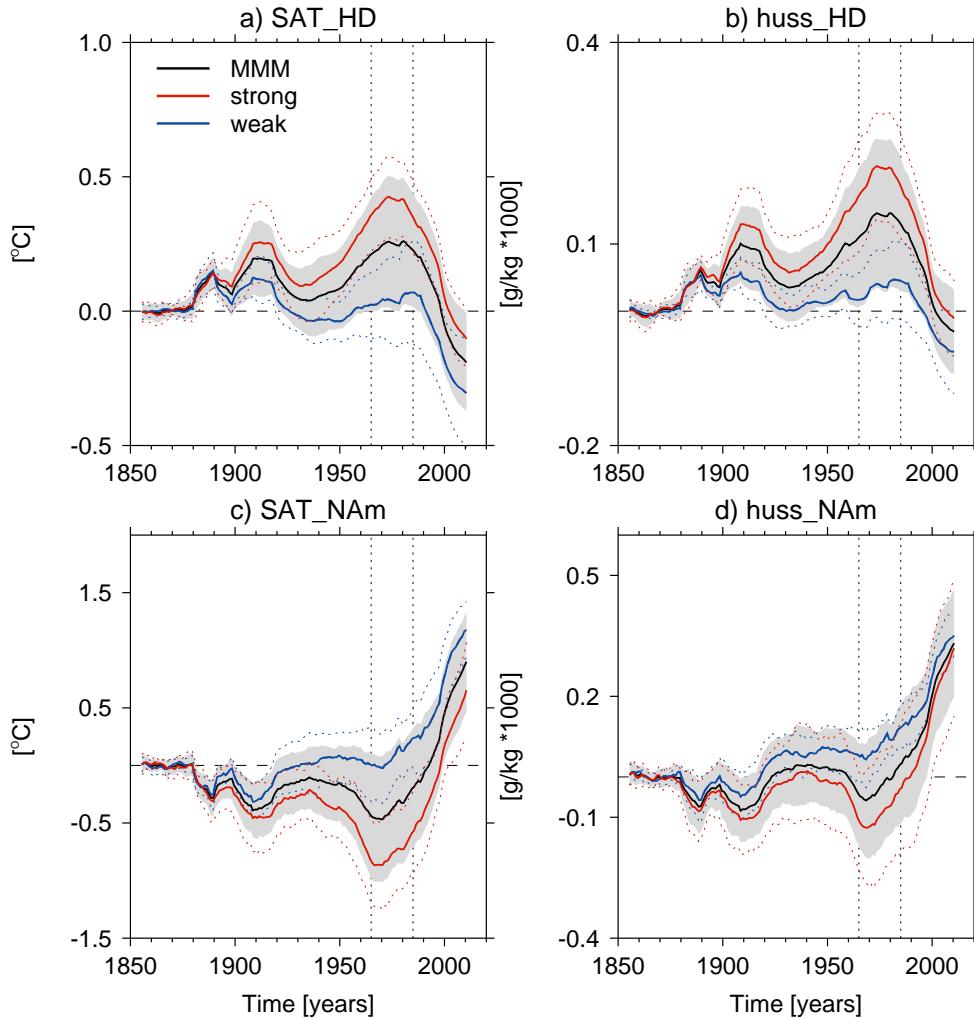


FIG. 7. shows key time-series of surface temperature and surface humidity. a) and b) shows the inter-hemispheric difference in surface air temperature and surface specific humidity (SAT_HD and huss_HD, respectively). Both are computed as SH–NH. c) and d) shows surface air temperature and surface specific humidity averaged over the North American continent (115–80°W, 32.5–50°N, see purple box on fig. 5i, SAT_NAm and huss_NAm respectively). Solid thick lines show the ensemble means for the MMM (black), ‘strong’ (red) and ‘weak’ (blue). 1σ spread of the MMM is shown in grey shading, and thin red and blue dashed lines for the strong and weak respectively. Vertical dashed lines show the period 1965–1985 period. All time-series are smoothed with an 10-year running mean.

and humidity. In other words, the cool and dry continental anomalies were advected over the North Atlantic, and this drove increased turbulent heat flux cooling of the SPNA. Such a hypothesis

is broadly supported by the analysis in figure 8, which shows the longitude-time of annual-mean differences (i.e., ‘strong’ *minus* ‘weak’) turbulent heat flux anomalies and their drivers. In particular, it is clear from figure 8 that the cooling and drying of the atmosphere over the North American continent is substantially larger than the cooling and drying over the SPNA (see fig. 8c and g). Furthermore, the continental cooling leads the atmospheric cooling over the SPNA, which is especially clear when comparing the ΔT and ΔH terms (fig 8e and i) to SAT and huss (fig 8c and g).

Finally, figure 8a shows that the anomalous heat fluxes are first seen in the western SPNA, especially in the sensible heat fluxes (*hfss*, fig. 8b), and then in the East. This feature is most evident for anomalies in the period ~1900–1930, but occurs again from ~1940. The initial anomalies in the west could be consistent with cold advection from the North American Continent and the resulting ΔT in particular (i.e., the effect will be largest nearest the source of cold air as the atmospheric anomalies will come into balance with the SST as they are advected over the ocean). The anomalous heat fluxes occurring later in the Eastern SPNA after 1940 may be consistent with the intensification of the cold dry signal over North America (i.e. the colder and drier lower atmosphere anomalies advected over the SPNA take longer to come into equilibrium with the SSTs). However, it may also be consistent with the expected warming at this time due to the strengthened AMOC which can especially impact the Eastern SPNA (e.g., Moat et al. (2019)).

6. Oceanic feedbacks that act to amplify the AMOC response to anthropogenic aerosols

So far we have explored the forced changes in the AMOC from a surface density flux framework. However, we expect there to be feedbacks which amplify or damp the forced response on the AMOC. For example, through its role in ocean heat transports, an increase in the AMOC is usually associated with a warming of SST (Moat et al. 2019), which will affect the air-sea temperature contrasts and, hence, surface fluxes. Furthermore, transports of heat and salinity/freshwater could impact on ocean density (e.g., Menary et al. (2013)), which could alter dense water formation (Speer and Tziperman 1992; Petit et al. 2021).

We find that changes in SPNA SST that are due to forced changes in AMOC act as a *positive* feedback on turbulent heat fluxes. Figure 5m shows that for the 1965–1985 time-period the SPNA is anomalously warm in the MMM compared to the 1850–1879 climatology despite the large and

