The Evolving Role of External Forcing in North Atlantic SST Variability over the Last Millennium

JEREMY M. KLAVANS,[a,b] AMY C. CLEMENT,[b] MARK A. CANE,[c] AND LISA N. MURPHY[b]

[a] *Department of Atmospheric and Oceanic Science, University of Colorado Boulder, Boulder, Colorado*
[b] *Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida*
[c] *Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York*

(Manuscript received 27 April 2021, in final form 3 January 2022)

ABSTRACT: Atlantic multidecadal variability (AMV) impacts temperature, precipitation, and extreme events on both sides of the Atlantic Ocean basin. Previous studies with climate models have suggested that when external radiative forcing is held constant, the large-scale ocean and atmosphere circulation are associated with sea surface temperature (SST) anomalies that have similar characteristics to the observed AMV. However, there is an active debate as to whether these internal fluctuations driven by coupled atmosphere–ocean variability remain influential to the AMV on multidecadal time scales in our modern, anthropogenically forced climate. Here we provide evidence from multiple large ensembles of climate models, paleoreconstructions, and instrumental observations of a growing role for external forcing in the AMV. Prior to 1850, external forcing, primarily from volcanoes, explains about one-third of AMV variance. Between 1850 and 1950, there is a transitional period, where external forcing explains one-half of AMV variance, but volcanic forcing only accounts for about 10% of that. After 1950, external forcing explains three-quarters of AMV variance. That is, the role for external forcing in the AMV grows as the variations in external forcing grow, even if the forcing is from different sources. When forcing is relatively stable, as in earlier modeling studies, a higher percentage of AMV variations are internally generated.

KEYWORDS: North Atlantic Ocean; Climate variability; Multidecadal variability

1. Introduction

Our modern interest in Atlantic Ocean multidecadal sea surface temperature (SST) variability was born in the 1980s of an urgent need to explain the causes of severe droughts and associated famines in the Sahel (Folland et al. 1986, 1991; Hastenrath 1984; Lamb 1978; Lough 1986). Over the last three-plus decades, research shows that in addition to Sahel rainfall, Atlantic SSTs covary with North American and Brazilian precipitation (Enfield et al. 2001; Knight et al. 2006; Ruprich-Robert et al. 2017; Sutton and Hodson 2005), Atlantic hurricanes (Goldenberg et al. 2001), and European temperatures (Sutton and Hodson 2005). The outsized impact of this SST variability earned it a name: Atlantic multidecadal variability (AMV). Successful prediction of these impacts requires that we identify and quantify the contributors to the AMV [as in Qin et al. (2020)].

Zhang et al. (2019)'s recent review on the link between the Atlantic meridional overturning circulation (AMOC) and the AMV articulates the case for the ocean circulation. The basic argument, which has roots in Bjerknes (1964), states that multidecadal changes in Atlantic SSTs are a result of changes in the ocean thermohaline circulation. In contrast, interannual variations in Atlantic SSTs are the result of wind-driven changes in turbulent heat fluxes. Gulev et al. (2013) use the sign of the relationship between observed turbulent heat fluxes and observed SSTs to argue that ocean heat convergence drive SST changes at multidecadal time scales, whereas winds drive SSTs on interannual time scales, although Cane et al. (2017) illustrate the difficulty in inferring causality in low-pass-filtered data. Given the limited length of the observed AMOC record, climate models are a key method to build understanding. For example, in preindustrial control runs of climate models, Delworth et al. (2017) show that an interactive ocean model is required to reproduce observed relationships between climate indices in the North Atlantic. Kim et al. (2020) built on this work with their own experiments in a coupled model, showing that a North Atlantic Oscillation (NAO)-like heat flux anomaly imposed in the Labrador Sea can create SST patterns that looks like the AMV via changes in the thermohaline circulation (and atmospheric teleconnections).

The same Zhang et al. (2019) review paper recognizes the possibility that multidecadal Atlantic SST variability may not be caused by AMOC (Zhang et al. 2019). A series of recent papers compiles evidence that external radiative forcing plays fundamental role in the phasing and magnitude of the modern AMV (e.g., Bellomo et al. 2018; Bellucci et al. 2017; Birkel et al. 2018; Booth et al. 2012; Mann et al. 2020, 2021; Murphy et al. 2017, 2021; Otterå et al. 2010; Undorf et al. 2018; Watanabe and Tatebe 2019). For example, Murphy et al. (2021) use paired ensembles of a single climate model to show that an interactive ocean model, and therefore AMOC, are not necessary to reproduce the timing of the AMV after about 1930. Conversely, the inclusion of greenhouse gases and industrial aerosols seems necessary to produce the correct timing of the AMV index, in climate models (Booth et al.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Jeremy M. Klavans, jeremy.klavans@colorado.edu

DOI: 10.1175/JCLI-D-21-0338.1

2012; Undorf et al. 2018; Watanabe and Tatebe 2019). Prior to about 1950, in both models and observations, volcanic eruptions coincide with phase changes of the AMV (Otterå et al. 2010; Birkel et al. 2018). The Zhang et al. review criticized this hypothesis by noting, generally, that 1) these papers typically use only one climate model, which may have incorrect or incomplete ocean physics, 2) external forcing explains a small, fixed amount of AMV variance throughout the last millennium, and 3) climate models produce too little internally generated variability.

As for the twentieth century, there is also considerable debate over the AMV's link to either ocean currents or external forcing over the last millennium. Mjell et al. (2016) show that their reconstruction of Iceland Scotland Overflow Water (ISOW) speed is coherent with both marine and terrestrial proxy reconstructions of the AMV. That is, the AMV is associated with variations in part of the overturning ocean circulation. Mjell et al. (2016) show that ISOW speed lags the AMV by up to two decades. This reconstruction and a similar proxy for ISOW speed produce at least intermittent negative correlations with the NAO, which is related to North Atlantic SSTs (Boessenkool et al. 2007; McCarthy et al. 2015; Delworth et al. 2017). O'Reilly et al. (2019) show an in-phase relationship between the integral of an early instrumental record of the NAO and three AMV reconstructions, on multidecadal time scales. The authors interpret this as evidence for an ocean-driven AMV, following evidence in McCarthy et al. (2015) of a link between the integrated NAO and AMV in modern observations. Of course, establishing causal relationships between proxy records is difficult and may be impeded by errors in dating (Hu et al. 2017), or due to the diverse nature of AMV proxies, which include tree rings, corals, sclerosponges, speleothems, and land and ocean sediments—all of which may record different climatological phenomena through different proxy mechanisms (Waite et al. 2020).

Proxy records and climate models also offer evidence that the AMV responds to natural changes in external forcing over the last millennium. Analyses of terrestrial proxy records show that there is a linear relationship between external forcing and the AMV (Knudsen et al. 2014; O'Reilly et al. 2019). In one proxy reconstruction, external forcing explains up to 30% of North Atlantic SST variance (J. Wang et al. 2017). These and other proxy records show coherence with both solar and volcanic forcing (Otterå et al. 2010; Schmitt et al. 2020; Waite et al. 2020). Comparison of these records with climate models adds evidence that preindustrial AMV cold events coincide with volcanic eruptions (Waite et al. 2020). Without variable volcanic forcing, climate models do not produce an AMV index with the enhanced low-frequency variability seen in observations (Otto-Bliesner et al. 2016; Mann et al. 2021).

Given the evidence presented above, it seems likely that both internal atmosphere–ocean variability and external radiative forcing contribute to the AMV over the last millennium, and through present-day. We seek to understand the relative contributions of both internal variability and external forcing to the AMV. Here, we apply six state-of-the-art ensembles of climate models, modern observational products, and four paleoclimate records to build that understanding. We apply this large amount of data across such a long time period to address the concern that the role of forcing in the AMV may be model dependent and/or fixed at preindustrial levels with a small amount of explained variance (Zhang et al. 2019). Overall, we find that external forcing grows from explaining about 30% of AMV variance prior to 1950 to explaining 77% of AMV variance after 1950. This coincides with the sources of external forcing evolving from primarily natural to primarily anthropogenic. When forcing is large and variable, the AMV is primarily externally forced; when forcing is stable there is a larger role for internal variability. We use this large collection of climate model simulations and observations to further argue against the criticism that models produce too little internal variability, and instead we argue that the ratio of the forced signal to internal noise in SST in models is likely too small relative to observations (Smith et al. 2019; Klavans et al. 2021).

2. Methods

a. Models

To aid understanding of the AMV in the post-1850 period, we use output from the six models in the Multimodel Large Ensemble Archive (MMLEA) that contained the required output at the time of writing (Deser et al. 2020). Of those six models, two were initialized in 1850 (CSIRO Mk3.0 and MPI-ESM-LR), two were initialized in 1920 (GFDL CM3 and NCAR CESM-LE), and two were initialized in 1950 (CCCma CanESM2 and GFDL-ESM2M). In total, we consider ~30 000 model years of output across 271 ensemble members. We use the start of each run through the end of 2005, where all models experience identical time histories of external forcing (Schmidt et al. 2011). We place additional emphasis on 1950–2005, the time period common to all six ensembles. The postwar period has the added advantage of higher quality observational data. For this time period, we also construct a multimodel large ensemble of all 271 ensemble members. A brief summary of the models we use here is in Table 1. More details can be found in Deser et al. (2020) or on the website for MMLEA (http://www.cesm.ucar.edu/projects/community-projects/MMLEA/).

