Supporting Information for

**Quantifying contributions of internal variability and external forcing to Atlantic multidecadal variability since 1870**

Minhua Qin, Aiguo Dai, and Wenjian Hua

Correspondence to: [adai@albany.edu](mailto:adai@albany.edu) (A.D.), [wenjian@nuist.edu.cn](mailto:wenjian@nuist.edu.cn) (W.H.)

**Contents of this file**

Supplementary Text

Figures S1 to S10

Tables S1 to S2

**Supplementary Text**

***Methods***

To derive the multi-model ensemble mean (MMM), we first averaged the ensemble runs to derive the ensemble mean for each model and then averaged the ensemble mean over all the models with the same weight for each model. The results are similar if all the ensemble runs from all the models were weighted equally without averaging for each model first.

***S1. Estimates of the forced and internal components***

An implicit assumption in this study is that externally-forced (EX) responses and internally-generated (IV) variations are linearly separable in NASST. To separate the internally-generated and externally-forced variations in observed SSTs, we used the sum of rescaled global-mean SST (GMSST, exclude the area above 60°N and below 60°S because of low SST data coverage) time series from the MMM of individual foricngs (i.e., GHG, NAT and AA) as the first-order estimate of the externally-forced component (e.g., regressed it against the observed GMSST via multiple linear regression). And we removed this forced signal through linear regression from the observed SSTs at each grid point to produce the residual SST fields that contain primarily internal component. This method can remove model biases in response to individual forcings. Dai & Bloecker (2019) suggested that using the GMSST time series from the MMM simulations is a preferred way to define the forced signal, which has the same temporal evolution over all grid boxes because it is determined by the external forcing series. Following Dai et al. (2015), we also tested the use of GMSST of all-forcing MMM (with 321 runs from 27 CMIP6 models, Table S2) as the estimate of the forced signal (Figure S1b). There is a difference between these two methods, in particular since the 1990s (Figure S1). Note that the CMIP5 or CMIP6 models may overestimate the SST response to aerosol changes (Sato et al., 2018; Toll et al., 2019; Chylek et al., 2020), making the estimated amplitude of the aerosol-forced signal biased. To account for different rescaling factors in response to individual forcings, we used the re-scaling through multiple linear regression (see Methods) to remove any systematic biases in model simulated responses.

We also estimated the internally-generated and externally-forced GMSST variations in observations. The IV-induced multidecadal GMSST variations in observations show a negative anomaly since the early 2000s but most positively anomalies from about 1933-1970, whereas EX-induced GMSST variations (relative to the forced long-term trend) exhibit a positive anomaly since about 1970 (Figure S10). The apparent anti-correlation between the IV-induced and externally-forced GMSST multidecadal variations around the late 1990s is largely due to the phase change of the Interdecadal Pacific Oscillation (IPO, Dai et al., 2015) or Pacific Multidecadal Oscillation (PMO, Steinman et al., 2015) that have greatly contributed to the IV-induce variations. The AMV has only a secondary contribution to the global-mean temperature variations (Dai et al., 2015).

***S2. Definition of the AMV index***

Many different methods have been used to define the index and spatial pattern associated with AMV or AMO. A common method is to use the 10-year smoothed, area-weighted average of locally linearly detrended SSTs over the North Atlantic to define the AMO index (Enfield et al., 2001). This follows the generic definition of oscillations in a time series; thus, it provides a good measure of the total multidecadal variations in NASST from both internal dynamics and external forcing. However, the externally-forced long-term warming signal is nonlinear and thus the linear detrending would not accurately remove the externally-forced component from NASST (Yan et al., 2019). Trenberth & Shea (2006) proposed a revised AMO index by subtracting GMSST directly from NASST (without re-scaling through regression). Sutton and Dong (2012) defined the AMO index by removing the low-pass filtered GMSST (with the North Atlantic being excluded) from NASST. Similarly, Mohino et al. (2011) used the low-pass filtered time series of averaged SST between 45°S and 60°N to represent the global warming signal and removed it from NASST before computing the EOFs to derive an AMO index. Ting et al. (2009) extracted the externally-forced signal by applying a signal-to-noise (S/N) maximizing EOF analysis of the global surface temperature fields from individual model simulations, and found that the S/N-maximizing leading principal component and GMSST both can effectively isolate the global warming signal in observations.

Recently, some studies (Sutton et al., 2018; Yan et al., 2019; Zhang et al., 2019) suggest that AMV should be used to describe the multidecadal variability unique in its coherent multivariate feature and they removed GMSST and consider the residual as unique to the North Atlantic from a multivariate perspective. It does not provide any attribution information. Yan et al. (2019) used a multivariate AMV index (MAI) to define the AMV by combining the associated variations in NASST, surface salinity, upper (top 700m) ocean heat and salt content. They suggested that the observed AMV, after removal of the signal associated with GMSST, is not dominated by external forcing and is unique to the Atlantic and is linked to multidecadal variations of the AMOC. This conclusion is consistent with our findings regarding the unforced internal component.

Kim et al. (2018) suggested that the observed AMV anomaly pattern during 1996-2005 is dominated by internal variability, not by externally-forced variations based on CESM large ensemble simulations. However, the quantitative contributions from internal variability and external forcing since the 1990s remain unknown. We further decomposed the observed NASST into a component caused by internal variability and another caused by external forcing. The well-known horseshoe pattern in the observed NASST from 1996-2005 is indeed caused by internal variability (e.g., AMOC), as the IV-induced SST anomaly patterns roughly match the observed SST anomaly patterns (not shown). However, external forcing (including GHGs and aerosols) is in phase with and amplifies the IV-induced SST anomaly, although the externally-forced SST anomaly pattern differs from the observed or IV-induced patterns in the North Atlantic. Furthermore, our results are consistent with previous study (Hua et al., 2019) who showed the CMIP5 MMM of all-forcing runs shows a weak AMV in phase with the observed AMV. Yan et al. (2019) also showed a weak but non-negligible AMV component from the CMIP5 MMM of all-forcing runs (after linear detrending over 1945-2012, see their Fig. 2j). Thus, these earlier results based on CMIP5 MMM support our conclusion that the observed AMV over the past ~150 years is mainly IV-driven but with significant enhancements by volcanic and anthropogenic aerosols, especially for the recent period since the 1930s.