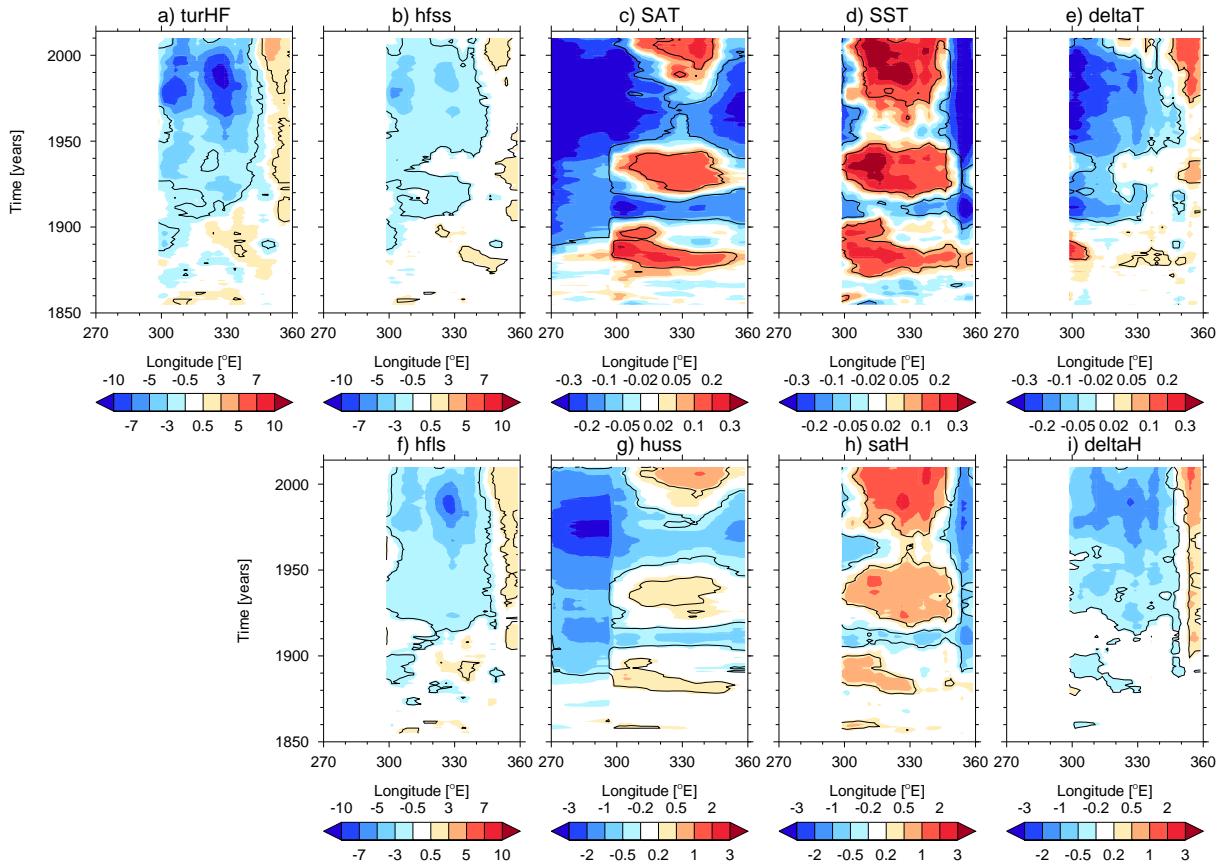


FIG. 8. Longitude-time plots of differences (i.e., ‘strong’–‘weak’) in surface heat fluxes and their drivers averaged over the SPNA (50–65°N) between 60–0°W and atmospheric anomalies over the North American Continent (115–80°W, 32.5–50°N (i.e., the box on fig. 5i). a) shows anomalies in turbulent heat flux (turHF) [W m^{-2}], where negative anomalies is increased cooling of the Ocean. b), c), d) and e) show the same as a) but now for sensible heat flux (hfss), surface air temperature (SAT), sea surface temperature (SST) and deltaT (e.g. SAT *minus* SST), respectively. f)-i) shows the same as b)-e) but now for latent heat flux (hfls), specific humidity (huss), surface saturated humidity (satH), and ΔH (e.g. huss *minus* satH). satH is computed as in equation 4 from Large and Yeager (2009). Specifically, $\text{satH} = 0.98\rho^{-1}640380(\text{kg m}^{-3})e^{-5107.4/SST}$ where ρ is the density of air and assumed to be 1.22 kg m^{-3} (Large and Yeager 2009). Note that for SAT and huss, values west of 60°W are averaged over North American Continent. Black contours show where the differences are significant at the $p \leq 0.05$ level. Bottom shows the same, but for the humidity and latent heat flux terms. Note that IPSL-CM6 and NESM3 have been removed from panels f)-i) due to problems with humidity data.

consistent cooling of the SPNA (e.g., fig. 4a), which indicates the surface warming originates from the ocean rather than the atmosphere. We also find that the SPNA is warmer in ‘strong’ models than in ‘weak’ (see fig. 5p). These changes in SST are consistent with the simulated relationship between SPNA SST and AMOC changes (Menary et al. 2020). As turbulent heat fluxes are proportional to the air-sea temperature and humidity contrasts, and as the warming SST will increase the contrast between the atmosphere and ocean, the simulated warming of the SPNA SST will further enhance the turbulent heat flux cooling in the ‘strong’ models.

By exploring the role of atmosphere and ocean in driving the air-sea contrasts in temperature and humidity (e.g. ΔT and ΔH), we find that generally the atmosphere is the main driver of the TurHF anomalies. Figure 9 shows the role of the ocean (e.g., SST) and atmosphere (e.g., SAT and huss) in shaping ΔT and ΔH over the SPNA by showing the difference in these terms (i.e., ‘strong’ minus ‘weak’). It is clear that the atmosphere plays an important role in ΔT and ΔH ; in particular, negative SAT anomalies contribute to ΔT between ~ 1900 – 1930 and ~ 1940 – 2000 (fig. 9a), and negative huss anomalies contribute to ΔH after ~ 1940 – 1990 (fig. 9b). Furthermore, by comparing the tendency of the smoothed SAT and SST terms, we find periods consistent with the atmosphere driving the ocean between ~ 1900 – 1915 and ~ 1940 – 1970 (i.e., the atmosphere is cooling faster, see purple horizontal lines). Therefore, figure 9 is consistent with the hypothesis that the SHFs are driving AMOC, and not vice versa.

However, although the SAT anomalies play a pivotal role in driving air-sea contrasts over the SPNA, we also find that SST changes do indeed contribute to broaden and delay the peak in turbulent heat flux cooling. For example, the minimum of ΔT and ΔH occurs in ~ 1985 , consistent with the peak heat loss in the ‘strong’ models (e.g. fig. 4). However, the minimum in SPNA SAT and huss occurs earlier, in ~ 1970 . Note this earlier minimum in atmospheric variables is consistent with the minimum SAT and huss averaged over the North American Continent in the ‘strong’ models (see fig 7c and d) and is, hence, consistent with our hypothesis about the importance of AA forced continental temperature anomalies and their resultant advection over the North Atlantic. Therefore, SSTs also contribute to the evolution of ΔT and ΔH , and appear important between ~ 1920 – 1940 and, especially, after ~ 1960 where SSTs become anomalously warm in the ‘strong’ models. This warming of SSTs after ~ 1960 is particularly clear in winter, which is the season with the strongest turbulent heat flux cooling (see figure S1 in supplementary material).

Therefore, the changes in SST in the ‘strong’ models are clearly acting to increase the differences in surface heat fluxes between ‘strong’ and ‘weak’ and to broaden the peak heat flux cooling especially after ~1980.

The second feedback involves changes in salinity that *positively* feedback onto surface ocean density anomalies. Figure 5q shows that for the 1965–1985 time-period in the MMM the SPNA is also anomalously salty compared to the 1850–1879 climatology. As with other variables, the salinification of the SPNA is dominated by the ‘strong’ models (fig. 5r), although there is a small salinification of the ‘weak’ models (fig. 5s). When considering the difference between ‘strong’ and ‘weak’ models we find that there is a salinification of the SPNA, but a relative freshening of the western subtropical Atlantic (fig. 5t). These changes in salinity are, again, consistent with the changes in AMOC and its impact on freshwater transports, with increased AMOC leading to increased northward transport of salt (Zhu and Liu 2020).

We find that the changes in SPNA salinity are a key driver of SPNA surface density anomalies. Figure 10 shows surface density anomalies averaged over the SPNA are also very similar to the evolution of the AMOC (see fig 1). The MMM index shows an increase in surface density from ~1880 to ~1910 before decreasing slightly and leveling off until ~1960 and increasing up to ~1980. However, the MMM is again dominated by the ‘strong’ models after ~1900. By decomposing the surface density changes into both the part due to changes in surface salinity and surface temperatures (RhoS and RhoT respectively, see section 2b for details of computation), we find that the long-term increase in surface density is dominated by changes in salinity (i.e. RhoS) in the ‘strong’ models. Salinity also contributes to the increase in SPNA density up to ~1910 in both ‘strong’ and ‘weak’ models.

In contrast, RhoT appears to contribute little to the overall changes in surface density until after 1970, when SPNA warming causes a relatively fast decline in SPNA surface density anomalies. Note, that this result is in contrast to Hassan et al. (2021), who argued that SST changes dominated surface density, and suggests their analysis is dominated by the decreasing AMOC (and related RhoT changes) after 1980. However, we also note there are small but significant positive RhoT anomalies in ~1900–1910 and in ~1950–1980 in the ‘strong’ models. These are both periods of large deviations in ΔT (see fig. 9) and SHF between ‘strong’ and ‘weak’ (see fig. 4, suggesting that AA driven atmospheric cooling did contribute to the surface SPNA density anomalies in these

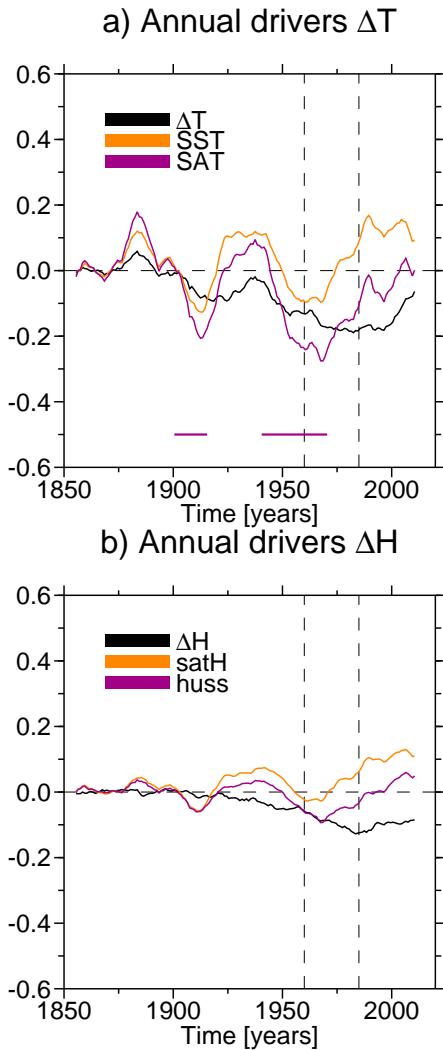


FIG. 9. Atmospheric and oceanic contributions to air-sea contrasts in temperature and humidity (e.g., ΔT and ΔH) averaged over the SPNA ($50\text{--}65^\circ\text{N}$ and $70\text{--}0^\circ\text{W}$ - note the smaller latitudinal area used is chosen to represent the region of positive SST anomalies in figure 5). a) shows the difference in annual-mean ΔT (e.g. SAT *minus* SST, black), SST (orange), and SAT (purple) between the ‘strong’ and ‘weak’ ensemble (e.g., ‘strong’ *minus* ‘weak’). b) shows the same but now for ΔH (e.g. huss *minus* satH), surface specific humidity (huss), and saturated humidity (satH) computed from SST. $satH = 0.98\rho^{-1}640380(kg\ m^{-3})e^{-5107.4/SST}$ where ρ is the density of air and assumed to be $1.22\ kg\ m^{-3}$ (Large and Yeager 2009). Note that IPSL-CM6 and NESM3 have been removed from panel b) due to problems with humidity data. Purple horizontal bars in a) approximately indicates where the negative tendency in SAT is larger than for SST, and so indicates where the atmosphere is driving the ocean rather than the other way round.