To reach farther back into time, we also consider the CESM Last Millennium Ensemble (CESM-LME; Otto-Bliesner et al. 2016). These simulations use the same version of CESM as CESM-LE, although at a coarser resolution in the atmosphere and land components (~2° × 2° instead of 1° × 1°). All runs considered in this paper begin in 850 and extend to 2005. A total of 13 members of this ensemble experience the best estimates of all historical external forcings over the entire period, which we call "all-forcings."

We do not consider any model years beyond 2005 in this analysis because the CESM-LME only extends through the end of 2005. Including the scenario-forced portion of the MMLEA models provides results that are qualitatively consistent with the results we present below. If anything, the growth of external forcing between 2006 and 2021 strengthens our arguments.

TABLE 1. Description of the large ensembles from the MMLEA and CESM-LME used in this study [Deser et al. (2020); individual ensembles are from Jeffrey et al. (2013), Kay et al. (2015), Rodgers et al. (2015), Kirchmeier-Young et al. (2017), Sun et al. (2018), Maher et al. (2019), and Otto-Bliesner et al. (2016)]. Hawkins et al. (2016) established the "method of initialization" dichotomy. "Micro" initialization refers to those runs that begin from the same climate state with minute (round-off level) prescribed differences in the atmosphere alone. "Macro" initialization means that each ensemble member begins from a different climate state; this climate state does not necessarily correspond to observations. Details on downloading model output are available from MMLEA online (http://www.cesm.ucar.edu/projects/community-projects/MMLEA/). Note that the ensemble size appears to be larger than on the MMLEA website because we do not use the projection portions of CESM, allowing us to use an additional member.

Institution	Model	Start year	Horizontal resolution (atmosphere/ocean)	Method of initialization	No. of members
National Center for Atmospheric Research (NCAR)	Community Earth System Model, version 1 (CESM)	1920	~1.3° × 0.9°/nominal 1.0°	Micro	41
Geophysical Fluid Dynamics Laboratory (GFDL)	Climate Model, version 3 (CM3)	1920	~2.0° × 2.5°/~1.0° × 0.9°	Micro	20
GFDL	Earth System Model with MOM, version 4 component (ESM2M)	1950	~2.0° × 2.5°/~1.0° × 0.9°	Macro	30
Canadian Centre for Climate Modelling and Analysis (CCCma)	Second Generation Canadian Earth System Model (CanESM2)	1950	~2.8° × 2.8°/~1.4° × 0.9°	Macro and micro	50
Commonwealth Scientific and Industrial Research Organisation (CSIRO)	Mark 3.0 (Mk3.0)	1850	~1.9° × 1.9°/~1.9° × 1.0°	Macro	30
Max Planck Institute for Meteorology (MPI)	Earth System Model, low resolution (ESM-LR)	1850	~1.9° × 1.9°/nominal 1.5°	Macro	100
NCAR	CESM1 Last Millennium Ensemble	850	~2° × 2°/nominal 1.0°	Micro	13

In addition, this millennium ensemble contains five "single-forcing" small ensembles in which each forcing can be evaluated separately. In these ensembles, greenhouse gases (3 members), land-use change (3 members), orbital (3 members), volcanic (5 members), and solar (4 members) forcing are individually allowed to vary while all other forcings are held constant at 850 levels.

To study the influence of volcanic forcing on the AMV, we follow Otto-Bliesner et al. (2016) in partitioning CESM-LME into those runs that experience volcanic forcing (i.e., all-forcings and the volcanic single-forcing ensemble) and those that to not (all other single forcing runs). For convenience we call these new ensembles "Volc" and "NoVolc," respectively. The ensemble mean AMV of the 18 member Volc ensemble can roughly be interpreted as the volcanically forced AMV, especially prior to the industrial revolution. Conversely, the ensemble mean of the 13 member No Volc ensemble can be roughly be interpreted as the AMV index from a counterfactual world without volcanic eruptions. More details on CESM-LME can be found in Otto-Bliesner et al. (2016) or online (https://www.cesm.ucar.edu/projects/community-projects/LME/).

Last, we consider a long preindustrial control run of CESM1 for comparison with CESM-LME. In this run, all forcings are held constant at 1850 levels (Kay et al. 2015).

b. Indices

For the AMV index, we use the 10-yr low-pass-filtered, linearly de-trended, spatially weighted average SST anomalies over the Atlantic (0°–60°N, 80°W–0°; Enfield et al. 2001). Anomalous SST is calculated by subtracting the monthly climatology and then averaging over the annual cycle. We subdivide the Atlantic into midlatitude (40°–60°N, 80°W–0°) and tropical (0°–20°N, 80°W–0°) regional indices that we call AMVmid and AMVtrop, respectively. When we de-trend a subsection of a model run, we de-trend over only that subsection. To ensure results are not an artifact of de-trending, we also consider the North Atlantic SST (NASST) index, which is the low-pass-filtered, spatially weighted average SST over the AMV region. The NASST index is not de-trended. Our time series of the response to all external forcings is taken to be the ensemble mean AMV index from the CESM-LME all-forcings runs. Likewise, the response to volcanic forcing is taken to be the ensemble mean AMV index from the CESM-LME "Volc" ensemble.

c. Observations and observational products

To calculate our instrumentally observed AMV index, we use the Extended Reconstructed Sea Surface Temperature, version 5.1 (ERSSTv5), dataset at 2° × 2° resolution for 1854–2005 (Huang et al. 2017).

d. Proxy records and reconstructions

We analyze four relatively recent reconstructions of the AMV, three of which allow for direct comparison with O'Reilly et al. (2019). Gray et al. (2004) reconstructed the AMV from 1567 to 1990 based on a network of 12 tree-ring records from around the Atlantic basin. In total, 9 of the 12 records were from Europe or the Middle East and 3 were from the southeastern United States. Mann et al. (2009) reconstructed global SSTs from a multiproxy network composed of tree-ring, ice core, coral, sediment, and other records. They take an area average in

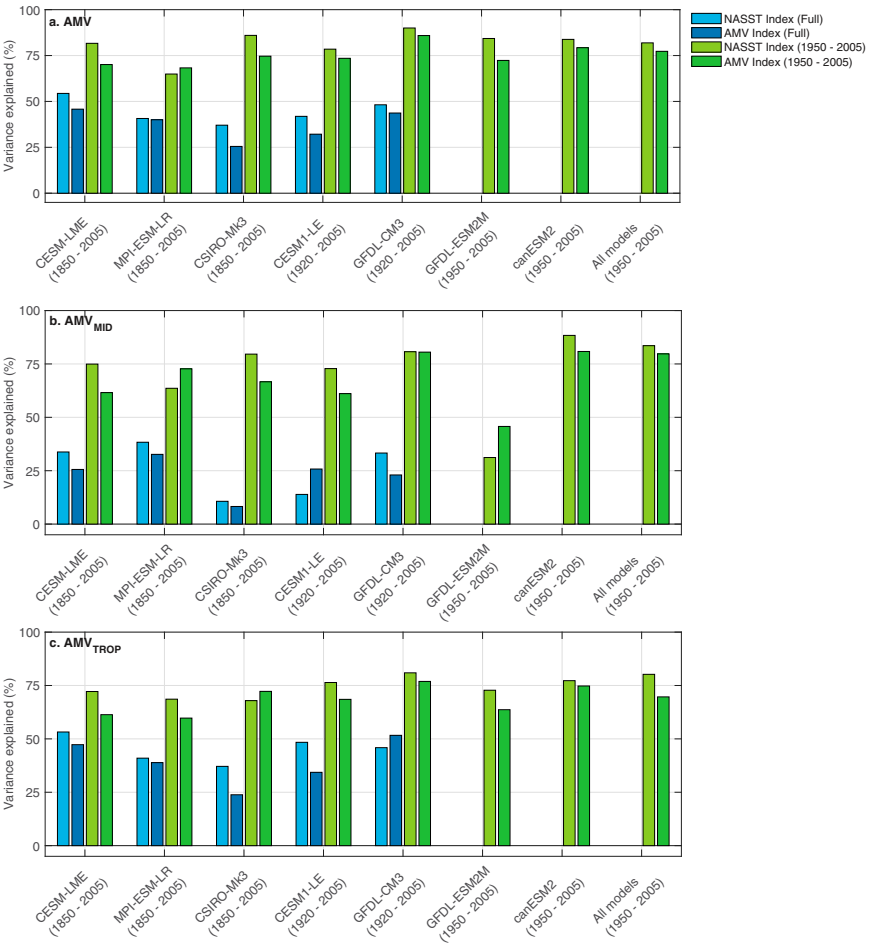

FIG. 1. The variance explained in the observed NASST and AMV indices by the ensemble mean of each individual ensemble and the agglomeration of all ensemble members. Values are displayed for two periods. First, for MMLEA models we display the "full" time period in shades of blue. As shown in Table 1, these models are initialized in 1850, 1920, or 1950. For CESM-LME, we consider the "full" time period to be 1850–2005. The period considered is listed below each model. Second, we show results for the period common to all ensembles (1950–2005) in shades of green. We evaluate the influence of de-trending by comparing the NASST time series (not de-trended; lighter shades) with the AMV time series (de-trended; darker shades).

the Atlantic to produce an AMV index from 1650 to 1980. Some overlapping terrestrial records were used in the more recent J. Wang et al. (2017) reconstruction of the AMV index. This reconstruction captures summertime Atlantic variability from 800 to 2010, which is similar to the annual average. Last, we include a single sclerosponge proxy record from the Bahamas that extends from 1385 to 1993 (Waite et al. 2020). External forcing is known to be a contributor to both the J. Wang et al. (2017) and Waite et al. (2020) records.