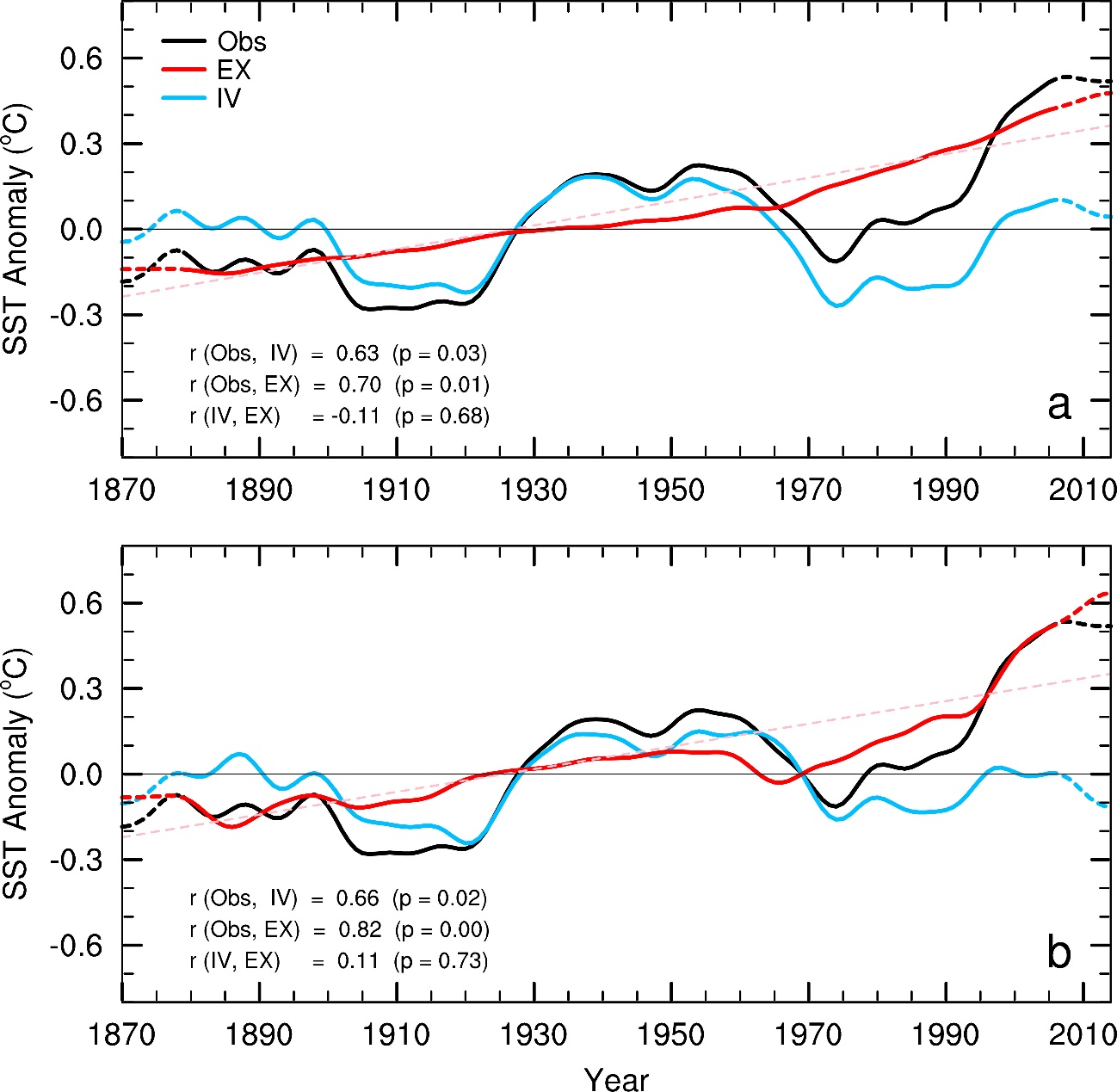
***S3. Oceanic response to GHGs***

SST and surface ocean and wind circulation change patterns similar to that due to decreasing aerosol forcing are seen in coupled model simulations with increasing GHGs (Chemke et al., 2020; Figure S8), and the strong cooling (warming) along the U.S. Northeast coast (Figure. S7a, d) is a common response in climate models to an AMOC strengthening (weakening), which is due to the interaction of AMOC-induced changes in the deep western boundary current (DWBC) and bottom topography (Zhang & Vallis, 2007; Zhang, 2008; Zhang et al., 2011; Caesar et al., 2018). Such a dipole pattern (i.e., opposite signs between the subpolar gyre and the Gulf Stream path) is a distinctive fingerprint of AMOC variability (Zhang, 2008). Bottom vortex stretching induced by a downslope DWBC near the south of the Grand Banks leads to the formation of a cyclonic northern recirculation gyre and keeps the path of the depth-integrated western boundary current downstream of Cape Hatteras separated from the North American coast (Zhang & Vallis, 2007).

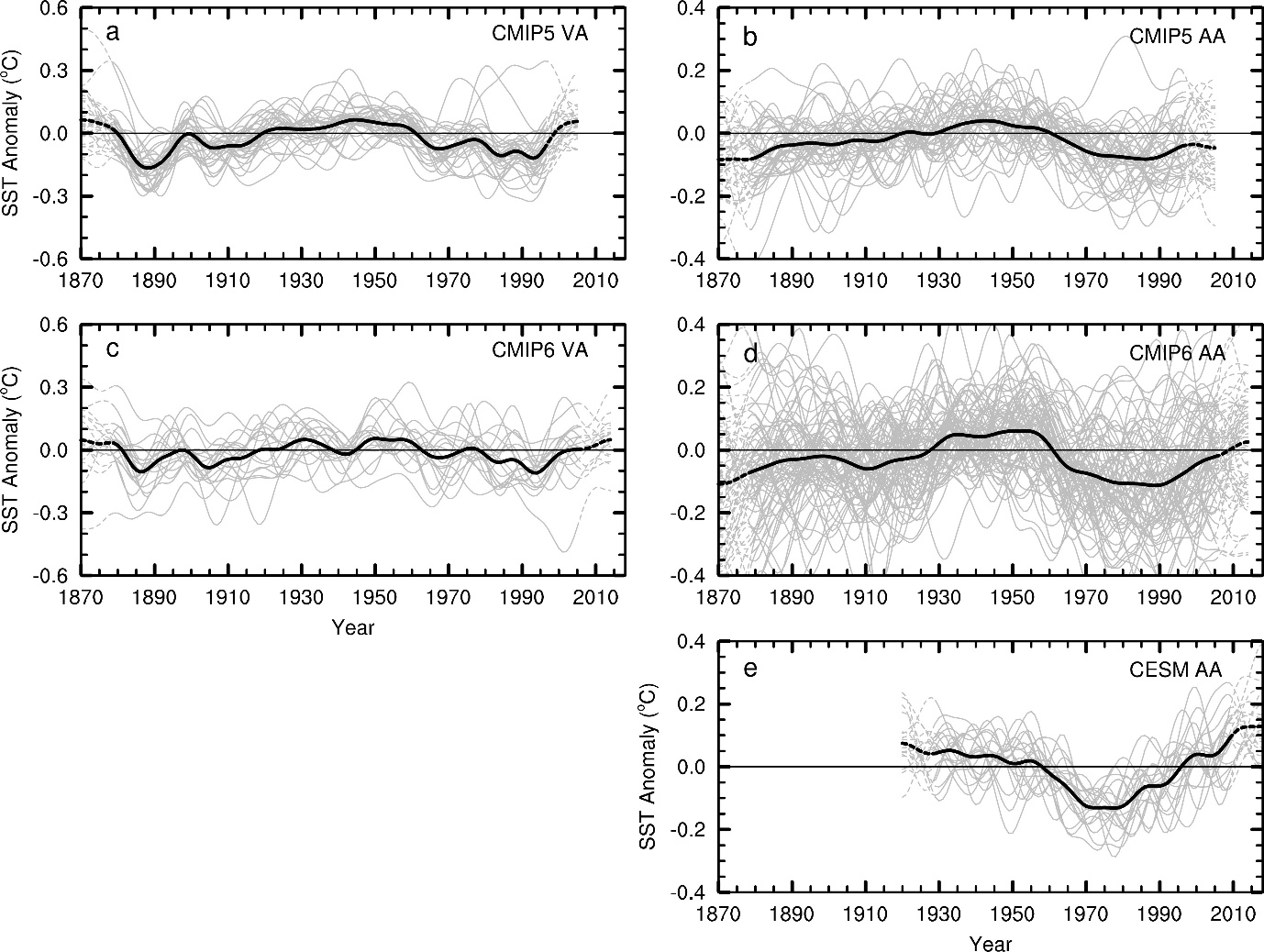
***S4. Comparison of simulated externally-forced AMOC changes with AMOC-related observations***