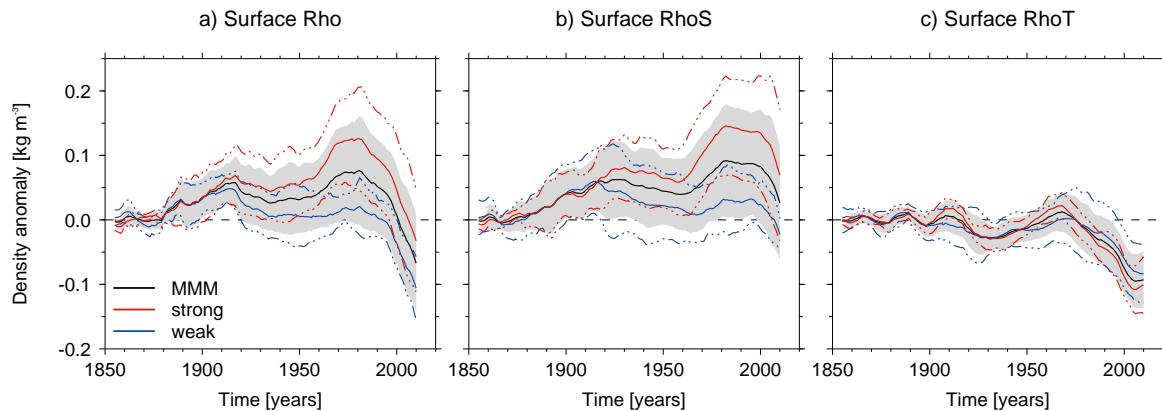


FIG. 10. Evolution of SPNA annual-mean surface density anomalies in CMIP6 historical simulations, and the role of salinity and temperature. a) shows the surface density anomalies in CMIP6 MMM (black), and ‘strong’ (red) and ‘weak’ (blue) sub-ensembles respectively. Grey shading and thin coloured lines show the $\pm 1\sigma$ spread. b) shows the same as a) but now for the contribution of salinity, e.g. by using climatological SST (see text for details). c) shows the same as b), but now for the role of temperature by using climatological SSS. All timeseries are smoothed with a 10-year running mean.

time periods. However, it is clear that the overall influence of SST on ocean surface density trends was small.

At first glance, these changes in surface density due to salinity are consistent with previous work that explore the impact of AA on AMOC via analysis of surface density anomalies (e.g. Delworth and Dixon (2006) and Menary et al. (2013)). However, we argue that these salinity driven surface density anomalies are primarily a feedback onto AMOC and not an initial driver of the AMOC. For example, PmE only diverges between ‘strong’ and ‘weak’ after ~ 1940 (see fig. 3), but surface density diverges around ~ 1900 . Furthermore, we also find that the initial divergence in surface density appears to be due to surface temperature, whereas the surface density anomaly due to salinity only diverges at about ~ 1915 (see fig 10). Similarly, lagged correlations computed using the whole time-series 1850–2014, shows that the maximum correlation is found when AMOC leads the surface density anomalies due to salinity by a few years (not shown). For models where data is available, we also find that ocean freshwater flux divergence anomalies dominate the full-depth salinity budget in the SPNA and that the freshwater flux divergence in ‘strong’ and ‘weak’ models diverge in ~ 1900 , consistent with changes in the AMOC (not shown).

7. Discussion

a. Summary of the proposed mechanism

As already shown by Menary et al. (2020), the evolution of externally-forced changes in the AMOC in historical simulations made for CMIP6 is highly dependent on the strength of the AA forcing / climate system response to AA forcing. Here we have shown that the spread in the model AMOC response appears to be mainly related to the strength of anomalous turbulent heat flux cooling over the SPNA which is, in turn, related to the fact that the atmosphere is anomalously cooler and drier in models with strong AA forcing. We hypothesize that this difference in turbulent heat flux cooling is largely a result of the differences in local cooling over the continents, and specifically the North American continent. Nevertheless, we emphasize that this mechanism is in addition to the more ‘direct’ impact of changes in downwelling SW over the SPNA. In ‘weak’ models, the negative anomalies in netSW over the SPNA clearly counteract positive turHF anomalies (see fig. 4) and, therefore, likely delays a reduction in AMOC. However, this ‘direct’ impact on the SPNA heat budget does not explain the difference between ‘strong’ and ‘weak’, as shown by the lack of correlation in figure 4g.

The hypothesized mechanism to explain the difference in ‘strong’ and ‘weak’ is summarised in the schematic shown in figure 11 and progresses as follows. First, increased aerosol and aerosol precursor emissions drive negative radiative forcing over the North American continent. The AA forcing leads to the rapid onset of cold and dry surface anomalies over land relative to the ocean (fig. 7c). These cold and dry anomalies are then advected across the North Atlantic Ocean (likely by the climatological winds) and lead to increased air-sea temperature and humidity contrasts and, hence, increased SHF cooling of the SPNA due to turHF (fig 6). The increased SHF cooling drives increased transformation of surface water due to increased surface density flux (fig. 3) and, subsequently, a strengthening of the AMOC.

The strengthening of the AMOC also feeds back onto itself positively due to related changes in ocean heat and salt transports. In particular, increased ocean heat transport warms the SPNA SST, which further amplifies air-sea contrasts in temperature and humidity and, hence, turbulent heat fluxes. The SST feedback broadens and delays the peak SHF cooling of the SPNA and, potentially, the AMOC strengthening. Increased salt (or reduced freshwater) transport also causes

forcing. Therefore, we now address whether we can constrain the historically-forced AMOC response by comparing the simulation of key variables against observations. Specifically, we compare SPNA surface properties and surface temperatures over the North American continent and over the Northern Hemisphere. We do this by computing linear trends from all individual ensemble members and compare them with contemporaneous trends from the observations. Due to the lack of SPNA data over 1850–1900, we focus on trends over 1900–1985.

As in Menary et al. (2020), we find that SPNA SST is a weak constraint on the externally-forced AMOC response. Figure 12a shows that the MMM SPNA SST shows noticeable variability throughout the historical period, with a warming peak ~1930, a minimum ~1970, and a warming post-~1985. However, there is little difference in the ‘strong’ and ‘weak’ models ensemble-mean overall evolution, although ‘strong’ models show slightly colder anomalies in ~1910 and ~1970, consistent with increased surface flux cooling in these time-periods (e.g. fig. 9). However, it is clear that the ensemble-mean temperature anomalies (i.e. the externally forced response) are small in terms of long-term trends and also small compared to observed or simulated multi-decadal variability. Therefore, when computing trends of the SPNA SST (i.e. fig. 12e) the observed trend falls within the 5–95% confidence interval of nearly all models.

SPNA SSS increases in the historical period up to ~1980 in the ‘strong’ models, and these changes appear inconsistent with observations. Figure 12b shows the time-series of SSS. In the ‘strong’ models, SPNA surface salinity increases from 1850, but with a more rapid increase between 1950–1985 largely in response to the simulated AMOC increase. In comparison, ‘weak’ models shows a smaller increase from 1850–1910, but then show little trend thereafter. However, in contrast, observations show a reduction in SPNA surface salinity over the historical period, and especially following the 1960s, the well-known Great Salinity Anomaly (Dickson et al. 1988). When we compare trends between the models and observations in Figure 12f, we find that the observed trend is more consistent with ‘weak’ models. The observed trend falls outside the 5–95% confidence interval for the majority of ‘strong’ models (7 of 9), but is within the majority of ‘weak’ models.