3. Results

a. Variations in the modern AMV are externally forced in multiple large ensembles

In the second half of the twentieth century, external forcing can explain a majority of the time history of the AMV. On average, the ensemble mean AMV index explains 40%, 39%, or 83% of the observed AMV variance for those ensembles initialized in 1850, 1920, or 1950, respectively (Fig. 1a). All models, regardless of when they are initialized, explain a larger amount of variance between 1950 and 2005 than in any other time period (the average over all 6 MMLEA single-model ensembles and the LME is 79%), suggesting a dominant role for external forcing. When we use all models to create a multimodel ensemble mean (weighting each of the 271 members equally), we find that this new ensemble mean explains 77% of the observed variance. The growing role for external forcing over the 19th and 20th centuries appears with or without linear de-trending. On average the ensemble mean NASST index (not de-trended) explains 44%, 45%, and 84% of the observed NASST variance for those ensembles initialized in 1850, 1920, or 1950, respectively. Like for the AMV,

the ensemble mean NASST in all models explains more variance in the second half of the twentieth century than for any other time period (multimodel average: 81%). The ensemble mean of all individual ensemble members explains 82% of the observed variance. The growing role we identify for external forcing in the twentieth century is consistent with prior work that shows a similar metric in CMIP5 models and CESM-LENS (Murphy et al. 2017, 2021).

After 1950, external forcing contributes consistently in the midlatitude and tropical Atlantic as well. Previous work suggests different mechanisms contribute to the AMV in each of these regions (e.g., Buckley et al. 2012, 2014). We find that, for both the AMV_{MID} and the AMV_{TROP}, external forcing explains most of the observed variance (Figs. 1b,c). Over the 1950–2005 period, on average across all seven single-model ensembles, external forcing explains 74% and 72% of the AMVmid and AMVtrop variance, respectively. For the same time period, the external forcing in the multimodel large ensemble explains 80% of the observed AMVmid variance and 70% of the observed AMVtrop variance. This stands in contrast to time periods that include the late 19th and early twentieth century. For the years 1850–2005, the ensemble mean only explains 28% of observed AMV_{MID} and 44% of observed AMV_{TROP} variance. For the years 1920–2005, the ensemble mean only explains 24% of observed AMV_{MID} and 47% of observed AMV_{TROP} variance. We therefore suggest that in the late twentieth century, external forcing contributes to the AMV signal in all parts of the North Atlantic. As with the whole-basin AMV index, there is a smaller role for external forcing prior to 1950. Further, external forcing in the pre-1950 midlatitudes is less influential than in the pre-1950 tropics.

External forcing increases AMV variance at low frequencies and the influence of forcing is more prevalent later in the twentieth century. Murphy et al. (2017) shows that historically forced CMIP5 climate models have higher AMV variance than preindustrial control runs of the same models. In one model, Bellomo et al. (2018) show how forcing influences the AMV at the lowest resolvable frequencies by comparing the spectra of the AMV index with the spectra of the unforced component of the AMV index. We expand this analysis and find similar results, particularly in those MMLEA models initialized in the twentieth century (Fig. 2). The separation between the total AMV index and the unforced AMV index indicates a role for external forcing at those frequencies. The two models initialized in 1850 (MPI-ESM-LR and CSIRO Mk3.0; Figs. 2a,b) show less separation between the spectra, indicating that for this time period the AMV has a large internal component. For those four models initialized later in the twentieth century, with the possible exception of GFDL-ESM2M, a separation between the two spectra begins to emerge at periods longer than 10–20 years (Figs. 2c–f). This again suggests that as we move farther into the twentieth century, the influence of external forcing on the AMV magnitude grows.

The outsized role for external forcing and sources of that forcing are discernable in paleorecords of the AMV. As described above, we evaluate four paleorecords of the AMV, all of which are positively correlated with instrumental observations after 1850 (Fig. 3a). Like in Fig. 1, there are high correlations between the MMLEA ensemble mean and each record of the AMV (not shown). We evaluate the sources of this forcing using the CESM Last Millennium Ensemble and its associated single forcing runs (see Methods). In both the CESM-LME Volc and No Volc ensembles, the ensemble mean is better correlated with the Waite et al. Wang et al. and Gray et al. records than all individual ensemble members (Fig. 3b). That is, both volcanic and non-volcanic forcings contribute to the AMV. But, looking closer, most of the correlation between the Volc ensemble mean and observations are driven by the all-forcings ensemble. For example, the Volc ensemble mean explains 66% of the ERSST AMV variance between 1854 and 2005. Removing the three volcanic-only runs from this ensemble only reduces the variance explained to 62%. From 1854 to 1950, the Volc ensemble explains 48% of ERSST variance with the volcanic-only runs included and 39% without. After 1950, the Volc ensemble explains 78% of ERSST variance with the volcanic-only runs included and 82% without. We find similar patterns in the paleoreconstructions. The Waite et al. record shows the strongest influence of volcanic variability, although this may be due to the site of the sclerosponge proxy. In general, we find that volcanic forcing does not explain the phasing of the post 1950 AMV. Volcanic forcing does offer a 9% marginal improvement in our ability to explain the 1854–1950 AMV. Proxies and instrumental observations yield similar results, therefore building trust that the paleorecords can be helpful in examining the role and sources of external forcing in the pre-1850 AMV.

b. Forced variations in the pre-1850 AMV in one model and proxy records

Previous studies have unearthed signals from both solar and volcanic variability in proxy records of Atlantic SST (e.g., Mann et al. 2021; O'Reilly et al. 2019; Schmitt et al. 2020; Waite et al. 2020; J. Wang et al. 2017). Here, we confirm previous findings that there is a role for volcanic forcing in the pre-1850 AMV in CESM-LME using different metrics. As expected, no two proxy records are exactly alike (Fig. 4a). Over the period for which we have data from all records (1650–1980), all combinations of proxy AMV time series are statistically significantly positively correlated, except for the Waite et al. and Mann et al. records (Table 2). We note that the Waite et al. record is from a single sclerosponge, whereas the other records include information from multiple records. When we consider only the period prior to 1850 for which we have data from all four records (1650–1849), the correlations between the records decrease. We note that this is outside the calibration period for all three proxy reconstructions. Here, only the Wang et al. record is statistically significantly positively correlated with the Gray et al. and Mann et al. records (Table 3). To capture the variability common to all four of these records during the 1650–1849 time period we also report on the first principal component (PC) of these time series. The Volc ensemble mean has positive significant correlations at the 95% level with all four proxy datasets and their first PC