Note that the externally-forced AMOC changes in the historical simulations have opposite phases with AMOC-related changes in observations (e.g., Yan et al., 2019). For example, the AMOC changes from historical simulations (i.e. strengthening of AMOC during 1950-1980 and weakening of AMOC after 1980, Figures S7 and S8) are opposite to the observational-based AMOC proxies (e.g. weakening of AMOC during 1950-1980 and strengthening of AMOC over 1990-2005, see their Figure 3a in Chen & Tung, 2018). Furthermore, direct observations from the Labrador Sea western boundary sections suggest a strengthening of the AMOC outflow from the late 1990s to the early 2000s, opposite to the decline found in the GHG-forced and AA-forced simulations (Figures S7 and S8). These discrepancies suggest that the observed multidecadal AMOC changes are dominated by internal variability, while external forcing (e.g., GHG and aerosols) may play a role, especially since the 1990s.

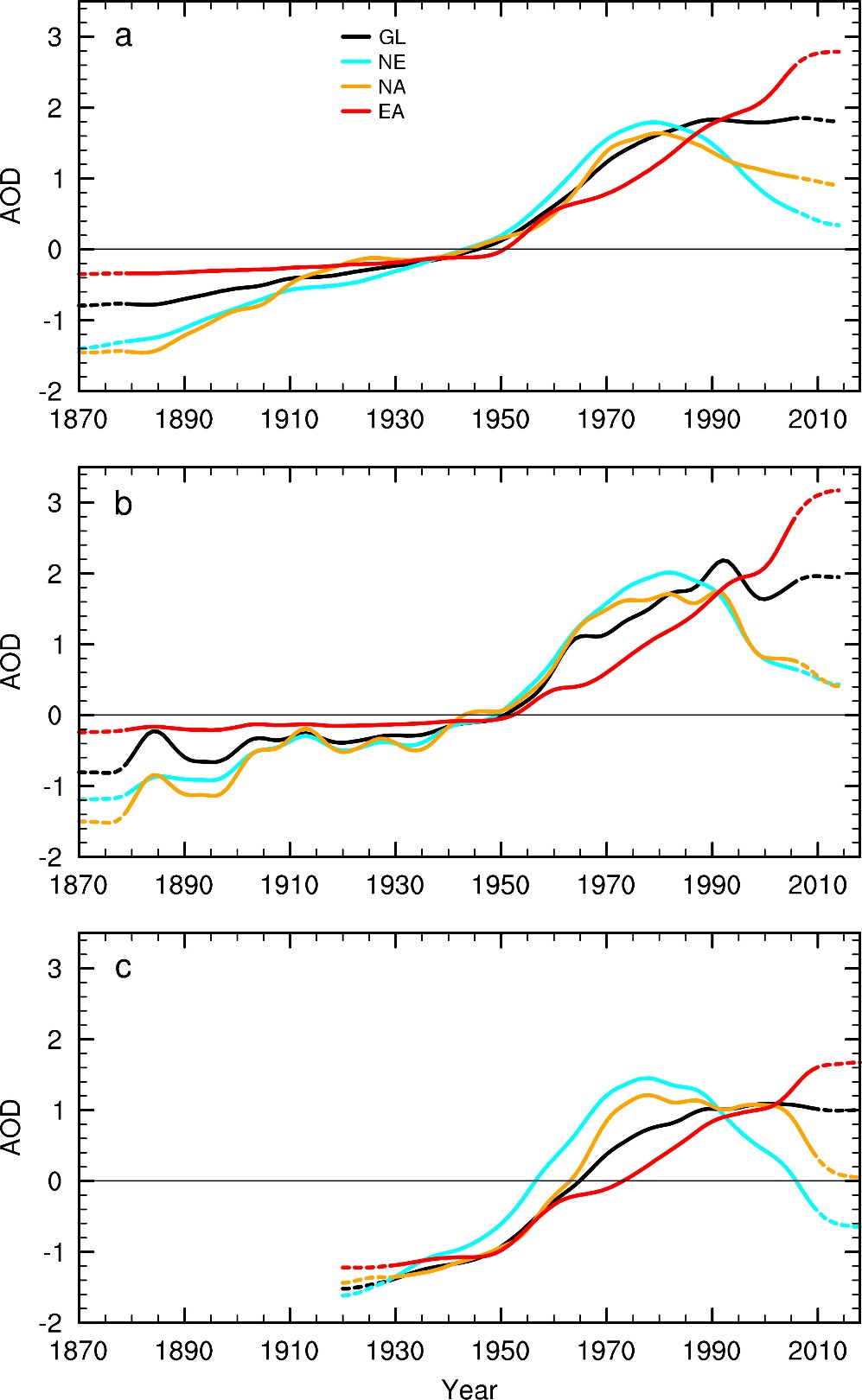
In addition, the magnitudes of IV-induced SST variations during the recent positive AMV phase (1997-2005) is smaller than that during the negative AMV phase (1966-1996) (see blue lines in Figure 2a and S1a). The weak IV-induced SST signal is inconsistent with the recent strengthening of AMOC changes since 1990 (Chen & Tung, 2018), the strong internally induced AMV-related subpolar North Atlantic sea surface salinity (SSS) signal since 1990 (Yan et al. 2019), and surface turbulent heat flux signal during 1990-2005 (Kim et al. 2018). Therefore, the magnitudes of externally-forced positive AMV shift since 1990 may exist uncertainties. One of the causes is that current climate models overestimate the responses to volcanic and anthropogenic aerosols (Chylek et al., 2020; Sato et al., 2018; Toll et al., 2019).