We also find that the simulated changes in surface temperatures in the ‘strong’ models are not consistent with observations. Figure 12c shows that North American continental surface air temperature (i.e., SAT_NAm) cools in the MMM, largely due to the substantial cooling in

the ‘strong’ models. However, we find that the observed continental surface temperature warms almost continuously over the historical period following ~ 1890 . There is a small observed cooling over ~ 1940 – 1970 , but this cooling is substantially smaller than is simulated in the ‘strong’ model ensemble-mean. When comparing the 1900–1985 trends in SAT_NAm in figure 12g we find that the observed trend is not consistent with the ‘strong’ models (trend is outside the 5–95% confidence interval of trends in almost all ‘strong’ models (8 out of 9)), but is consistent with all the ‘weak’ models. Furthermore, figure 12d shows that there is also only as small warming in Northern Hemisphere surface temperatures (ST_NH) before ~ 1985 in the MMM, which is again dominated by no net warming in the ‘strong’ models and a hemispheric cooling circa 1950–1975. However, in contrast, the observed Northern Hemisphere generally warms apart from ~ 1940 – 1970 , but this cooling is again substantially smaller than in the ‘strong’ model ensemble-mean. When comparing the 1900–1985 trends in ST_NH in figure 12h we find that the observed trend is not consistent with the ‘strong’ models (trend is outside the 5–95% confidence interval of trends in all ‘strong’ models), but it is consistent with most of the ‘weak’ models (5 out of 8). Therefore, we infer that the anomalous turbulent heat flux anomalies, which are driven by the cooling atmosphere, are very likely to be over estimated within the strong models.

Taken together, the evidence presented here suggests that the magnitude of the historical externally-forced AMOC strengthening in the ‘strong’ CMIP6 models is not consistent with observations. In particular, we show clearly that the simulation of North American surface air temperatures and Northern Hemisphere surface temperatures are not consistent with observations in the ‘strong’ models (i.e., see fig 12c). As these are key regions affected by AA forcing, we conclude that the AA forcing, or the response to AA forcing, is too large in the ‘strong’ models. Such a conclusion is consistent with a number of studies that document a too large AA driven cooling in CMIP6 historical simulations (Flynn and Mauritsen 2020; Dittus et al. 2020; Wang et al. 2021; Zhang et al. 2021). As there is a strong relationship between the strength of the AA forcing/response and the changes of the AMOC in CMIP6-class models (i.e. fig. 1), it follows that the AMOC response is overestimated too, irrespective of the exact mechanism. This conclusion is also supported overall by the comparison of simulated SPNA salinity with observations, which indicates that the SPNA salinity increases too much in the ‘strong’ models, which appears consistent with a too large AMOC increase.

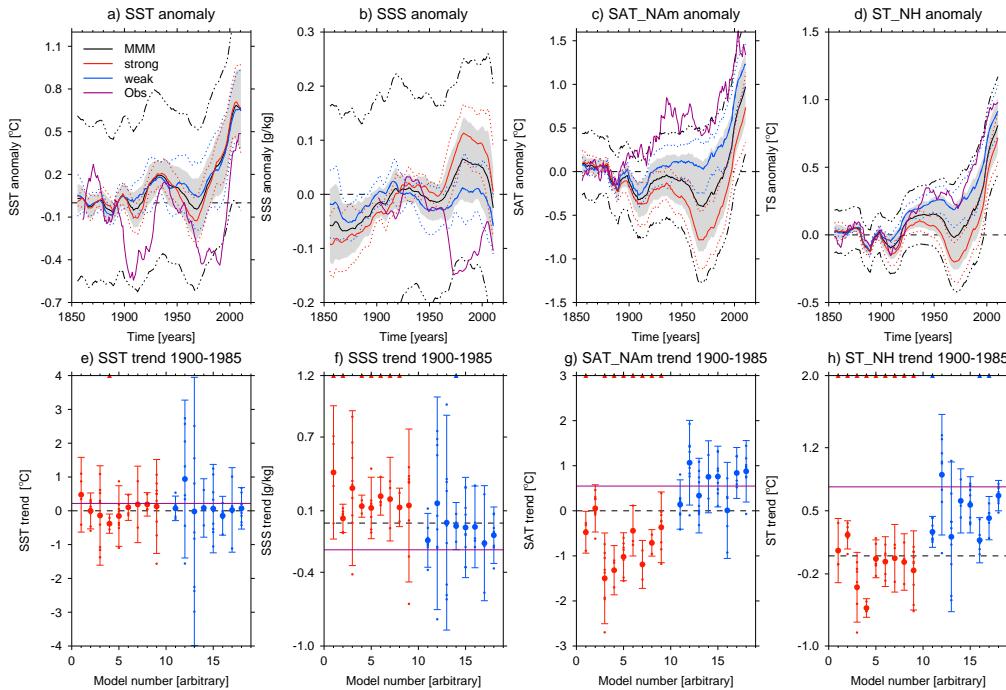


FIG. 12. Shows the comparison of historical simulations with observations for key metrics. a) shows anomalies for SPNA sea surface temperature (SST) relative to 1850-1899, for the MMM (black), strong (red) and weak (blue) ensembles and observations (purple, ERSSTv5). 1σ spread of the individual models ensemble-means is shown by the black shading or red and dotted lines for MMM and strong and weak models respectively. Black dot-dash line shows the 5–95% confidence interval of the MMM ensemble mean based on all the individual members. b), c) and d) shows the same for SPNA sea surface salinity (SSS), North American surface air temperature (SAT_NAm), and Northern Hemisphere surface temperature (ST_NH, e.g., SAT and SST combined and averaged over 0–60°N). Observations are given by the ‘Anually binned Sea Surface Salinity’ data set from Reverdin et al. (2019) for b), and BEST for c) and d). Note anomalies for SSS are computed over 1900-1950 due to shorter observations. Note also that panel c) shows the same time-series as shown in figure 7c, but now for the 1850–1900 climatological period. e) shows the comparison of observed and simulated linear trends of SPNA sea surface temperature (SST) computed over 1900-1985 for each member of each model (small dots show each member, and large dot shows the ensemble mean for each model). Error bars show the 5–95% confidence interval for the model ensemble means computed by using the individual model members. Horizontal purple line shows the observed trend computed over the same time period. Triangles at the top of the plot shows where observations are outside the 5–95% confidence interval. f)-h) shows the same as e) but now for SSS, SAT_NAm, and TS_NH.

c. Wider implications and caveats

In this paper we have tried to present a broad and in-depth analysis of the role of AA forcing in driving the AMOC in CMIP6 historical simulations. The analysis presented here raises a range of questions which will need to be addressed in future research. In particular, questions remain regarding the exact role of the positive feedbacks outlined in section 6, the role of the atmosphere in driving SHF anomalies, the regions most important for mediating the AA-driven AMOC response, and on the wider implications.

Although positive feedbacks appear in operation in the simulated evolution of the externally forced AMOC, their relative importance remains unclear. For example, to what extent does the positive feedback on heat fluxes contribute to the increasing AMOC? Furthermore, how important are the salinity driven surface density changes, which have been the focus of previous studies (e.g., Menary et al. (2013)), compared to the AA driven SHF anomalies? It is also important to emphasize that the feedbacks discussed in section 6 are always positive *only* in the presence of continued externally forced cooling of the SPNA due to AA forcing. Otherwise, the role of ocean heat transports and the subsequent SPNA warming could dominate density anomalies and contribute to a reduction in SPNA surface density. Indeed, this temperature-driven negative feedback may be a key driver of the decreasing AMOC in ‘strong’ models after ~1985, but more work is needed to understand this time period which is complicated by large GHG forced changes (Menary et al. 2020). Future work could address these questions through computation of surface water mass transformations across the SPNA in models in order to understand the relative role of surface density fluxes as opposed to surface density anomalies in driving anomalous transformations (i.e., as in Petit et al. (2021)), or to focus on the importance of different regions (e.g. Yeager et al. (2021)). Given the importance of salinity highlighted here, further detailed analysis and budgets are also necessary to understand the key drivers of the SPNA salinity changes in ‘strong’ models. However, many of the important variables related to freshwater budgets (particularly those related to sea ice) were not available for analysis across all models, and full-depth ocean transports may not be sufficient. Therefore, further careful computation of freshwater transports to focus on the upper ocean is needed.

Despite a significant atmospheric circulation response in the historical simulations related to AA forcing (e.g., figure 5), the AMOC mechanism presented here is essentially a thermodynamically driven mechanism. In other words, changes in the atmospheric circulation strength play little role

in explaining the difference in AMOC evolution between ‘strong’ and ‘weak’. However, the ΔT anomalies could, in part, be related to changes in surface atmospheric temperature associated with changes in atmospheric circulation (e.g., a shift in the jet position (Ma et al. 2020)), which we have not explored here. Therefore, further analysis of the forced response in atmospheric circulation is needed, not least to understand the mechanistic origin of the response and its realism. It is also important to emphasize that models have known deficiencies in simulating the atmospheric circulation response to a range of drivers, including external-forcing - the signal-to-noise problem (Scaife and Smith 2018). Therefore, the extent by which AA forcing drove the AMOC in the real world is still unclear as we cannot trust that the atmospheric circulation response is adequately represented.