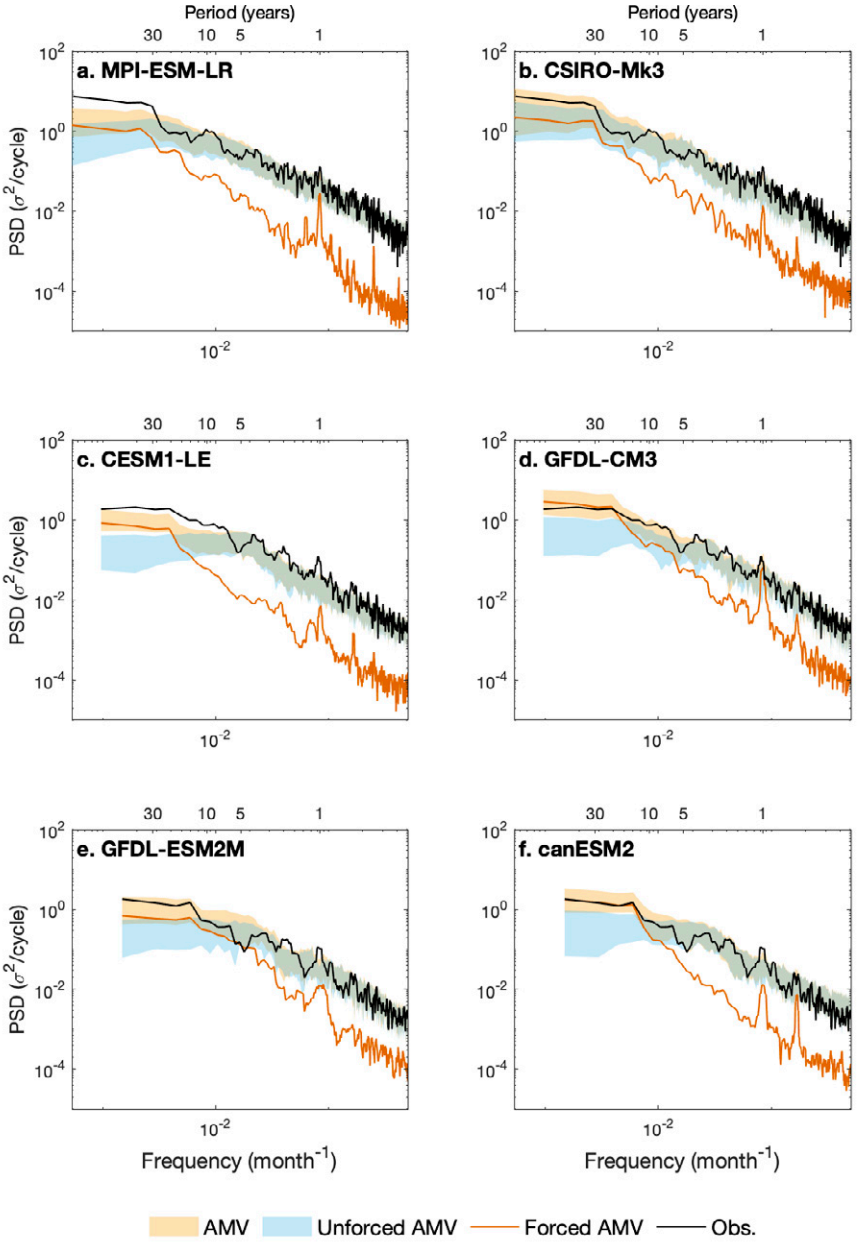

FIG. 2. Spectra of the AMV index from observations (black) and MMLEA models over the full length of each run. The orange cloud represents the ~95% confidence interval of spectra of the AMV index from individual ensemble members in each model. The orange solid line is the spectrum of the ensemble mean AMV index. The blue cloud is the ~95% confidence interval from the internal portion of the AMV index in each model. To calculate the internal portion of the AMV, we subtract the ensemble mean AMV index from the AMV index in each individual ensemble member.

($r = 0.54$; $p < 0.01$), in contrast to the unforced (control) simulation. Like in J. Wang et al. (2017), volcanic forcing explains about 30% of the 1650–1849 AMV variance. We reenforce their finding that volcanic forcing is an important influence on the SST in this pre-instrumental period (see also Otterå et al. 2010). We note that the earlier portions of the Wang et al. and Waite et al. records do not have as high a correlation with the Volc ensemble (not shown). The No Volc ensemble mean correlation is less than or equal to zero in all four records and first PC ($r = -0.30$; $p < 0.01$). This negative correlation does not preclude the possibility that the non-volcanic forcings have a detectable influence in this period. Overall, the results indicate that volcanic aerosols are a major forcing of the AMV over this time period.

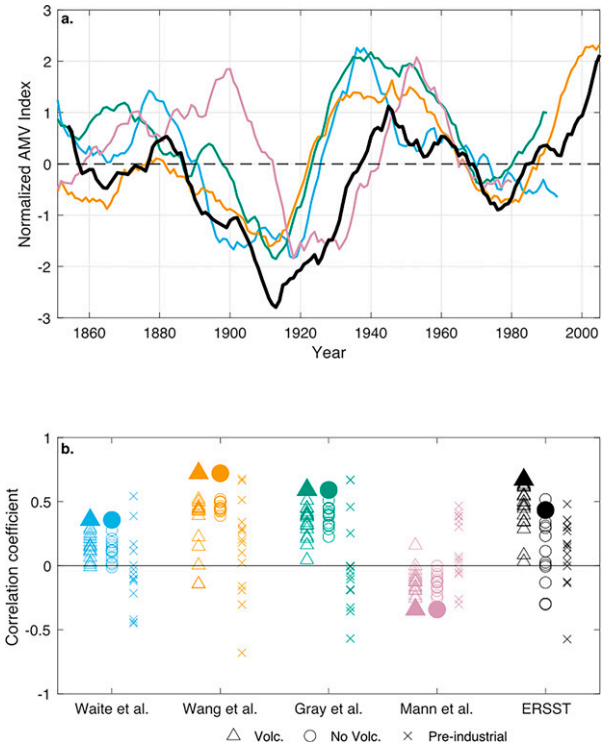

FIG. 3. The role of volcanic forcing in the post-1850 AMV: (a) time series of four proxy records of the AMV (colored lines) and observational product (black line), along with (b) correlations between the AMV reconstructions/observations and the ensembles with and without volcanic forcing (triangles and circles, respectively). Correlations between data and individual ensemble members are shown with open symbols; correlations between data and ensemble means are shown with filled symbols. For comparison, we also include bootstrapped correlations from a long preindustrial control run of CESM (crosses).

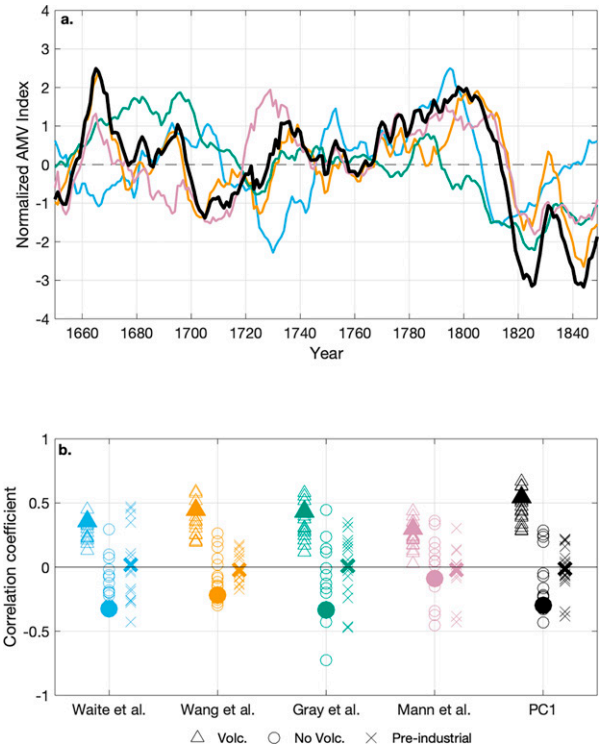

FIG. 4. The role of volcanic forcing in the pre-1850 AMV: (a) time series of four proxy records of the AMV for 1650–1849 (colored lines) and the first principal component (PC1) of these time series (black line), along with (b) correlations between the AMV reconstructions and the ensembles with and without volcanic forcing (triangles and circles, respectively). Correlations between data and individual ensemble members are shown with open symbols; correlations between data and ensemble means are shown with filled symbols. For comparison, we also include bootstrapped correlations from a long preindustrial control run of CESM (crosses).

Volcanic forcing has a stronger influence on the pre-1850 AMV than on the post-1850 AMV. We follow Swingedouw et al. (2017) and O'Reilly et al. (2019) in applying a "superposed epoch analysis" to the CESM-LME (Fig. 5). We differ from those two previous studies in dividing our time series into pre and post-1850 sections. Prior to 1850, we find an average −1.18°C ($p < 0.01$) response in the AMV index in the first full year after a large volcanic eruption (Fig. 5a). This basinwide response lingers for nearly a decade (Fig. 5b). Conversely, after 1850 there is not a statistically significant AMV response to the 10 largest eruptions after one year (Fig. 5c, 0.16°C; $p = 0.25$). Interestingly, the SST response to volcanic forcing after 1850 is spatially heterogeneous, with local maxima in cooling in the subpolar gyre and eastern Atlantic (Fig. 5d). There is even a warm response near the equator. However, the spatial variations in this response appear to be attributable to the greenhouse warming signal. Volcanic forcing induces a spatially homogenous response in linearly de-trended SSTs after 1850 (Fig. 5e), more like pre-1850. Overall, we conclude that in the pre-1850 period, volcanic eruptions strongly influence SST (Robock 2000) and that volcanic influence lingers for nearly a decade. But, as in Fig. 4, the influence of volcanoes wanes as we move into the twentieth century.

There is a larger influence of volcanic forcing on SST prior to 1850 because there were larger eruptions over this longer time period. We arrive at this parsimonious result by comparing the column mass of volcanic aerosols in CESM-LME with the 10-yr integrated temperature response (Fig. 6). Predictably, the largest eruptions are those that impose durable SST responses on the Atlantic. To date, those eruptions only occurred prior to the twentieth century. We also test the

TABLE 2. Correlations between low-pass-filtered proxy records over the period common to all records (1650–1980). Asterisks represent statistical significance at the 95% level.

	Waite	Wang	Gray	Mann
Waite	—			
Wang	0.245*	—		
Gray	0.243*	0.509*	—	
Mann	−0.097	0.366*	0.162*	—

TABLE 3. Correlations between low-pass-filtered proxy records over the period common to all records before 1850 (1650–1849). Asterisks represent statistical significance at the 95% level.

	Waite	Wang	Gray	Mann
Waite	—			
Wang	0.099	—		
Gray	0.097	0.308*	—	
Mann	−0.131	0.523*	0.110	—

possibility that volcanic forcing is less efficacious in a more polluted, modern atmosphere (similar to Shindell et al. 2001). We find that when we pick eruptions in the 850–1849 period that are of similar size to those observed over the post-1850 period the SST response matches that of the modern period (−0.23°C; $p = 0.26$; Figs. 5f,g). That is, these smaller eruptions fail to elicit a statistically significant SST response, even in a less polluted atmosphere. This also points toward the pithy conclusion that large eruptions are better at altering the AMV. Last, we find no evidence that "clusters" of consecutive eruptions can together force large changes in the AMV (not shown). Again, this suggests that volcanic activity has induced changes in the AMV in the past and implies that future large volcanic eruptions will drive future changes in the AMV.