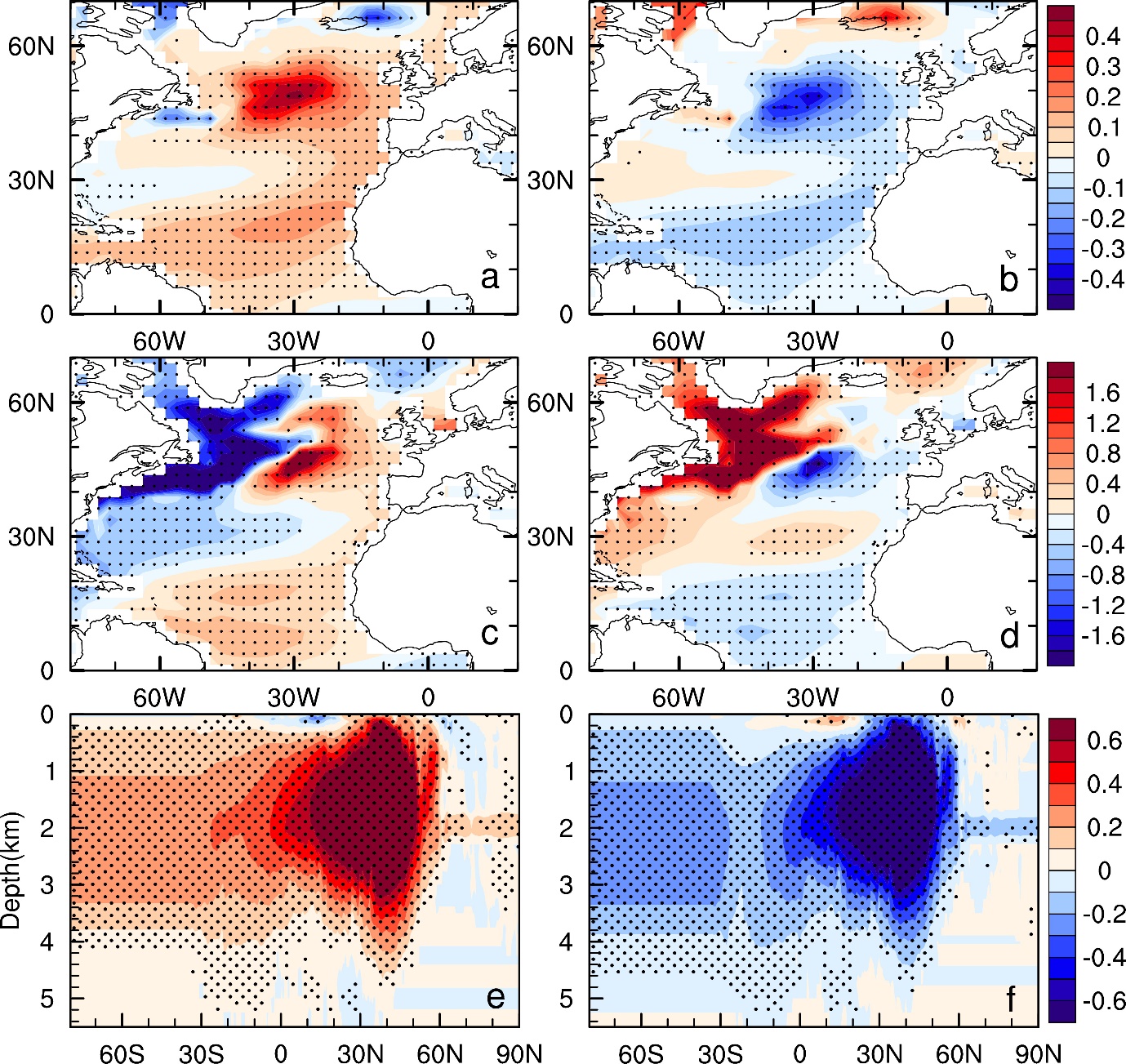
**Figure S1. Forced and internal decadal NASST variations.** (**a**) Time series of the low-pass filtered, annual-mean SST anomalies (ºC, relative to 1901-1970 mean) averaged over the North Atlantic (80°W–0°W, 0°–60°N) from observations (HadISST1, Obs, black line) and the estimated externally-forced (EX, red) and internally-generated (IV, blue) components in the observations from 1870-2014. The data near the two ends are derived with mirrored data in the filtering and thus are less reliable, they are marked by the dashed lines. We used the global-mean SST (GMSST) time series from CMIP6 multi-model ensemble mean (MMM) of individual forcing (i.e., GHG, NAT, AA) simulations as the estimate of externally forced signal. We regressed the GMSST MMM of individual forcings against the observed GMSST time series via multiple linear regression to obtain the forced signal and removed it from the observed SSTs at each grid point to produce the residual SST fields that contain primarily internal component. The thin red dashed line represents the linear trend of the EX component. The correlation coefficients (r) between the Obs and IV or EX, together with the attained significance (p), are also shown. (**b**) Same as (a), but using the GMSST time series from CMIP6 MMM of all-forcing historical simulations as the externally-forced signal and regressed it against local observed SST to estimate the EX component. The residual from the observation was used to estimate the IV component.



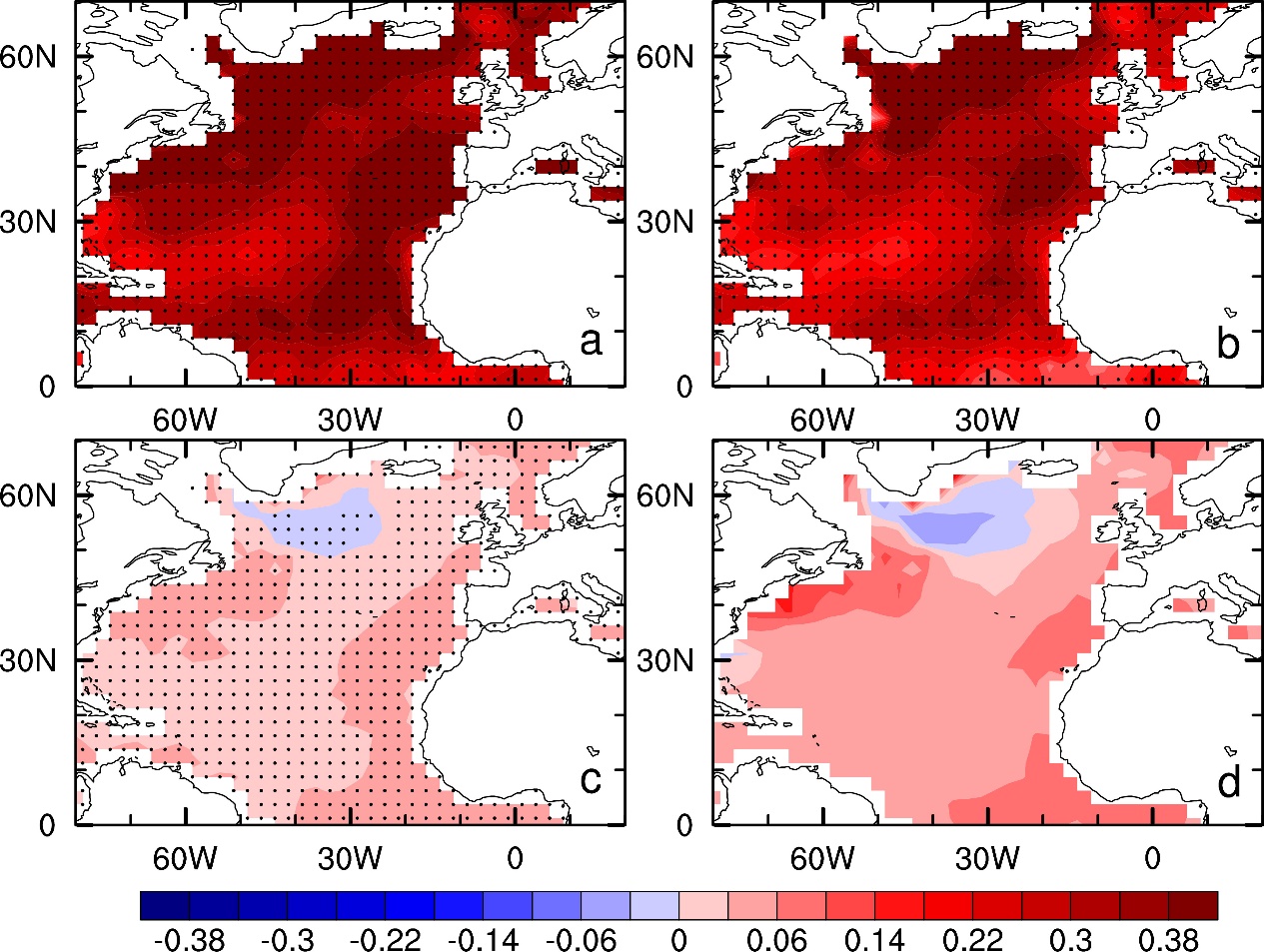
**Figure S2**. **Regionally-averaged time series of the low-pass filtered SST anomalies.** SST anomalies (°C, relative to 1901-1970 mean) averaged over the North Atlantic (80°W–0°W, 0°–60°N) from the CMIP5 single forcing only simulations from 1870-2005 for (**a**) VA and (**b**) AA forcing. The model simulations were linearly detrended before calculating the anomalies. The black lines represent the ensemble mean of the individual simulations represented by the grey lines. The data near the two ends are derived with mirrored data in the filtering and thus are less reliable, they are marked by the dashed lines. (**c-d**) Same as a-b, but for the SST anomalies (°C) from CMIP6 single forcing-only simulations from 1870-2014. (**e**) Same as (b), but for the SST anomalies (°C, relative to 1920-2018 mean) from CESM1 AER simulations from 1920-2018.