We also recognise that the hypothesis that North American continental region is *the* key region for mediating the AMOC response to AA in the ‘strong’ models is also uncertain. This is because AA forcing is the ultimate driver of a range of climate signals in the ‘strong’ models (e.g., surface radiation budgets and surface temperature over the Northern Hemisphere, see figure S2). Furthermore, the AMOC and climate related feedbacks described above and other confounding factors (e.g., contemporaneous external forcing) make it more difficult to ascribe causality. In particular, the role of volcanic forcing has also not been assessed here, and natural forcings do contribute to AMOC variability in CMIP6 historical simulations Menary et al. (2020). Changes in land surface use can also be an important driver of North American continental surface temperatures (Andrews et al. 2017), and could thus contribute to the differences between models via the proposed mechanism here. However, initial analysis of fixed land use simulations made with UKESM1-0-LL as part of the Land Use Model Intercomparison Project (LUMIP, see Lawrence et al. (2016)) suggests a relatively small impact on ASR_{HD} (not shown) and, we assume, AMOC, but further analysis is needed. Nevertheless, this range of uncertainties highlights the difficulty in fully unraveling the full mechanism using only diagnostic analysis of historical simulations. Therefore, the importance of AA forcing in different regions and the processes that mediate the AMOC response still need to be tested through further targeted experiments; for example, by introducing idealised regional downwelling shortwave anomalies. Future work could also examine the importance of continental surface air temperature responses in mediating the AMOC response to other forcings such as volcanoes etc.

The results also raise a number of additional questions, which need further investigation. For example, it is not clear if the mismatch between the evolution of historical AMOC in MMM CMIP6 simulations and that interpreted through indirect measurements (e.g., Caesar et al. (2018, 2021)) is explained by the too large AMOC response in the ‘strong’ models. Therefore, questions still remain about how models represent the evolution of AMOC over the historical period. There are also questions about what these results mean for future AMOC changes. For example, Weijer et al. (2020) showed that the rate of AMOC decline in CMIP6 projections is model dependent. Given the large reduction in AMOC in ‘strong’ models after ~1985 (see fig. 1) it is plausible that some of this difference is related to the too large historical AMOC changes in ‘strong’ models and the sensitivity to AA forcing, but it is unclear if this is the case. It is also not clear what aspects of the AA forcing and climate response is leading to the large ASR_{HD} and AMOC response in the ‘strong’ models, or whether the processes are the same in all time periods. Indeed, there is still large uncertainty in the models’ Effective Radiative Forcing and the spatial distribution of forcing (Smith et al. 2021). Furthermore, Weijer et al. (2020) and Menary et al. (2020) both show a correlation between models with a strong transient climate response (TCR) and historical AMOC trends (i.e. models with a larger AMOC trend have a larger TCR). Therefore, it is likely that both the time-varying instantaneous AA radiative forcing and the range of climate adjustments and feedbacks (including temperature related feedbacks) are important for shaping the overall response of AMOC to AA forcing. Hence, future work is needed to better constrain the response of the climate system to AA forcing and to understand the ASR_{HD} and AMOC relationship in different time periods. However, for now, it is crucial that research reflects the diversity, and uncertainty, in the AA forcing and response rather than focusing on the MMM.

We should also note briefly that there are also inevitably caveats to the analysis presented here. For example, we have not included all CMIP6 models, which reflects the need for multiple variables and ensemble members. Therefore, we chose to focus on a smaller ensemble of models where we have complete datasets for at least 3 members. However, Menary et al. (2020) has already shown that the relationship between ASR_{HD} and AMOC holds in a larger range of CMIP5 and CMIP6 models and we assume that our results will broadly hold across other models. Nevertheless, the exact details of the mechanism will likely differ from model to model due to model differences. The limited number of ensemble members will also inevitably lead to some uncertainties for model-

dependent results (Tandon and Kushner 2015). Although we mitigate this sampling uncertainty through the creation of the ‘strong’ and ‘weak’ sub-ensembles, we note that the representation of internal variability remains another key uncertainty in understanding AMOC variability and its impact (Yan et al. 2018; Bonnet et al. 2021; Jackson et al. 2022).

One aspect that needs further exploring is the exact role of sea ice in mediating the response of SPNA heat fluxes to AA forcing. Indeed, sea ice grows in the ‘strong’ models (not shown, but see Robson et al. (2020) for analysis of UKESM1-0-LL) and so questions arise about how this affects heat and freshwater fluxes and SHF in the SPNA and Arctic (which can also be important for AMOC (Sévellec et al. 2017; Zhang and Thomas 2021)). Sea ice can also be heavily biased, which affects relationship with SPNA SHFs (not shown) - for this reason we chose not to include CanESM5 in the analysis due to a large positive sea ice bias in the western SPNA (Swart et al. 2019).

8. Conclusions

We have analysed CMIP6 historical simulations in order to understand the processes leading to the anthropogenic aerosol (AA) forced increase in Atlantic Meridional Overturning Circulation (AMOC) over the period 1850–1985 shown in Menary et al. (2020). We explore ensemble-mean changes in AMOC between models, and using an inter-hemispheric proxy of AA forcing response we split models between ‘strong’ or ‘weak’ AA forcing. The key results are as follows.

- There is a large spread in the simulated AMOC changes in CMIP6 historical simulations which is related to the spread in the strength of the AA forcing response. In particular, the multi-model mean (MMM) response is dominated by anomalous AMOC in the ‘strong’ models. In contrast, the ‘weak’ models show little change in AMOC.
- Surface density fluxes due to surface heat flux (SHF) cooling and resultant density fluxes over the subpolar North Atlantic (45–65°N, SPNA) appear to dominate the simulated increase in AMOC in ‘strong’ models over 1850–1985 (consistent with Hassan et al. (2021) who found a similar relationship over 1940–2014 for the MMM).
- Differences in downwelling shortwave radiation over the SPNA do not explain the difference between net SHF over the SPNA, nor the evolution of AMOC in models. Instead, turbulent

heat fluxes dominate the spread in AMOC responses. Hence, AA forcing does not act directly on the AMOC through only changes in the radiation balance over the SPNA.

- The anomalies in annual-mean turbulent heat flux appear to be dominated by changes in temperature and humidity across the air-sea interface (ΔT and ΔH , respectively), rather than changes in wind speed. Therefore, AMOC is primarily driven by a colder and drier atmosphere in the ‘strong’ models compared to the ‘weak’.
- The colder and drier atmosphere is consistent with larger AA forced changes in downwelling shortwave and, hence, colder and drier temperatures over the continental regions, especially over North America. Hence, the increase in AMOC appears consistent with the advection of cold and dry air from the continents over the ocean which drives increased turbulent heat fluxes through increased air-sea contrasts.
- AMOC-related changes in ocean salinity and surface temperatures in the SPNA act to amplify the AMOC increase. Specifically, the increase in AMOC leads to a saltier SPNA that increases surface density anomalies. Additionally, despite the strong SHF cooling in the ‘strong’ models, the SPNA warms in response to the strengthened AMOC. This warming acts as a positive feedback on the SHF cooling and surface density flux by increasing the air-sea temperature and humidity contrasts. However, it is less clear if this SHF feedback affects the AMOC.
- The evolution of surface temperature over North America and the Northern Hemisphere in ‘strong’ models over 1900–1985 is not consistent with observations. Furthermore, ‘strong’ models also simulate the wrong sign of SPNA surface salinity trends. Therefore, comparisons with observations suggests that the AMOC increase in the ‘strong’ models is likely to be overestimated.

In this paper we have highlighted the uncertainty in the historically forced AMOC response in CMIP6 models, and its dependence on the uncertainty in AA forcing. We have presented the surprising findings that diversity in AA forced AMOC change is explained by the spread in turbulent heat flux anomalies over the SPNA (rather than the changes in downwelling shortwave), and that the AA driven cooling over the continental regions mediates the AMOC response to AA forcing. In all models, AA forcing drives the AMOC directly by changing the radiation budget over the North Atlantic, but in ‘strong’ models the indirect effect of AA forcing on AMOC, through its

impact on atmospheric temperature and humidity, is at least as important. Furthermore, we provide evidence that this indirect “thermodynamically” driven mechanism, and hence the increase in the externally forced AMOC change, are likely too large in the ‘strong’ models, and thus in the CMIP6 MMM.

Given the uncertainty in AA forcing, and the AMOC response to it, care is needed in using and interpreting simulated changes in the North Atlantic in CMIP6 historical runs. More research is needed to understand the wider implications of the spread in historical AMOC responses, including the potential impact on future projections and in constraining past changes. Further research is also needed to understand if the mechanisms presented here can be generalized to other forcings (e.g., volcanic) and other time periods as well as to understand the relevance to the real world. Indeed, the uncertainties in the external forcing and the simulated response in ocean and atmosphere make it difficult to attribute past AMOC changes based on CMIP6 models alone.