If we could predict these future volcanic eruptions, we would enhance our ability to predict the AMV [relative to the AMV's ability to predict itself, i.e., persistence as modeled by Eq. (1) below]. To quantify the volcanic influence on AMV predictability, we evaluate the same data within the context of Granger causality. Following McGraw and Barnes (2018), we first create two regression models:

$$\mathrm{AMV} = b_1\mathrm{AMV}_{t-1} + b_2\mathrm{AMV}_{t-2} + \cdots + b_n\mathrm{AMV}_{t-n} + e \text{ and} \quad (1)$$

$$\mathrm{AMV} = b_1\mathrm{AMV}_{t-1} + b_2\mathrm{AMV}_{t-2} + \cdots + b_n\mathrm{AMV}_{t-n} + c_1\mathrm{Volc}_{t-1} + c_2\mathrm{Volc}_{t-2} + \cdots + c_n\mathrm{Volc}_{t-n} + e, \quad (2)$$

where b is a coefficient on the lagged AMV, c is a coefficient on the lagged time history of volcanic forcing, t is the time at lag 0, and e is the error term. We determine the number of lags n as the model with the minimum Akaike information criterion value (Mosedale et al. 2006; McGraw and Barnes 2018). (We also choose the number of lags using the Bayesian information criterion to similar effect.) We determine if volcanic forcing Granger causes the AMV if it meets two criteria:

1) At least one value of c is statistically significant by a two-sided t test.
2) The F statistic is greater than the critical value for that model.

We first apply this test to the "all forcings" CESM-LME. The analysis above suggests we ought to find that volcanic forcing influences the AMV prior to 1850, but not afterward. We find that in the pre-1850 period volcanic forcing Granger causes the AMV in all 13 members of CESM-LME. After 1850, volcanic forcing only Granger causes the AMV in 4 out of 13 members (30%). We find similar results for the AMVmid and AMVtrop. Volcanic forcing Granger causes the AMV in all 13 members in both regions prior to 1850. After 1850, volcanic forcing only Granger causes the AMV in 1 of 13 and 8 of 13 members in the AMVmid and AMVtrop, respectively. Overall, this test has confirmed our priors. Volcanic forcing generally Granger causes the AMV prior to 1850, but volcanic forcing rarely Granger causes the AMV after 1850.

We next test our proxy records via the same analysis. Both the Waite et al. and Wang et al. records meet the two criteria to claim that volcanic eruptions Granger cause the AMV prior to 1850 (either for the common period, 1659–1849, or for the full length of the record prior to 1850). This does not rule out the possibility that nonvolcanic forcings are also contributing to the AMV in the Waite et al. and/or Wang et al. records (as mentioned above; Fig. 4). Conversely, none of the proxy records produce a Granger causal relationship between volcanic forcing and the AMV after 1850. We interpret this as additional evidence that large eruptions (like those occurring prior to 1850) help to explain the AMV and that small eruptions (like most of those occurring after 1850) do not explain the forced component of the AMV.

4. Discussion

In this paper, we show that external forcing helps guide the AMV over the last millennium and takes a leading role in the AMV after 1950. Our main conclusions are straightforward:

1) When changes in external radiative forcing are large, the AMV responds to those changes in external forcing. More granularly, when there are large volcanic eruptions, volcanism paces the AMV. When greenhouse gases and aerosol forcing is large, anthropogenic forcing guides the AMV.
2) When external radiative forcing is nearly stable, there is room for internally generated variability to dominate.

In the last millennium, the largest explosive volcanic eruptions occurred prior to 1850. Climate models and proxy records show a consistent low-frequency Atlantic SST response to these large eruptions (Waite et al. 2020; Mann et al. 2021). In the periods between these large eruptions, the AMV may have a meaningful internally generated component (forcing only explains about one-third of AMV variance). In the industrial era, volcanoes have been relatively quiet, abdicating their leading role in forcing the AMV. Between 1850 and 1950, there is a role for external forcing in the AMV, including a waning role for volcanic forcing, but it is not the sole explanatory variable. This 100-yr window may be a good time period to study internal mechanisms for the AMV in instrumental observations. After 1950, anthropogenic forcing asserts its dominance on both global climate and the AMV. Above we estimate that external forcing explains 77%–79% of the observed variance in the post-1950 AMV

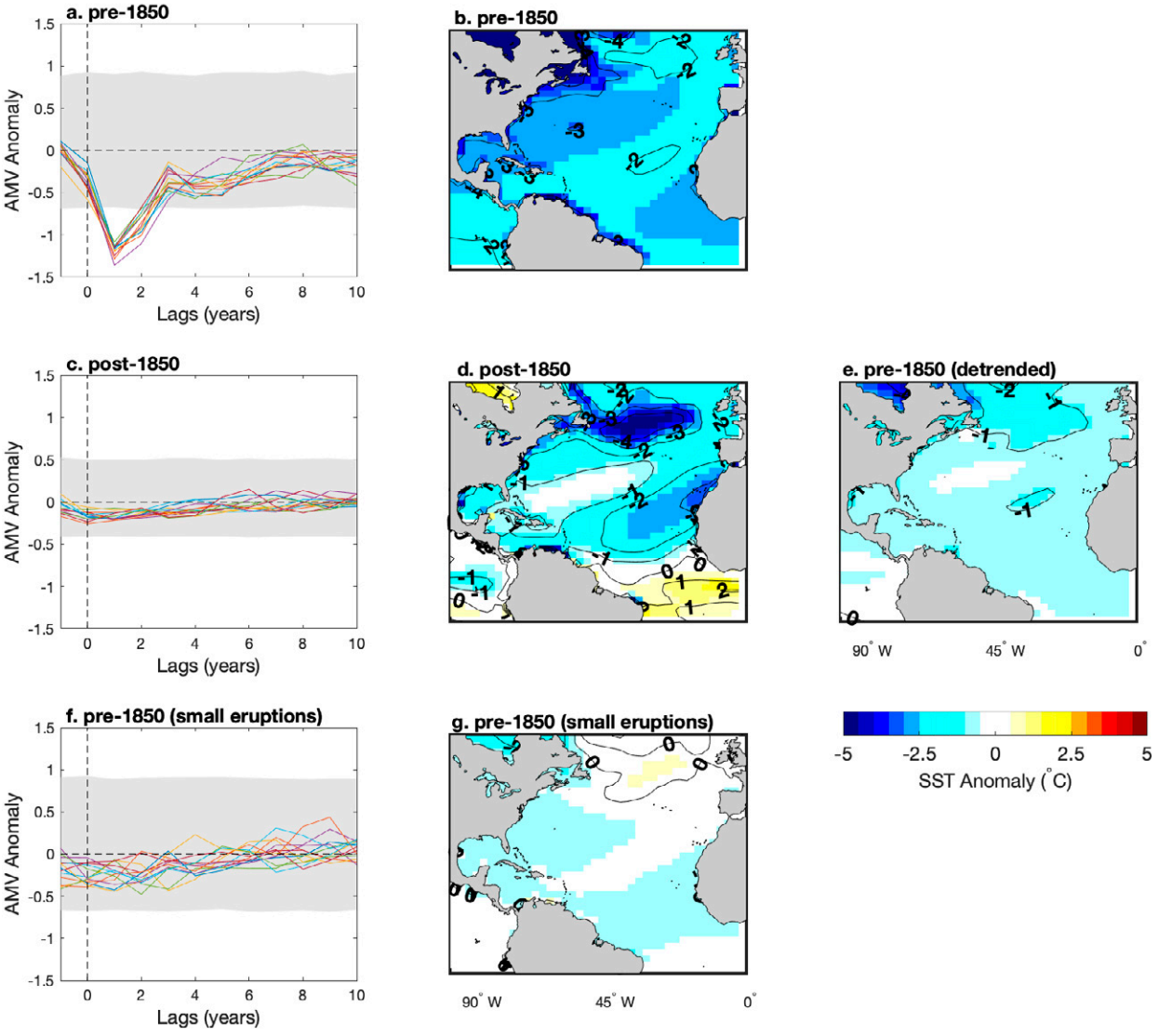

FIG. 5. Composite response to volcanic forcing in the CESM-LME before and after 1850: (a),(c),(f) Superposed epoch analysis of the AMV index after the 10 largest eruptions in each period. The 95% confidence interval (gray shaded box) is calculated by randomly subsampling 8-yr time slices, with replacement. (b),(d),(g) Maps of the composite SST response to the 10 largest eruptions in each time period. (e) As in (b),(d),(g), but for linearly de-trended SST after 1850. In (f) and (g), we control for the differences in eruption magnitude for the two time periods that we consider. That is, we identify the eruptions in the pre-1850 period that are most similar in size to the 10 largest eruptions in the post-1850 period. We use these 10 eruptions to construct the superposed epoch analysis in (f) and the composite map in (g).

index. This is consistent with the results of Bellomo et al. (2018), Otto-Bliesner et al. (2016), and Waite et al. (2020). Bellomo et al. (2018) uses single forcing runs from the CESM-LME to show that greenhouse gas forcing and anthropogenic aerosols are the key causes of the forced component of the AMV (see also Booth et al. 2012; Murphy et al. 2017, 2021). We similarly find that volcanic forcing has a negligible influence on the AMV after 1950. The externally forced signal may be particularly visible in the Atlantic because of the history of fossil-fuel-burning industry (e.g., Booth et al. 2012), because there is less internally generated noise in the Atlantic than in the Pacific Ocean (Bellomo et al. 2018), or because forcing excites an internal mode of variability unique to the Atlantic (Swingedouw et al. 2017).