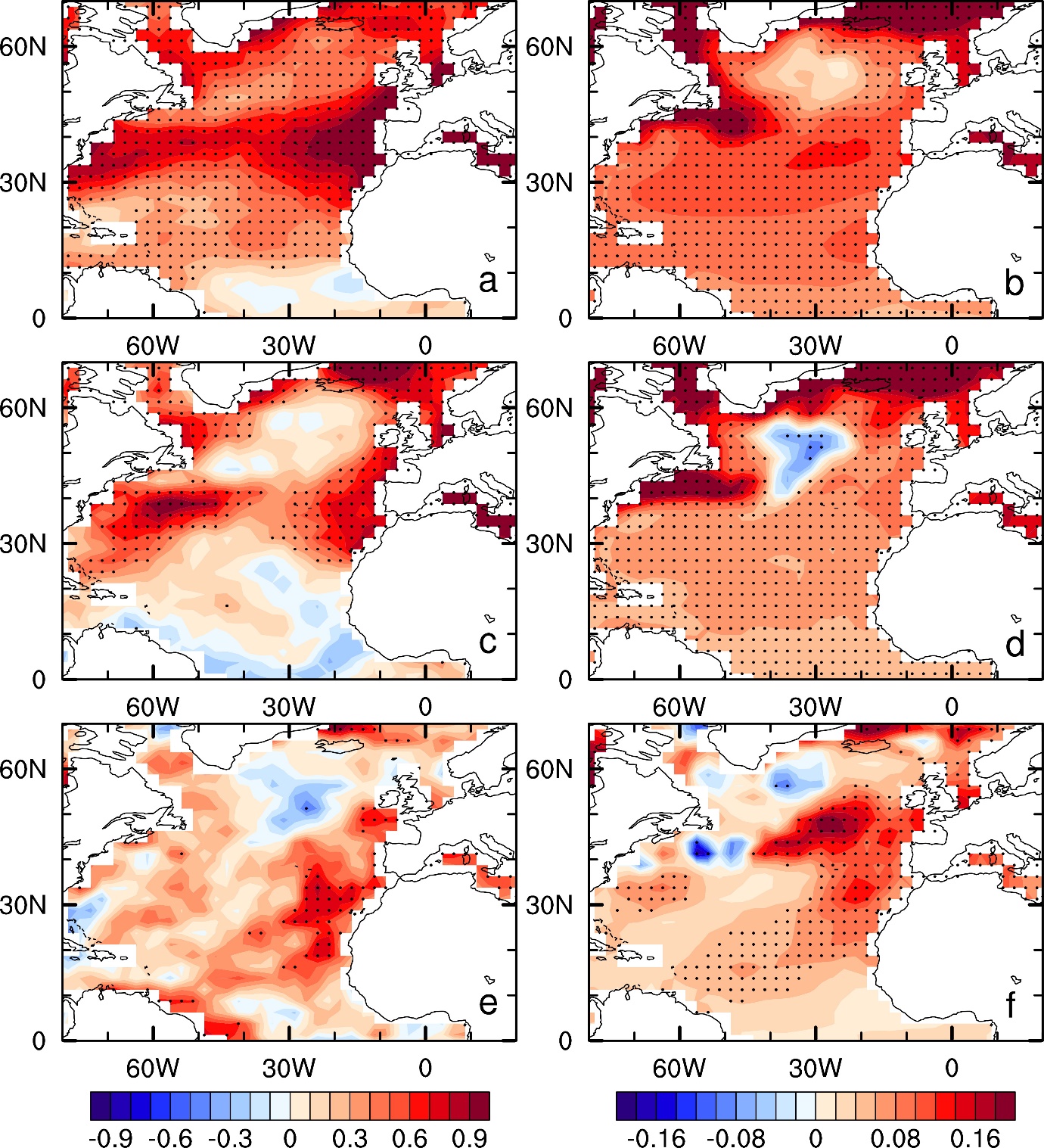
**Figure S3. Time series of anthropogenic aerosol loading.** Regionally-averaged, normalized anomaly (relative to 1901-1970 for (a) and (b), 1920-2018 for (c)) time series of the smoothed ambient aerosol optical depth (AOD) at λ = 550 nm averaged over the near-globe (GL: 60°S–60°N, black), North America and Europe (NE: 20°N–65°N, 130°W–70°E, land only, blue), North Atlantic (NA: 0°N–60°N, 80°W–0°W, yellow) and East Asia (EA: 20°N–40°N, 100°E–160°E, red) derived from the (**a**) CMIP5 and (**b**) CMIP6 simulations from 1870-2014, and (**c**) CESM1 simulations from 1920-2018.