To make progress it is crucial to minimise the uncertainty in historical aerosol forcing and how the climate system responds to those changes. It is also important to better understand the representation and drivers of AMOC and how the ocean and atmosphere responds to different external forcing in models (including the signal-to-noise problems (Scaife and Smith 2018)) and in the real world. Note that we were only able to discern the role of turbulent heat fluxes and continental temperature anomalies in CMIP6 models by contrasting models with ‘strong’ and ‘weak’ AMOC response. Therefore, these results highlight the continuing need to understand the range of model responses and mechanisms to AA forcing, and not just focus on one model or the multi-model mean.

Acknowledgments. We thank John Fasullo and two anonymous reviewers for their thoughtful comments that improved the clarity of the manuscript. This study was funded by the NERC ACSIS program (NE/N018001/1 and NE/N018044/1). JR was additionally funded by NERC via NCAS, as well as the WISHBONE (NE/T013516/1) and SNAP-DRAGON (NE/T013494/1) projects. CJ was funded by the NERC UKESM program (NE/N017978/1). We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. This work used JASMIN, the UK collaborative data analysis facility, and model data was analysed with CDO (Schulzweida 2021).

Data availability statement. The CMIP6 model data used in this study was downloaded from the ESGF (e.g. <https://esgf-index1.ceda.ac.uk/>) and are freely available. References for the simulations used are in the supplementary material. The binned North Atlantic subpolar surface salinity data was made freely available by the French Sea Surface Salinity Observation Service (<http://www.legos.obs-mip.fr/observations/sss/>). ERSST was freely available from the National Centers for Environmental Information (<https://www.ncei.noaa.gov/>). The Berkeley Earth Surface Temperature data set was downloaded from <https://berkeleyearth.org/data/>.

References

- Andrews, M. B., and Coauthors, 2020: Historical simulations with hadgem3-gc3.1 for cmip6. *Journal of Advances in Modeling Earth Systems*, **12** (6), e2019MS001995, <https://doi.org/10.1029/2019MS001995>.
- Andrews, T., R. A. Betts, B. B. Booth, C. D. Jones, and G. S. Jones, 2017: Effective radiative forcing from historical land use change. *Climate Dynamics*, **48** (11), 3489–3505, <https://doi.org/10.1007/s00382-016-3280-7>.
- Bellomo, K., M. Angeloni, S. Corti, and J. von Hardenberg, 2021: Future climate change shaped by inter-model differences in atlantic meridional overturning circulation response. *Nature Communications*, **12** (1), 1–10, <https://doi.org/10.1038/s41467-021-24015-w>.

- Blaker, A. T., J. J.-M. Hirschi, G. McCarthy, B. Sinha, S. Taws, R. Marsh, A. Coward, and B. de Cuevas, 2015: Historical analogues of the recent extreme minima observed in the atlantic meridional overturning circulation at 26 n. *Climate dynamics*, **44** (1-2), 457–473, <https://doi.org/https://doi.org/10.1007/s00382-014-2274-6>.
- Bonnet, R., and Coauthors, 2021: Increased risk of near term global warming due to a recent amoc weakening. *Nature communications*, **12** (1), 1–9, <https://doi.org/https://doi.org/10.1038/s41467-021-26370-0>.
- Buckley, M. W., and J. Marshall, 2016: Observations, inferences, and mechanisms of the atlantic meridional overturning circulation: A review. *Reviews of Geophysics*, **54** (1), 5–63, <https://doi.org/https://doi.org/10.1002/2015RG000493>.
- Caesar, L., G. McCarthy, D. Thornalley, N. Cahill, and S. Rahmstorf, 2021: Current atlantic meridional overturning circulation weakest in last millennium. *Nature Geoscience*, **14** (3), 118–120, <https://doi.org/https://doi.org/10.1038/s41561-021-00699-z>.
- Caesar, L., S. Rahmstorf, A. Robinson, G. Feulner, and V. Saba, 2018: Observed fingerprint of a weakening atlantic ocean overturning circulation. *Nature*, **556** (7700), 191–196, <https://doi.org/https://doi.org/10.1038/s41586-018-0006-5>.
- Cai, W., D. Bi, J. Church, T. Cowan, M. Dix, and L. Rotstayn, 2006: Pan-oceanic response to increasing anthropogenic aerosols: Impacts on the southern hemisphere oceanic circulation. *Geophysical Research Letters*, **33** (21), <https://doi.org/https://doi.org/10.1029/2006GL027513>.
- Cheng, W., J. C. H. Chiang, and D. Zhang, 2013: Atlantic meridional overturning circulation (amoc) in cmip5 models: Rcp and historical simulations. *Journal of Climate*, **26** (18), 7187 – 7197, <https://doi.org/10.1175/JCLI-D-12-00496.1>, URL <https://journals.ametsoc.org/view/journals/clim/26/18/jcli-d-12-00496.1.xml>.
- Couldrey, M. P., and Coauthors, 2021: What causes the spread of model projections of ocean dynamic sea-level change in response to greenhouse gas forcing? *Climate Dynamics*, **56** (1), 155–187, <https://doi.org/https://doi.org/10.1007/s00382-020-05471-4>.

- Delworth, T., S. Manabe, and R. Stouffer, 1993: Interdecadal variations of the thermohaline circulation in a coupled ocean–atmosphere model. *J. Climate*, **6**, 1993–2011, [https://doi.org/https://doi.org/10.1175/1520-0442\(1993\)006<1993:IVOTTC>2.0.CO;2](https://doi.org/https://doi.org/10.1175/1520-0442(1993)006<1993:IVOTTC>2.0.CO;2).
- Delworth, T. L., and K. W. Dixon, 2006: Have anthropogenic aerosols delayed a greenhouse gas-induced weakening of the north atlantic thermohaline circulation? *Geophysical Research Letters*, **33** (2), <https://doi.org/https://doi.org/10.1029/2005GL024980>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL024980>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2005GL024980>.
- Delworth, T. L., and F. Zeng, 2016: The impact of the north atlantic oscillation on climate through its influence on the atlantic meridional overturning circulation. *Journal of Climate*, **29** (3), 941 – 962, <https://doi.org/10.1175/JCLI-D-15-0396.1>, URL <https://journals.ametsoc.org/view/journals/clim/29/3/jcli-d-15-0396.1.xml>.
- Dickson, R. R., J. Meincke, S.-A. Malmberg, and A. J. Lee, 1988: The “great salinity anomaly” in the northern north atlantic 1968–1982. *Progress in Oceanography*, **20** (2), 103–151, [https://doi.org/https://doi.org/10.1016/0079-6611\(88\)90049-3](https://doi.org/https://doi.org/10.1016/0079-6611(88)90049-3), URL <https://www.sciencedirect.com/science/article/pii/0079661188900493>.
- Dittus, A. J., E. Hawkins, J. I. Robson, D. M. Smith, and L. J. Wilcox, 2021: Drivers of recent north pacific decadal variability: the role of aerosol forcing. *Earth’s Future*, **n/a** (n/a), e2021EF002 249, <https://doi.org/https://doi.org/10.1029/2021EF002249>.
- Dittus, A. J., E. Hawkins, L. J. Wilcox, R. T. Sutton, C. J. Smith, M. B. Andrews, and P. M. Forster, 2020: Sensitivity of historical climate simulations to uncertain aerosol forcing. *Geophysical Research Letters*, **47** (13), e2019GL085 806, <https://doi.org/https://doi.org/10.1029/2019GL085806>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085806>, e2019GL085806 10.1029/2019GL085806, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL085806>.
- Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, **9** (5), 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, URL <https://gmd.copernicus.org/articles/9/1937/2016/>.

- Fiedler, S., and D. Putrasahan, 2021: How does the north atlantic sst pattern respond to anthropogenic aerosols in the 1970s and 2000s? *Geophysical Research Letters*, **48** (7), e2020GL092142, <https://doi.org/https://doi.org/10.1029/2020GL092142>.
- Flynn, C. M., and T. Mauritsen, 2020: On the climate sensitivity and historical warming evolution in recent coupled model ensembles. *Atmospheric Chemistry and Physics*, **20** (13), 7829–7842, <https://doi.org/10.5194/acp-20-7829-2020>, URL <https://acp.copernicus.org/articles/20/7829/2020/>.
- Gillett, N. P., and Coauthors, 2016: The detection and attribution model intercomparison project (damip v1.0) contribution to cmip6. *Geoscientific Model Development*, **9** (10), 3685–3697, <https://doi.org/10.5194/gmd-9-3685-2016>, URL <https://gmd.copernicus.org/articles/9/3685/2016/>.
- Gregory, J. M., and Coauthors, 2016: The flux-anomaly-forced model intercomparison project (fafmip) contribution to cmip6: investigation of sea-level and ocean climate change in response to co₂ forcing. *Geoscientific Model Development*, **9** (11), 3993–4017, <https://doi.org/10.5194/gmd-9-3993-2016>, URL <https://gmd.copernicus.org/articles/9/3993/2016/>.
- Griffies, S. M., and Coauthors, 2014: An assessment of global and regional sea level for years 1993–2007 in a suite of interannual core-ii simulations. *Ocean Modelling*, **78**, 35–89, <https://doi.org/https://doi.org/10.1016/j.ocemod.2014.03.004>, URL <https://www.sciencedirect.com/science/article/pii/S1463500314000407>.
- Grist, J. P., S. A. Josey, R. Marsh, Y.-O. Kwon, R. J. Bingham, and A. T. Blaker, 2014: The surface-forced overturning of the north atlantic: Estimates from modern era atmospheric reanalysis datasets. *Journal of Climate*, **27** (10), 3596 – 3618, <https://doi.org/10.1175/JCLI-D-13-00070.1>, URL <https://journals.ametsoc.org/view/journals/clim/27/10/jcli-d-13-00070.1.xml>.
- Hassan, T., R. J. Allen, W. Liu, and C. A. Randles, 2021: Anthropogenic aerosol forcing of the atlantic meridional overturning circulation and the associated mechanisms in cmip6 models. *Atmospheric Chemistry and Physics*, **21** (8), 5821–5846, <https://doi.org/10.5194/acp-21-5821-2021>, URL <https://acp.copernicus.org/articles/21/5821/2021/>.