Prior to 1950, there are periods of time when external forcing does not explain the majority of AMV variance. Our analysis leaves three possible explanations for the remaining variance:

1) Observations of the AMV are less precise farther back in time. Differences in timing between the AMV in climate models and older instrumental records or proxy observations may be a result of observational errors.

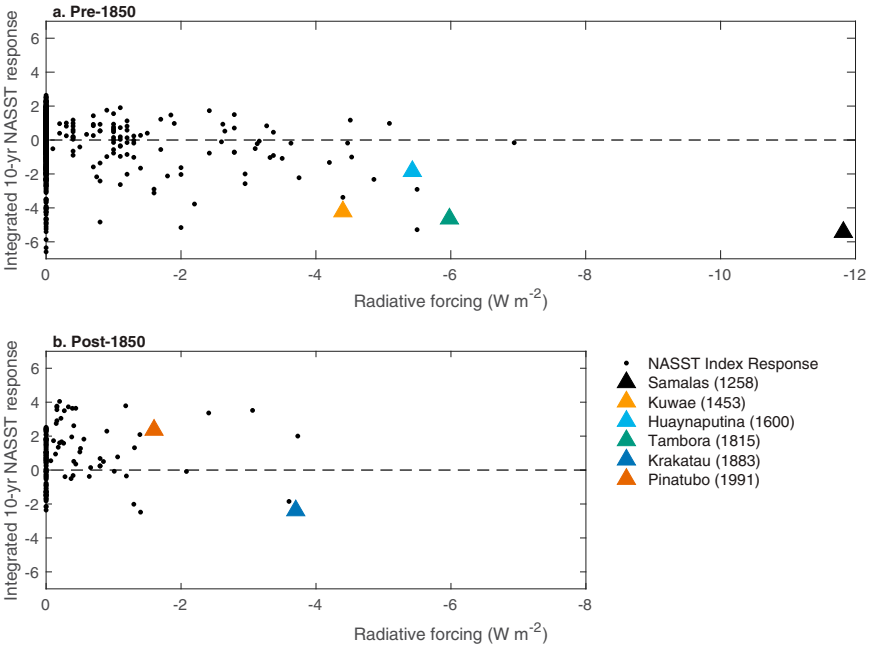

FIG. 6. The size of the volcanic eruption influences the long-term SST response in CESM-LME (a) prior to 1850 and (b) after 1850. Values of radiative forcing associated with volcanic eruptions are borrowed from Crowley (2000). We integrate the 10-yr SST response to each eruption to approximate the ability of the eruption to project onto the AMV. Triangles represent the large, well-known eruptions in each period.

2) Reconstructions of external forcing from proxy records is also less precise farther back in time. Climate models may not be responding to the true time series of external forcing.
3) There is more room for internal variability in the AMV to be dominant prior to the strong anthropogenic forcing post-1950 and between large volcanic eruptions. In CESM-LME, total variance appears to be the linear combination of internally generated and externally forced variability (Waite et al. 2020, their Table S3.1). Therefore, as forcing is reduced, the ratio of forced to total AMV variance nonlinearly decreases.

Our results differ from some previous interpretations of the drivers of the AMV by placing an emphasis on external forcing. In the next three paragraphs we suggest three structural differences between our study and previous work that argue for a larger role for internal variability: 1) the definition of the AMV index, 2) omitted variable bias, and 3) the "signal-to-noise paradox" in climate models.

Several previous studies used alternative definitions of the AMV index. We consider the whole-basin SST anomaly both with and without its trend. These indices, which do not exclude external forcing, are the ones most often used in studies of impacts of the AMV (e.g., Enfield et al. 2001; Trenberth and Shea 2006; Klavans et al. 2020). Other studies deliberately exclude the externally forced component of the index. Ting et al. (2009) do so via signal-to-noise maximizing empirical orthogonal function (see also Ting et al. 2014). Yan et al. (2019) do so by privileging information about the ocean circulation. Different definitions of the AMV yield different projections of the AMV in the future (e.g., Frajka-Williams et al. 2017). Better understanding of which indices best explain observed impacts may be useful to clarify the literature.

Prior workers may have inadvertently introduced statistical bias into their results by excluding the externally forced component of the AMV. For example, O'Reilly et al. (2019) identified a high correlation between the integrated NAO index and the AMV over the last several centuries. They interpret the integrated NAO index to be a representation of the large-scale ocean circulation (following McCarthy et al. 2015). Of course, the NAO itself may also have a meaningful externally forced component (Klavans et al. 2021). O'Reilly et al. [2019; their Eq. (2)] formalizes their model of the AMV as follows:

$$\mathrm{AMV}(t) \propto \int_{t_0}^{t} \mathrm{NAO}(t')\, dt'. \tag{3}$$

In the appendix, we show that the high correlations reported from this model are exaggerated by omitted variable bias, which arises by omitting the possibility of external forcing directly influencing both the SST and the NAO.

Recent publications offer contradictory criticisms of the signal-to-noise ratio produced by climate models. These models are accused of both 1) generating an overexuberant forced signal, relative to internal variability (e.g., X. Wang et al. 2017, Kravtsov 2017; Kim et al. 2018a; Sato et al. 2018; Toll et al. 2019 Chylek et al. 2020), and 2) underrepresenting the

"predictable signal" relative to unpredictable noise (i.e., the "signal-to-noise paradox"; Scaife and Smith 2018; Smith et al. 2019). In ensembles of CMIP5 models, this predictable signal is not only associated with model initialization and thus likely also related to the external forcing (Scaife and Smith 2018; Borchert et al. 2021; Klavans et al. 2021; Zhang et al. 2021). If the forced signal is too weak, that can explain the discrepancy in amplitude between the observed AMV index and the ensemble mean index in CMIP5 models (Murphy et al. 2017; Kim et al. 2018b). In the context of our results, we suggest that if the signal-to-noise ratio in climate models could be corrected, the separation between the total and unforced spectra in Fig. 2 would be more dramatic and that fewer ensemble members will be required to capture the forced component of the AMV. While we continue to unravel the signal-to-noise paradox, large ensembles can help us to understand and predict the AMV as they have helped us to understand other components of North Atlantic climate variability (e.g., Borchert et al. 2021; Menary et al. 2020).

Overall, our results read as a synthesis of the role of external forcing in the AMV. In constructing this synthesis, we hope that we have rejoined some of the outstanding critiques of this hypothesis [as reviewed by Zhang et al. (2019)]. To reprise, we group the collective criticism into three categories: 1) prior studies typically use only one climate model, which may have incorrect or incomplete ocean physics or initial states, 2) external forcing explains a small, fixed amount of AMV variance throughout the last millennium, and 3) climate models produce too little internally generated variability. To address the first point, we find a large role for external forcing after 1950 in six large ensembles of uninitialized climate models (Figs. 1 and 2). For the post-1950 time period, this clearly rebuts the statement that "hindcasts with prescribed changes in external radiative forcings but no initialization in ocean states are not able to predict the observed AMV" (Zhang et al. 2019). Second, the other paleorecords agree with J. Wang et al. (2017) that external forcing explains about one-third of preindustrial AMV variance. However, we show that the forced component of the AMV explains more variance over time, as anthropogenic forcing grows (see Fig. 1). Third, we suggest that an ensemble mean AMV that is highly correlated with observations but too weak in amplitude is consistent with the signal-to-noise paradox (Fig. 2; Murphy et al. 2017; Scaife and Smith 2018). We hope that by addressing some of these gaps in the literature, we can move toward building a fuller understanding of how AMOC and external forcing may interact in a highly forced environment.

Acknowledgments. NOAA ERSSTv5.1 and paleoclimatic proxy data were provided by NOAA/OAR/ESRL PSD (https://www.esrl.noaa.gov/psd/). Access and download to the MMLEA were facilitated by NCAR's supercomputing resources provided by NSF/CISL/Cheyenne (see http://www.cesm.ucar.edu/projects/community-projects/MMLEA/). HadISST data were obtained online (https://www.metoffice.gov.uk/hadobs/hadisst/) and are 2020 British Crown Copyright by the Met Office, provided under a "Non-Commercial Government Licence" (http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/). We acknowledge grants from the National Science Foundation (NSF) Climate and Large-Scale Dynamics program (Grants AGS 1735245 and AGS 1650209) and the NSF Paleo Perspectives on Climate Change program (Grant AGS 1703076) as well as from the NOAA Climate Program Office.

Data availability statement. The model simulations analyzed in this study are available online (https://www.cesm.ucar.edu/projects/community-projects/MMLEA/).