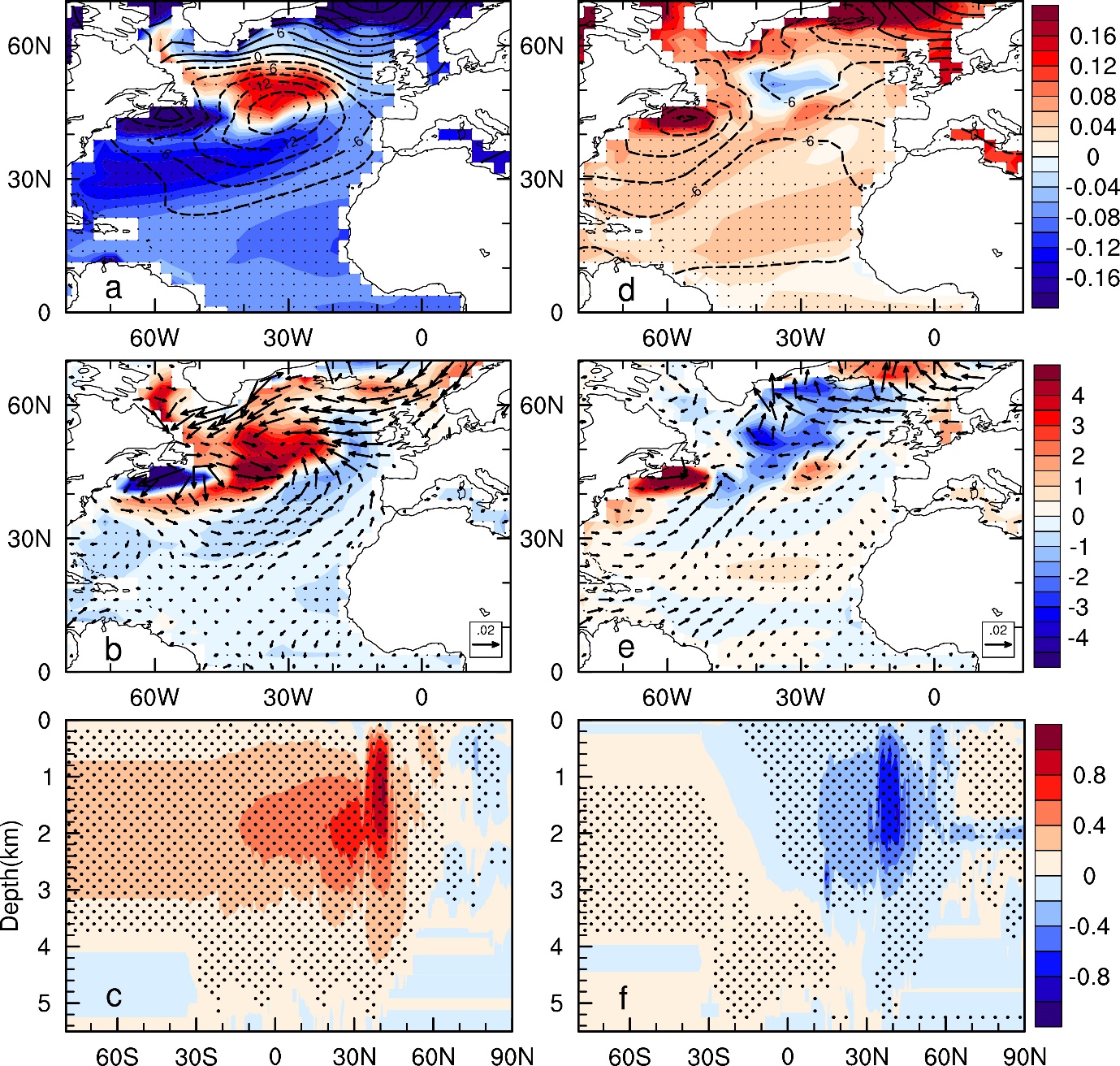
**Figure S4. Differences in spatial response between warm and cold AMOC phases.** Composite (**a-b**) SST (°C), (**c-d**) SSH (cm) and (**e-f**) Atlantic meridional overturning circulation stream-function (Sv) anomaly patterns during the positive (left column) and negative (right column) phases of the AMOC in the CESM1 preindustrial control run. The different AMOC phases are defined as the periods with the AMOC index larger or lower than one standard deviation. Stippled regions indicate the anomalies are statistically significant at the 5% level based on a Student's t-test.



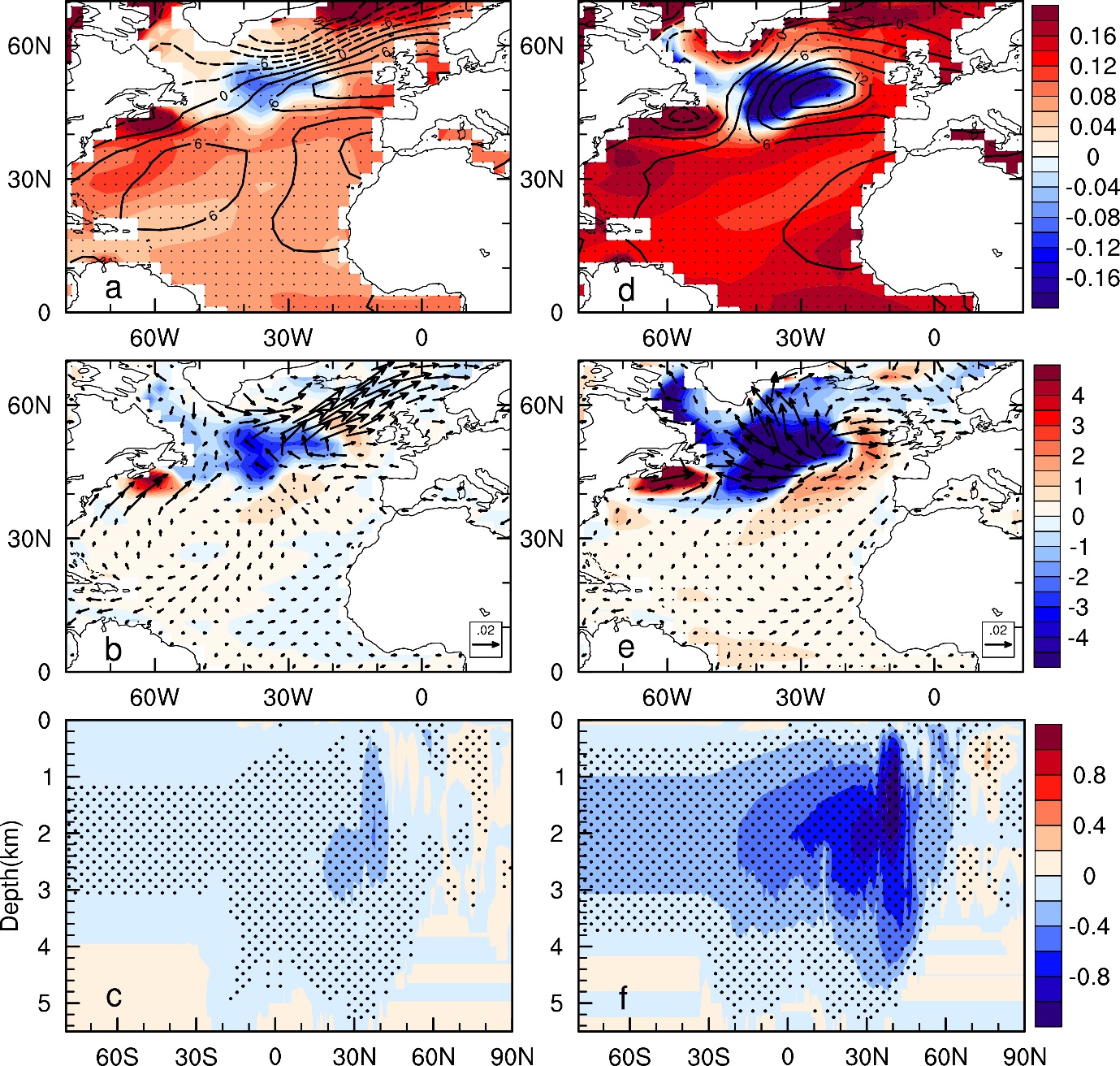
**Figure S5. Contributions of internal variability and external forcing.** Annual-mean SST anomalies (oC) for the warm periods (1927-1965 and 1997-2005) relative to the cold periods (1900-1926 and 1966-1996) based on the IV-induced AMV index (blue line in Figure 2a) from (**a**) observations, (**b**) IV-induced variations, (**c**) non-GHG forced and (**d**) GHG-forced changes. To derive the IV-induced variations, we removed the externally-forced changes and variations in observed SSTs (see Methods). The differences between (a) and (b) denotes the externally-forced (EX) variations. We regressed the local observed SST against the GMSST from the GHG-only MMM to estimate the GHG-forced component in observations (shown in d). The differences between the EX and GHG-forced variations show the non-GHG forced signal (shown in c). **The aerosol forcing enlarges the IV-induced AMV amplitude over most of the North Atlantic, especially along the eastern part, but with little influence over the subpolar region (c). The GHG forcing also enhances the IV-induced AMV amplitude over most of the North Atlantic, except the inner subpolar region where it damps the SST anomalies (d).**