- Huang, B., and Coauthors, 2017: Extended reconstructed sea surface temperature, version 5 (ersstv5): Upgrades, validations, and intercomparisons. *Journal of Climate*, **30** (20), 8179 – 8205, <https://doi.org/10.1175/JCLI-D-16-0836.1>, URL <https://journals.ametsoc.org/view/journals/clim/30/20/jcli-d-16-0836.1.xml>.
- Jackson, L. C., A. Biastoch, M. W. Buckley, D. G. Desbruyères, E. Frajka-Williams, B. Moat, and J. Robson, 2022: The evolution of the North Atlantic Meridional Overturning Circulation since 1980. *Nature Reviews Earth & Environment*, **3** (4), 241–254, <https://doi.org/10.1038/s43017-022-00263-2>, URL <https://doi.org/10.1038/s43017-022-00263-2>.
- Jackson, L. C., and Coauthors, 2020: Impact of ocean resolution and mean state on the rate of amoc weakening. *Climate Dynamics*, **55** (7), 1711–1732, <https://doi.org/https://doi.org/10.1007/s00382-020-05345-9>.
- Joshi, M. M., J. M. Gregory, M. J. Webb, D. M. Sexton, and T. C. Johns, 2008: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Climate dynamics*, **30** (5), 455–465, <https://doi.org/10.1007/s00382-007-0306-1>.
- Kang, S. M., S.-P. Xie, C. Deser, and B. Xiang, 2021: Zonal mean and shift modes of historical climate response to evolving aerosol distribution. *Science Bulletin*, **66** (23), 2405–2411, <https://doi.org/https://doi.org/10.1016/j.scib.2021.07.013>.
- Kostov, Y., K. C. Armour, and J. Marshall, 2014: Impact of the atlantic meridional overturning circulation on ocean heat storage and transient climate change. *Geophysical Research Letters*, **41** (6), 2108–2116, <https://doi.org/https://doi.org/10.1002/2013GL058998>.
- Large, W., and S. Yeager, 2009: The global climatology of an interannually varying air–sea flux data set. *Climate dynamics*, **33** (2-3), 341–364, <https://doi.org/10.1007/s00382-008-0441-3>.
- Lawrence, D. M., and Coauthors, 2016: The land use model intercomparison project (lumip) contribution to cmip6: rationale and experimental design. *Geoscientific Model Development*, **9** (9), 2973–2998, <https://doi.org/10.5194/gmd-9-2973-2016>, URL <https://gmd.copernicus.org/articles/9/2973/2016/>.
- Lund, M. T., G. Myhre, and B. H. Samset, 2019: Anthropogenic aerosol forcing under the shared socioeconomic pathways. *Atmospheric Chemistry and Physics*, **19** (22), 13 827–13 839,

<https://doi.org/10.5194/acp-19-13827-2019>, URL <https://acp.copernicus.org/articles/19/13827/2019/>.

- Ma, L., T. Woollings, R. G. Williams, D. Smith, and N. Dunstone, 2020: How does the winter jet stream affect surface temperature, heat flux, and sea ice in the north atlantic? *Journal of Climate*, **33** (9), 3711 – 3730, <https://doi.org/10.1175/JCLI-D-19-0247.1>.
- Marshall, J., A. Donohoe, D. Ferreira, and D. McGee, 2014: The ocean's role in setting the mean position of the inter-tropical convergence zone. *Climate Dynamics*, **42** (7), 1967–1979, <https://doi.org/10.1007/s00382-013-1767-z>.
- Megann, A., A. Blaker, S. Josey, A. New, and B. Sinha, 2021: Mechanisms for late 20th and early 21st century decadal amoc variability. *Journal of Geophysical Research: Oceans*, **126** (12), e2021JC017865, <https://doi.org/https://doi.org/10.1029/2021JC017865>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JC017865>, e2021JC017865 2021JC017865, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021JC017865>.
- Menary, M. B., D. L. R. Hodson, J. I. Robson, R. T. Sutton, R. A. Wood, and J. A. Hunt, 2015: Exploring the impact of cmip5 model biases on the simulation of north atlantic decadal variability. *Geophysical Research Letters*, **42** (14), 5926–5934, <https://doi.org/https://doi.org/10.1002/2015GL064360>.
- Menary, M. B., and A. A. Scaife, 2014: Naturally forced multidecadal variability of the atlantic meridional overturning circulation. *Climate dynamics*, **42** (5-6), 1347–1362, <https://doi.org/https://doi.org/10.1007/s00382-013-2028-x>.
- Menary, M. B., and Coauthors, 2013: Mechanisms of aerosol-forced amoc variability in a state of the art climate model. *Journal of Geophysical Research: Oceans*, **118** (4), 2087–2096, <https://doi.org/https://doi.org/10.1002/jgrc.20178>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrc.20178>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/jgrc.20178>.
- Menary, M. B., and Coauthors, 2020: Aerosol-forced amoc changes in cmip6 historical simulations. *Geophysical Research Letters*, **47** (14), e2020GL088166, <https://doi.org/https://doi.org/10.1029/2020GL088166>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL088166>.

2020GL088166, e2020GL088166 2020GL088166, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL088166>.

- Moat, B. I., and Coauthors, 2019: Insights into decadal north atlantic sea surface temperature and ocean heat content variability from an eddy-permitting coupled climate model. *Journal of Climate*, **32** (18), 6137–6161, <https://doi.org/10.1175/JCLI-D-18-0709.1>, URL <https://journals.ametsoc.org/view/journals/clim/32/18/jcli-d-18-0709.1.xml>.
- Otterå, O. H., M. Bentsen, H. Drange, and L. Suo, 2010: External forcing as a metronome for atlantic multidecadal variability. *Nature Geoscience*, **3** (10), 688–694, <https://doi.org/https://doi.org/10.1038/ngeo955>.
- Petit, T., M. S. Lozier, S. A. Josey, and S. A. Cunningham, 2021: Role of air–sea fluxes and ocean surface density in the production of deep waters in the eastern subpolar gyre of the north atlantic. *Ocean Science*, **17** (5), 1353–1365, <https://doi.org/10.5194/os-17-1353-2021>, URL <https://os.copernicus.org/articles/17/1353/2021/>.
- Polo, I., J. Robson, R. Sutton, and M. A. Balmaseda, 2014: The importance of wind and buoyancy forcing for the boundary density variations and the geostrophic component of the amoc at 26°n. *Journal of Physical Oceanography*, **44** (9), 2387 – 2408, <https://doi.org/10.1175/JPO-D-13-0264.1>.
- Reverdin, G., A. R. Friedman, L. Chafik, N. P. Holliday, T. Szekely, H. Valdimarsson, and I. Yashayaev, 2019: North atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. part 1: dataset validation and rms variability. *Ocean Dynamics*, **69** (3), 385–403, <https://doi.org/https://doi.org/10.1007/s10236-018-1240-y>.
- Robson, J., R. Sutton, K. Lohmann, D. Smith, and M. D. Palmer, 2012: Causes of the rapid warming of the north atlantic ocean in the mid-1990s. *Journal of Climate*, **25** (12), 4116 – 4134, <https://doi.org/10.1175/JCLI-D-11-00443.1>.
- Robson, J., and Coauthors, 2020: The evaluation of the north atlantic climate system in ukesm1 historical simulations for cmip6. *Journal of Advances in Modeling Earth Systems*, **12** (9), e2020MS002126, <https://doi.org/https://doi.org/10.1029/2020MS002126>.

org/10.1029/2020MS002126, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002126>, e2020MS002126 10.1029/2020MS002126, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020MS002126>.

Rohde, R. A., and Z. Hausfather, 2020: The berkeley earth land/ocean temperature record. *Earth System Science Data*, **12** (4), 3469–3479, <https://doi.org/10.5194/essd-12-3469-2020>, URL <https://essd.copernicus.org/articles/12/3469/2020/>.