APPENDIX

Omitted Variable Bias and the AMV

In section 3c we asserted that the model in O'Reilly et al. (2019) suffers from omitted variable bias by not explicitly including external forcing. They correlate the integral of an early instrumental NAO index with three of the four AMV proxies we consider here. [The Waite et al. record was published after O'Reilly et al. (2019)]. They find that the Wang et al. proxy AMV is well correlated with the integrated NAO over the period 1659–1999 ($r = 0.65$). The Mann et al. and Gray et al. records are also positively correlated, but with a smaller magnitude ($r = 0.51$ for both). We reproduce the model from O'Reilly et al. (2019) in Eq. (3) and again here:

$$\mathrm{AMV}(t) = \beta_1 \int_0^t \mathrm{NAO}(t')\, dt', \tag{A1}$$

where β_1 is a constant of proportionality, $t = 0$ is the initial time of the NAO time series, and $\mathrm{AMV}(t = 0)$ is ignored. This relation may also be taken as a linear regression:

$$\mathrm{AMV}(t) = \beta_1 \int_0^t \mathrm{NAO}(t')\, dt' + u, \tag{A2}$$

where β_1 is the regression coefficient on the integrated NAO and u is the error term. In this model, any potential role for external forcing is confined to the error term, u or through an indirect influence on the NAO. Therefore, we consider a second model that explicitly represents the integral of external forcing:

$$\mathrm{AMV}(t) = \beta_1 \int_0^t \mathrm{NAO}(t')\, dt' + \beta_2 \int_0^t F(t')\, dt' + \upsilon, \tag{A3}$$

where β_2 is the regression coefficient on external forcing, F is a time series of external radiative forcing, and υ is our new error term. O'Reilly et al. (2019) also show that there is an externally forced component to the NAO [their Fig. 8; in agreement with Klavans et al. (2021)]. Both this paper and O'Reilly et al. (2019) also show external forcing influences the AMV. These implies that $\beta_2 \neq 0$ and $\mathrm{cov}(F, u) \neq 0$. The conditional mean zero assumption therefore fails; the ordinary least squares estimate will be biased. Further, because we know the relationships between forcing

and the NAO and forcing and the AMV are both positive, the estimate of β_1 has a positive bias.

We illustrate the omitted variable bias by calculating regression models with [Eq. (A3)] and without [Eq. (A2)] forcing. We estimate the time series of integrated external forcing $F(t)$ as the ensemble mean AMV from CESM-LME; that should average out internal forcing and leave only the forcing common to all ensemble members. For the NAO index, we follow O'Reilly et al. (2019)'s method to calculate an NAO index from an early instrumental record of sea level pressure (Luterbacher et al. 2002). In the NAO-only model in Eq. (A2), regressing on the Wang et al. AMV record gives $\beta_1 = 0.60$ ($p = 0.94$). When we include integrated external forcing [Eq. (A3)], the estimate of the regression coefficient decreases to 0.46 ($p = 0.88$) and the only remaining statistically significant period is in the mid-twentieth century. The regression coefficients for the NAO from all other proxies also decrease when we go from Eq. (A2) to Eq. (A3).

REFERENCES

Bellomo, K., L. N. Murphy, M. A. Cane, A. C. Clement, and L. M. Polvani, 2018: Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble. *Climate Dyn.*, **50**, 3687–3698, https://doi.org/10.1007/s00382-017-3834-3.

Bellucci, A., A. Mariotti, and S. Gualdi, 2017: The role of forcings in the twentieth-century North Atlantic multidecadal variability: The 1940–75 North Atlantic cooling case study. *J. Climate*, **30**, 7317–7337, https://doi.org/10.1175/JCLI-D-16-0301.1.

Birkel, S. D., P. A. Mayewski, K. A. Maasch, A. V. Kurbatov, and B. Lyon, 2018: Evidence for a volcanic underpinning of the Atlantic multidecadal oscillation. *npj Climate Atmos. Sci.*, **1**, 24, https://doi.org/10.1038/s41612-018-0036-6.

Bjerknes, J., 1964: Atlantic air-sea interaction. *Advances in Geophysics*, Vol. 10, Academic Press, 1–82, https://doi.org/10.1016/S0065-2687(08)60005-9.

Boessenkool, K. P., I. R. Hall, H. Elderfield, and I. Yashayaev, 2007: North Atlantic climate and deep-ocean flow speed changes during the last 230 years. *Geophys. Res. Lett.*, **34**, L13614, https://doi.org/10.1029/2007GL030285.

Booth, B. B. B., N. J. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232, https://doi.org/10.1038/nature10946.

Borchert, L. F., V. Koul, M. B. Menary, D. J. Befort, D. Swingedouw, G. Sgubin, and J. Mignot, 2021: Skillful decadal prediction of unforced southern European summer temperature variations. *Environ. Res. Lett.*, **16**, 104017, https://doi.org/10.1088/1748-9326/ac20f5.

Buckley, D. Ferreira, J.-M. Campin, J. Marshall, and R. Tulloch, 2012: On the relationship between decadal buoyancy anomalies and variability of the Atlantic meridional overturning circulation. *J. Climate*, **25**, 8009–8030, https://doi.org/10.1175/JCLI-D-11-00505.1.

——, R. M. Ponte, G. Forget, and P. Heimbach, 2014: Low-frequency SST and upper-ocean heat content variability in the North Atlantic. *J. Climate*, **27**, 4996–5018, https://doi.org/10.1175/JCLI-D-13-00316.1.

Cane, M. A., A. C. Clement, L. N. Murphy, and K. Bellomo, 2017: Low-pass filtering, heat flux, and Atlantic multidecadal variability. *J. Climate*, **30**, 7529–7553, https://doi.org/10.1175/JCLI-D-16-0810.1.

Chylek, P., C. Folland, J. D. Klett, and M. K. Dubey, 2020: CMIP5 climate models overestimate cooling by volcanic aerosols. *Geophys. Res. Lett.*, **47**, e2020GL087047, https://doi.org/10.1029/2020GL087047.

Crowley, T. J., 2000: Causes of climate change over the past 1000 years. *Science*, **289**, 270–277, https://doi.org/10.1126/science.289.5477.270.

Delworth, T. L., F. Zeng, L. Zhang, R. Zhang, G. A. Vecchi, and X. Yang, 2017: The central role of ocean dynamics in connecting the North Atlantic Oscillation to the extratropical component of the Atlantic multidecadal oscillation. *J. Climate*, **30**, 3789–3805, https://doi.org/10.1175/JCLI-D-16-0358.1.

Deser, C., and Coauthors, 2020: Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Climate Change*, **10**, 277–286, https://doi.org/10.1038/s41558-020-0731-2.

Enfield, D. B., A. M. Mestas-Nuñez, and P. J. Trimble, 2001: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the Continental U.S. *Geophys. Res. Lett.*, **28**, 2077–2080, https://doi.org/10.1029/2000GL012745.

Folland, C., T. N. Palmer, and D. E. Parker, 1986: Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature*, **320**, 602–607, https://doi.org/10.1038/320602a0.

——, J. Owen, M. N. Ward, and A. Colman, 1991: Prediction of seasonal rainfall in the Sahel region using empirical and dynamical methods. *J. Forecasting*, **10**, 21–56, https://doi.org/10.1002/for.3980100104.

Frajka-Williams, E., C. Beaulieu, and A. Duchez, 2017: Emerging negative Atlantic multidecadal oscillation index in spite of warm subtropics. *Sci. Rep.*, **7**, 11224, https://doi.org/10.1038/s41598-017-11046-x.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479, https://doi.org/10.1126/science.1060040.

Gray, S. T., L. J. Graumlich, J. L. Betancourt, and G. T. Pederson, 2004: A tree-ring based reconstruction of the Atlantic multidecadal oscillation since 1567 A.D. *Geophys. Res. Lett.*, **31**, L12205, https://doi.org/10.1029/2004GL019932.

Gulev, S. K., M. Latif, N. Keenlyside, W. Park, and K. P. Koltermann, 2013: North Atlantic Ocean control on surface heat flux on multidecadal timescales. *Nature*, **499**, 464–467, https://doi.org/10.1038/nature12268.

Hastenrath, S., 1984: Interannual variability and annual cycle: Mechanisms of circulation and climate in the tropical Atlantic sector. *Mon. Wea. Rev.*, **112**, 1097–1107, https://doi.org/10.1175/1520-0493(1984)112<1097:IVAACM>2.0.CO;2.

Hawkins, E., R. S. Smith, J. M. Gregory, and D. A. Stainforth, 2016: Irreducible uncertainty in near-term climate projections. *Climate Dyn.*, **46**, 3807–3819, https://doi.org/10.1007/s00382-015-2806-8.

Hu, J., J. Emile-Geay, and J. Partin, 2017: Correlation-based interpretations of paleoclimate data—Where statistics meet past climates. *Earth Planet. Sci. Lett.*, **459**, 362–371, https://doi.org/10.1016/j.epsl.2016.11.048.

Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1.

Jeffrey, S., L. Rotstayn, M. Collier, S. Dravitzki, C. Hamalainen, C. Moeseneder, K. Wong, and J. Syktus, 2013: Australia's CMIP5 submission using the CSIRO Mk3.6 model. *Aust. Meteor. Oceanogr. J.*, **63**, 1–14, https://doi.org/10.22499/2.6301.001.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.

Kim, W. M., S. Yeager, P. Chang, and G. Danabasoglu, 2018a: Low-frequency North Atlantic climate variability in the community Earth system model large ensemble. *J. Climate*, **31**, 787–813, https://doi.org/10.1175/JCLI-D-17-0193.1.