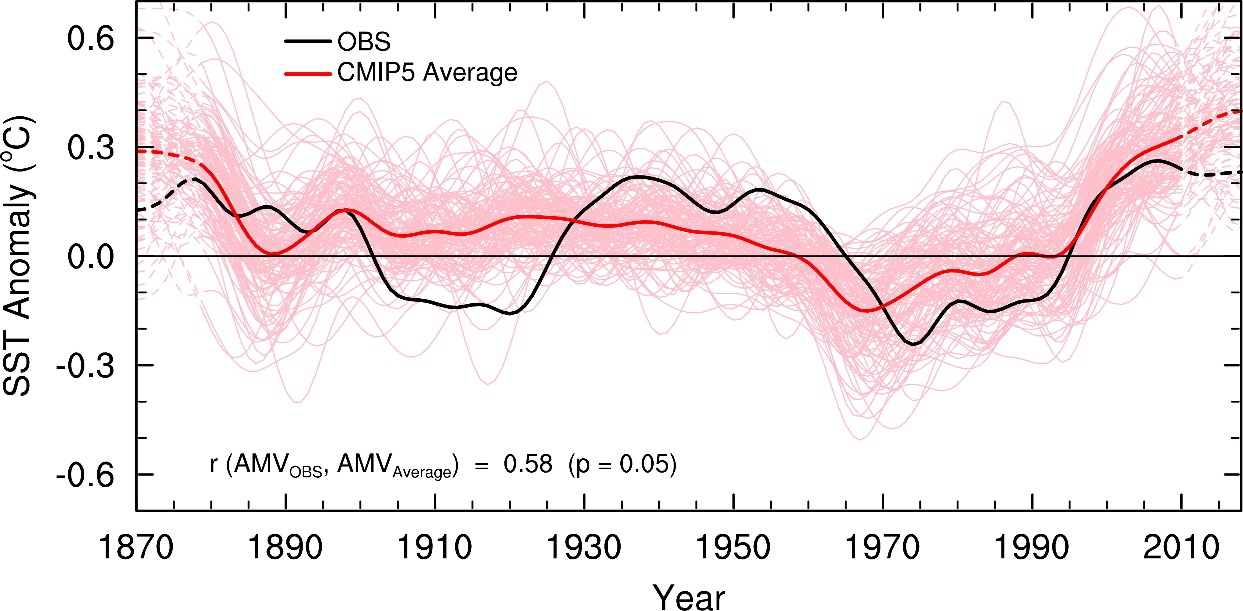
**Figure S6. Analysis of SST and surface solar radiation patterns in models.** Annual-mean anomalies for the warm period (1927-1965 and 1997-2005) relative to the cold period (1900-1926 and 1966-1996) based on the IV-induced AMO index (blue line in Figure 2a) for (**a**) surface net shortwave radiation (W m-2, downward positive), and (**b**) SST (ºC) from the CMIP6 multi-model ensemble mean (MMM) of the all-forcing simulations. (**c-d**) Same as (a-b), but for the CMIP6 AA forcing only simulations. (**e-f**) Same as (a-b), but for CMIP6 piAA simulations. The model simulations were linearly detrended before calculating the anomalies to highlight the decadal variability. The stippling indicates that the anomalies are statistically significant at the 5% level based on a Student's t-test.



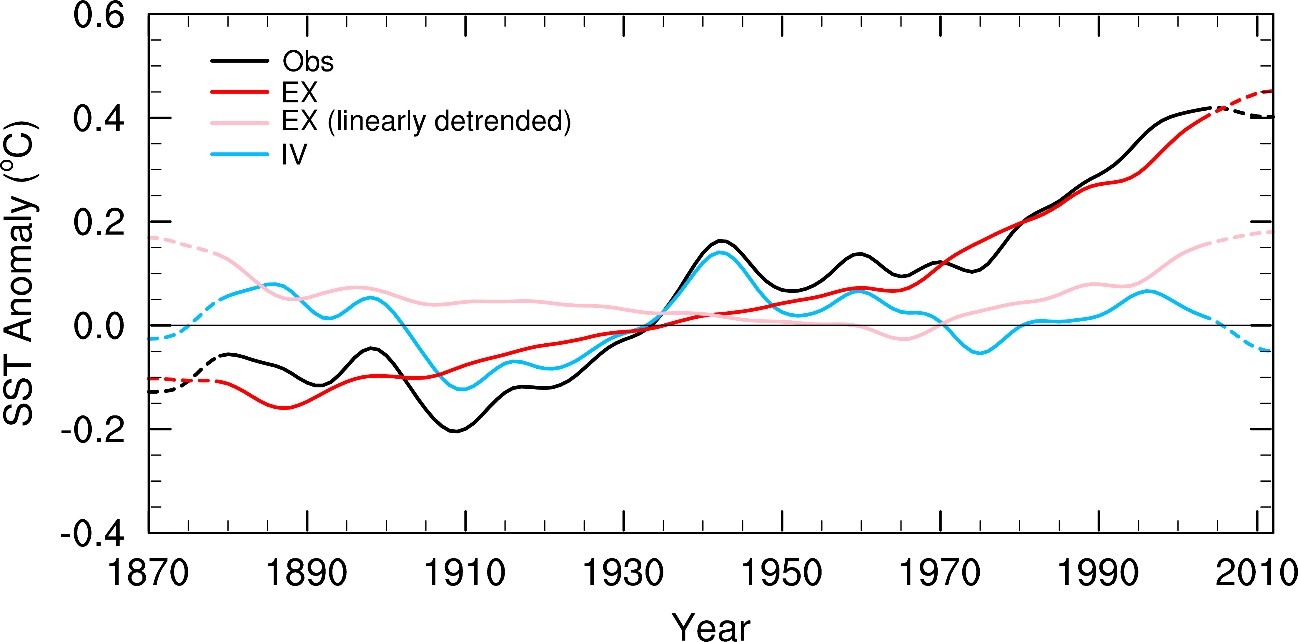
**Figure S7. North Atlantic responses to industrial aerosol forcing from CESM1 AER ensemble of simulations.** Shown are linear trends for (**a**) SST (shading, ºC decade-1) and sea level pressure (Pa decade-1, contours, dashed lines for negative values), (**b)** wind stress (vectors, dyn cm-2 decade-1) and surface heat flux (shading, W m-2 decade-1), and (**c**) Atlantic meridional overturning circulation stream-function (Sv decade-1) during 1950-1980 when North Atlantic aerosol loading was increasing. (**d-f**) Same as (a-c) but during 1980-2014 when North Atlantic aerosol loading was decreasing. The stippling indicates that the anomalies are statistically significant at the 5% level based on a Student's t-test.



**Figure S8. North Atlantic responses to GHG forcing from CESM1 GHG ensemble of simulations**. Shown are linear trends for (**a**) SST (shading, ºC decade-1) and sea level pressure (Pa decade-1, contours, dashed lines for negative values), (**b)** wind stress (vectors, dyn cm-2 decade-1) and surface heat flux (shading, W m-2 decade-1), and (**c**) Atlantic meridional overturning circulation stream-function (Sv decade-1) during 1950-1980 (with relatively slow warming). (**d-f**) Same as (a-c) but during 1980-2014 (with relatively rapid warming). The stippling indicates that the anomalies are statistically significant at the 5% level based on a Student's t-test.



**Figure S9. Regionally-averaged time series of SST anomalies in observations and models.** Time series of the local linearly-detrended, low-pass filtered, annual-mean SST anomalies (ºC, relative to 1901-1970 mean) averaged over the North Atlantic (80°W–0°W, 0°–60°N) from observations (HadISST1, Obs, black line) and individual CMIP5 model runs (121 runs in total, pink lines) from 1870-2018. The thick red line shows the ensemble mean of the individual runs. The data near the two ends are derived with mirrored data in the filtering and thus are less reliable, they are marked by the dashed lines.