Sarmiento, J. L., and C. Le Quéré, 1996: Oceanic carbon dioxide uptake in a model of century-scale global warming. *Science*, **274** (5291), 1346–1350, <https://doi.org/10.1126/science.274.5291.1346>, URL <https://www.science.org/doi/abs/10.1126/science.274.5291.1346>.

Scaife, A. A., and D. Smith, 2018: A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, **1** (1), 1–8, <https://doi.org/https://doi.org/10.1038/s41612-018-0038-4>.

Schulzweida, U., 2021: Cdo user guide. URL <https://doi.org/10.5281/zenodo.5614769>, <https://doi.org/10.5281/zenodo.5614769>.

Sévellec, F., A. V. Fedorov, and W. Liu, 2017: Arctic sea-ice decline weakens the atlantic meridional overturning circulation. *Nature Climate Change*, **7** (8), 604–610, <https://doi.org/https://doi.org/10.1038/nclimate3353>.

Smith, C. J., and Coauthors, 2021: Energy budget constraints on the time history of aerosol forcing and climate sensitivity. *Journal of Geophysical Research: Atmospheres*, **126** (13), e2020JD033 622, <https://doi.org/https://doi.org/10.1029/2020JD033622>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JD033622>, e2020JD033622 2020JD033622, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD033622>.

Speer, K., and E. Tziperman, 1992: Rates of water mass formation in the north atlantic ocean. *Journal of Physical Oceanography*, **22** (1), 93 – 104, [https://doi.org/10.1175/1520-0485\(1992\)022<0093:ROWMFI>2.0.CO;2](https://doi.org/10.1175/1520-0485(1992)022<0093:ROWMFI>2.0.CO;2), URL https://journals.ametsoc.org/view/journals/phoc/22/1/1520-0485_1992_022_0093_rowmfi_2_0_co_2.xml.

Sutton, R. T., B. Dong, and J. M. Gregory, 2007: Land/sea warming ratio in response to climate change: Ipccl ar4 model results and comparison with observations. *Geophysical Research Letters*, **34** (2), <https://doi.org/https://doi.org/10.1029/2006GL028164>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL028164>.

onlinelibrary.wiley.com/doi/abs/10.1029/2006GL028164, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006GL028164>.

Sutton, R. T., G. D. McCarthy, J. Robson, B. Sinha, A. T. Archibald, and L. J. Gray, 2018: Atlantic multidecadal variability and the u.k. acsis program. *Bulletin of the American Meteorological Society*, **99** (2), 415 – 425, <https://doi.org/10.1175/BAMS-D-16-0266.1>, URL <https://journals.ametsoc.org/view/journals/bams/99/2/bams-d-16-0266.1.xml>.

Swart, N. C., and Coauthors, 2019: The canadian earth system model version 5 (canesm5.0.3). *Geoscientific Model Development*, **12** (11), 4823–4873, <https://doi.org/10.5194/gmd-12-4823-2019>, URL <https://gmd.copernicus.org/articles/12/4823/2019/>.

Swingedouw, D., P. Ortega, J. Mignot, E. Guilyardi, V. Masson-Delmotte, P. G. Butler, M. Khodri, and R. Séférian, 2015: Bidecadal north atlantic ocean circulation variability controlled by timing of volcanic eruptions. *Nature communications*, **6** (1), 1–12, <https://doi.org/https://doi.org/10.1038/ncomms7545>.

Tandon, N. F., and P. J. Kushner, 2015: Does external forcing interfere with the amoc's influence on north atlantic sea surface temperature? *Journal of Climate*, **28** (16), 6309 – 6323, <https://doi.org/10.1175/JCLI-D-14-00664.1>, URL <https://journals.ametsoc.org/view/journals/clim/28/16/jcli-d-14-00664.1.xml>.

Thornalley, D. J., and Coauthors, 2018: Anomalously weak labrador sea convection and atlantic overturning during the past 150 years. *Nature*, **556** (7700), 227–230, <https://doi.org/https://doi.org/10.1038/s41586-018-0007-4>.

Trenberth, K. E., Y. Zhang, J. T. Fasullo, and L. Cheng, 2019: Observation-based estimates of global and basin ocean meridional heat transport time series. *Journal of Climate*, **32** (14), 4567 – 4583, <https://doi.org/10.1175/JCLI-D-18-0872.1>, URL <https://journals.ametsoc.org/view/journals/clim/32/14/jcli-d-18-0872.1.xml>.

Undorf, S., M. A. Bollasina, B. B. Booth, and G. C. Hegerl, 2018: Contrasting the effects of the 1850–1975 increase in sulphate aerosols from north america and europe on the atlantic in the cesm. *Geophysical Research Letters*, **45** (21), 11,930–11,940, <https://doi.org/https://doi.org/10.1029/2018GL079970>.

- Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus*, **34** (2), 187–195.
- Wang, C., B. J. Soden, W. Yang, and G. A. Vecchi, 2021: Compensation between cloud feedback and aerosol-cloud interaction in cmip6 models. *Geophysical Research Letters*, **48** (4), e2020GL091024, <https://doi.org/https://doi.org/10.1029/2020GL091024>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091024>, e2020GL091024 2020GL091024, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091024>.
- Wang, C., L. Zhang, S.-K. Lee, L. Wu, and C. R. Mechoso, 2014: A global perspective on cmip5 climate model biases. *Nature Climate Change*, **4** (3), 201–205, <https://doi.org/https://doi.org/10.1038/nclimate2118>.
- Wang, Y., J. H. Jiang, and H. Su, 2015: Atmospheric responses to the redistribution of anthropogenic aerosols. *Journal of Geophysical Research: Atmospheres*, **120** (18), 9625–9641, <https://doi.org/https://doi.org/10.1002/2015JD023665>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023665>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015JD023665>.
- Weijer, W., W. Cheng, O. A. Garuba, A. Hu, and B. T. Nadiga, 2020: C mip6 models predict significant 21st century decline of the atlantic meridional overturning circulation. *Geophysical Research Letters*, **47** (12), e2019GL086075, <https://doi.org/https://doi.org/10.1029/2019GL086075>.
- Wilcox, L. J., E. J. Highwood, and N. J. Dunstone, 2013: The influence of anthropogenic aerosol on multi-decadal variations of historical global climate. *Environmental Research Letters*, **8** (2), 024033, <https://doi.org/10.1088/1748-9326/8/2/024033>, URL <https://doi.org/10.1088/1748-9326/8/2/024033>.
- Xu, X., E. P. Chassignet, and F. Wang, 2019: On the variability of the atlantic meridional overturning circulation transports in coupled cmip5 simulations. *Climate Dynamics*, **52** (11), 6511–6531, <https://doi.org/10.1007/s00382-018-4529-0>.
- Yan, X., R. Zhang, and T. R. Knutson, 2018: Underestimated amoc variability and implications for amv and predictability in cmip models. *Geophysical Research Letters*, **45** (9),

4319–4328, <https://doi.org/https://doi.org/10.1029/2018GL077378>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL077378>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL077378>.

Yeager, S., and Coauthors, 2021: An outsized role for the labrador sea in the multidecadal variability of the atlantic overturning circulation. *Science Advances*, **7** (41), eabh3592, <https://doi.org/10.1126/sciadv.abh3592>, URL <https://www.science.org/doi/abs/10.1126/sciadv.abh3592>, <https://www.science.org/doi/pdf/10.1126/sciadv.abh3592>.

Yin, J., M. E. Schlesinger, and R. J. Stouffer, 2009: Model projections of rapid sea-level rise on the northeast coast of the united states. *Nature Geoscience*, **2** (4), 262–266.

Zelinka, M. D., T. Andrews, P. M. Forster, and K. E. Taylor, 2014: Quantifying components of aerosol-cloud-radiation interactions in climate models. *Journal of Geophysical Research: Atmospheres*, **119** (12), 7599–7615, <https://doi.org/https://doi.org/10.1002/2014JD021710>.

Zhang, J., and Coauthors, 2021: The role of anthropogenic aerosols in the anomalous cooling from 1960 to 1990 in the cmip6 earth system models. *Atmospheric Chemistry and Physics*, **21** (24), 18 609–18 627, <https://doi.org/10.5194/acp-21-18609-2021>, URL <https://acp.copernicus.org/articles/21/18609/2021/>.

Zhang, R., R. Sutton, G. Danabasoglu, Y.-O. Kwon, R. Marsh, S. G. Yeager, D. E. Amrhein, and C. M. Little, 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/https://doi.org/10.1029/2019RG000644>.

Zhang, R., and M. Thomas, 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/https://doi.org/10.1038/s43247-021-00182-y>.

Zhu, C., and Z. Liu, 2020: Weakening atlantic overturning circulation causes south atlantic salinity pile-up. *Nature Climate Change*, **10** (11), 998–1003, <https://doi.org/https://doi.org/10.1038/s41558-020-0897-7>.