**Figure S10. Regionally-averaged time series of GMSST anomalies.** Time series of low-pass filtered, annual-mean global (60°S–60°N) SST anomalies (ºC, relative to the mean of 1901-1970) for the observations from 1870-2012 (Obs, black line), estimated externally forced signal (EX, red) and internal variability (IV, blue) in the observations. The data near the two ends are derived with mirrored data in the filtering and thus are less reliable, they are marked by the dashed lines. We rescaled the CMIP5 multi-model ensemble mean (MMM) of the global-mean SST time series from the individual forcing (i.e., GHG, NAT, AA) simulations to the observed long-term change via multiple linear regression as the estimate of externally forced signal, and the difference between the observations and the external signal is used as the IV component in the observations. The pink line shows the multidecadal variations in EX after linearly detrending it.

**Table S1. The number of ensemble simulations for the 38 CMIP5 models used in this study.** Historical plus RCP4.5 runs were used the all-forcing (ALL) experiment. The single forcing historical runs are: GHG = greenhouse gases, NAT = natural forcing (solar plus volcanoes), VA = volcanic aerosols, SI = solar irradiance forcing, AA = anthropogenic aerosols.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Model name | ALL | GHG | NAT | VA | SI | AA |
| 1 | ACCESS1.0 | 1 |  |  |  |  |  |
| 2 | ACCESS1.3 | 1 |  |  |  |  |  |
| 3 | bcc-csm1.1 | 1 | 1 | 1 |  |  |  |
| 4 | bcc-csm1.1(m) | 1 |  |  |  |  |  |
| 5 | BNU-ESM | 1 | 1 | 1 |  |  |  |
| 6 | CanESM2 | 5 | 5 | 5 |  | 5 | 5 |
| 7 | CCSM4 | 6 | 3 | 4 | 3 | 3 | 3 |
| 8 | CESM1-BGC | 1 |  |  |  |  |  |
| 9 | CESM1-CAM5 | 3 | 1 | 3 | 3 | 2 | 3 |
| 10 | CMCC-CM | 1 |  |  |  |  |  |
| 11 | CMCC-CMS | 1 |  |  |  |  |  |
| 12 | CNRM-CM5 | 1 | 6 | 6 |  |  |  |
| 13 | CSIRO-Mk3.6.0 | 10 | 5 | 5 | 5 |  | 5 |
| 14 | FIO-ESM | 3 |  |  |  |  |  |
| 15 | GFDL-CM2p1 | 10 |  |  |  |  |  |
| 16 | GFDL-CM3 | 1 | 3 | 3 |  |  | 3 |
| 17 | GFDL-ESM2G | 1 |  |  |  |  |  |
| 18 | GFDL-ESM2M | 1 | 1 | 1 | 1 | 1 | 1 |
| 19 | GISS-E2-H-CC | 1 |  |  |  |  |  |
| 20 | GISS-E2-H | 16 | 5 | 5 | 10 | 10 | 10 |
| 21 | GISS-E2-R-CC | 1 |  |  |  |  |  |
| 22 | GISS-E2-R | 17 | 5 | 5 | 10 | 10 | 10 |
| 23 | HadCM3 | 10 |  |  |  |  |  |
| 24 | HadGEM2-CC | 1 |  |  |  |  |  |
| 25 | HadGEM2-ES | 4 | 4 | 4 |  |  |  |
| 26 | HadGEM2-AO | 1 |  |  |  |  |  |
| 27 | inmcm4 | 1 |  |  |  |  |  |
| 28 | IPSL-CM5A-LR | 4 | 3 | 3 |  |  | 1 |
| 29 | IPSL-CM5A-MR | 1 | 3 | 3 |  |  |  |
| 30 | IPSL-CM5B-LR | 1 |  |  |  |  |  |
| 31 | MIROC5 | 3 |  |  |  |  |  |
| 32 | MIROC-ESM-CHEM | 1 | 1 | 1 |  |  |  |
| 33 | MIROC-ESM | 1 | 3 | 3 |  |  |  |
| 34 | MPI-ESM-LR | 3 |  |  |  |  |  |
| 35 | MPI-ESM-MR | 3 |  |  |  |  |  |
| 36 | MRI-CGCM3 | 1 | 1 | 1 |  |  |  |
| 37 | NorESM1-ME | 1 |  |  |  |  |  |
| 38 | NorESM1-M | 1 | 1 | 1 |  |  | 1 |
|  | Total number | 121 | 52 | 55 | 32 | 31 | 42 |

**Table S2. The number of ensemble simulations for the 27 CMIP6 models used in this study.** ALL for historical (1850-2014) all-forcing runs, GHG for greenhouse gas-only runs, VA for volcanic aerosols only runs, AA for anthropogenic aerosols only runs, and piAA for runs with historical all forcing except AA fixed to the preindustrial level.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Model name | ALL | GHG | NAT | VA | AA | piAA |
| 1 | BCC-CSM2-MR | 3 | 3 | 3 |  | 3 |  |
| 2 | BCC-ESM1 | 3 |  |  |  |  |  |
| 3 | CanESM5 | 50 | 50 | 30 | 10 | 30 |  |
| 4 | CESM2 | 11 | 3 | 3 |  | 1 |  |
| 5 | CESM2-WACCM | 3 |  |  |  |  |  |
| 6 | CNRM-CM6-1 | 29 | 10 | 10 |  | 10 |  |
| 7 | CNRM-ESM2-1 | 9 |  |  |  |  |  |
| 8 | E3SM-1-0 | 5 |  |  |  |  |  |
| 9 | EC-Earth3-Veg | 4 |  |  |  |  |  |
| 10 | FGOALS-f3-L | 3 |  |  |  |  |  |
| 11 | FGOALS-g3 | 3 |  |  |  |  |  |
| 12 | GFDL-ESM4 | 3 | 1 | 3 |  | 1 | 1 |
| 13 | GISS-E2-1-G | 30 | 5 | 5 | 5 | 5 |  |
| 14 | GISS-E2-1-H | 23 |  |  |  |  |  |
| 15 | HadGEM3-GC31-LL | 4 | 4 | 4 |  | 4 |  |
| 16 | INM-CM5-0 | 10 |  |  |  |  |  |
| 17 | IPSL-CM6A-LR | 32 | 10 | 10 |  | 10 |  |
| 18 | MIROC-ES2L | 3 |  |  |  |  |  |
| 19 | MIROC6 | 10 | 3 | 3 | 3 | 3 | 3 |
| 20 | MPI-ESM-1-2-HAM | 2 |  |  |  |  | 1 |
| 21 | MPI-ESM1-2-HR | 10 |  |  |  |  |  |
| 22 | MPI-ESM1-2-LR | 10 |  |  |  |  |  |
| 23 | MRI-ESM2-0 | 6 | 3 | 3 |  | 3 |  |
| 24 | NESM3 | 5 |  |  |  |  |  |
| 25 | NorCPM1 | 30 |  |  |  |  |  |
| 26 | NorESM2-LM | 3 | 3 | 3 |  | 3 | 3 |
| 27 | UKESM1-0-LL | 17 |  |  |  |  | 3 |
|  | Total number | 321 | 95 | 77 | 18 | 73 | 11 |