——, ——, and G. Danabasoglu, 2018b: Key role of internal ocean dynamics in Atlantic multidecadal variability during the last half century. *Geophys. Res. Lett.*, **45**, 13 449–13 457, https://doi.org/10.1029/2018GL080474.

——, ——, and ——, 2020: Atlantic multidecadal variability and associated climate impacts initiated by ocean thermohaline dynamics. *J. Climate*, **33**, 1317–1334, https://doi.org/10.1175/JCLI-D-19-0530.1.

Kirchmeier-Young, M. C., F. W. Zwiers, and N. P. Gillett, 2017: Attribution of extreme events in Arctic sea ice extent. *J. Climate*, **30**, 553–571, https://doi.org/10.1175/JCLI-D-16-0412.1.

Klavans, J. M., A. C. Clement, L. N. Murphy, and H. Zhang, 2020: Identifying the externally forced Atlantic multidecadal variability signal through Florida rainfall. *Geophys. Res. Lett.*, **47**, e2020GL088361, https://doi.org/10.1029/2020GL088361.

——, M. A. Cane, A. C. Clement, and L. N. Murphy, 2021: NAO predictability from external forcing in the late 20th century. *npj Climate Atmos. Sci.*, **4**, 22, https://doi.org/10.1038/s41612-021-00177-8.

Knight, J. R., C. K. Folland, and A. A. Scaife, 2006: Climate impacts of the Atlantic multidecadal oscillation. *Geophys. Res. Lett.*, **33**, L17706, https://doi.org/10.1029/2006GL026242.

Knudsen, M. F., B. H. Jacobsen, M.-S. Seidenkrantz, and J. Olsen, 2014: Evidence for external forcing of the Atlantic multidecadal oscillation since termination of the Little Ice Age. *Nat. Commun.*, **5**, 3323, https://doi.org/10.1038/ncomms4323.

Kravtsov, S., 2017: Pronounced differences between observed and CMIP5-simulated multidecadal climate variability in the twentieth century. *Geophys. Res. Lett.*, **44**, 5749–5757, https://doi.org/10.1002/2017GL074016.

Lamb, P. J., 1978: Large-scale tropical Atlantic surface circulation patterns associated with subsaharan weather anomalies. *Tellus*, **30**, 240–251, https://doi.org/10.3402/tellusa.v30i3.10338.

Lough, J. M., 1986: Tropical Atlantic sea surface temperatures and rainfall variations in sub-Saharan Africa. *Mon. Wea. Rev.*, **114**, 561–570, https://doi.org/10.1175/1520-0493(1986)114<0561:TASSTA>2.0.CO;2.

Luterbacher, J., and Coauthors, 2002: Reconstruction of sea level pressure fields over the eastern North Atlantic and Europe back to 1500. *Climate Dyn.*, **18**, 545–561, https://doi.org/10.1007/s00382-001-0196-6.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *J. Adv. Model. Earth Syst.*, **11**, 2050–2069, https://doi.org/10.1029/2019MS001639.

Mann, M. E., and Coauthors, 2009: Global signatures and dynamical origins of the Little Ice Age and medieval climate anomaly. *Science*, **326**, 1256–1260, https://doi.org/10.1126/science.1177303.

——, B. A. Steinman, and S. K. Miller, 2020: Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nat. Commun.*, **11**, 49, https://doi.org/10.1038/s41467-019-13823-w.

——, ——, D. J. Brouillette, and S. K. Miller, 2021: Multidecadal climate oscillations during the past millennium driven by volcanic forcing. *Science*, **371**, 1014–1019, https://doi.org/10.1126/science.abc5810.

McCarthy, G. D., I. D. Haigh, J. J.-M. Hirschi, J. P. Grist, and D. A. Smeed, 2015: Ocean impact on decadal Atlantic climate variability revealed by sea-level observations. *Nature*, **521**, 508–510, https://doi.org/10.1038/nature14491.

McGraw, M. C., and E. A. Barnes, 2018: Memory matters: A case for granger causality in climate variability studies. *J. Climate*, **31**, 3289–3300, https://doi.org/10.1175/JCLI-D-17-0334.1.

Menary, M. B., and Coauthors, 2020: Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophys. Res. Lett.*, **47**, e2020GL088166, https://doi.org/10.1029/2020GL088166.

Mjell, T. L., U. S. Ninnemann, H. F. Kleiven, and I. R. Hall, 2016: Multidecadal changes in Iceland Scotland overflow water vigor over the last 600 years and its relationship to climate. *Geophys. Res. Lett.*, **43**, 2111–2117, https://doi.org/10.1002/2016GL068227.

Mosedale, T. J., D. B. Stephenson, M. Collins, and T. C. Mills, 2006: Granger causality of coupled climate processes: Ocean feedback on the North Atlantic Oscillation. *J. Climate*, **19**, 1182–1194, https://doi.org/10.1175/JCLI3653.1.

Murphy, L. N., K. Bellomo, M. Cane, and A. Clement, 2017: The role of historical forcings in simulating the observed Atlantic multidecadal oscillation. *Geophys. Res. Lett.*, **44**, 2472–2480, https://doi.org/10.1002/2016GL071337.

——, J. M. Klavans, A. C. Clement, and M. A. Cane, 2021: Investigating the roles of external forcing and ocean circulation on the Atlantic multidecadal SST variability in a large ensemble climate model hierarchy. *J. Climate*, **34**, 4835–4849, https://doi.org/10.1175/JCLI-D-20-0167.1.

O'Reilly, C. H., L. Zanna, and T. Woollings, 2019: Assessing external and internal sources of Atlantic multidecadal variability using models, proxy data, and early instrumental indices. *J. Climate*, **32**, 7727–7745, https://doi.org/10.1175/JCLI-D-19-0177.1.

Otterå, O. H., M. Bentsen, H. Drange, and L. Suo, 2010: External forcing as a metronome for Atlantic multidecadal variability. *Nat. Geosci.*, **3**, 688–694, https://doi.org/10.1038/ngeo955.

Otto-Bliesner, B. L., and Coauthors, 2016: Climate variability and change since 850 CE: An ensemble approach with the Community Earth System Model. *Bull. Amer. Meteor. Soc.*, **97**, 735–754, https://doi.org/10.1175/BAMS-D-14-00233.1.

Qin, M., A. Dai, and W. Hua, 2020: Quantifying contributions of internal variability and external forcing to Atlantic multidecadal variability since 1870. *Geophys. Res. Lett.*, **47**, e2020GL089504, https://doi.org/10.1029/2020GL089504.

Robock, A., 2000: Volcanic eruptions and climate. *Rev. Geophys.*, **38**, 191–219, https://doi.org/10.1029/1998RG000054.

Rodgers, K. B., J. Lin, and T. L. Frölicher, 2015: Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences*, **12**, 3301–3320, https://doi.org/10.5194/bg-12-3301-2015.

Ruprich-Robert, Y., R. Msadek, F. Castruccio, S. Yeager, T. Delworth, and G. Danabasoglu, 2017: Assessing the climate impacts of the observed Atlantic multidecadal variability using the GFDL CM2.1 and NCAR CESM1 global coupled models.

J. Climate, **30**, 2785–2810, https://doi.org/10.1175/JCLI-D-16-0127.1.

Sato, Y., D. Goto, T. Michibata, K. Suzuki, T. Takemura, H. Tomita, and T. Nakajima, 2018: Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model. *Nat. Commun.*, **9**, 985, https://doi.org/10.1038/s41467-018-03379-6.

Scaife, A. A., and D. Smith, 2018: A signal-to-noise paradox in climate science. *npj Climate Atmos. Sci.*, **1**, 28, https://doi.org/10.1038/s41612-018-0038-4.

Schmidt, G. A., and Coauthors, 2011: Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). *Geosci. Model Dev.*, **4**, 33–45, https://doi.org/10.5194/gmd-4-33-2011.

Schmitt, D., E. Gischler, D. Birgel, J. Peckmann, F. S. Anselmetti, and H. Vogel, 2020: Great Blue Hole (Lighthouse Reef, Belize): A continuous, annually-resolved record of common era sea surface temperature, Atlantic multidecadal oscillation and cyclone-controlled run-off. *Quat. Sci. Rev.*, **247**, 106570, https://doi.org/10.1016/j.quascirev.2020.106570.

Shindell, D. T., G. A. Schmidt, M. E. Mann, D. Rind, and A. Waple, 2001: Solar forcing of regional climate change during the maunder minimum. *Science*, **294**, 2149–2152, https://doi.org/10.1126/science.1064363.

Smith, D. M., and Coauthors, 2019: Robust skill of decadal climate predictions. *npj Climate Atmos. Sci.*, **2**, 13, https://doi.org/10.1038/s41612-019-0071-y.

Sun, L., M. Alexander, and C. Deser, 2018: Evolution of the global coupled climate response to Arctic sea ice loss during 1990–2090 and its contribution to climate change. *J. Climate*, **31**, 7823–7843, https://doi.org/10.1175/JCLI-D-18-0134.1.

Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science*, **309**, 115–118, https://doi.org/10.1126/science.1109496.

Swingedouw, D., J. Mignot, P. Ortega, M. Khodri, M. Menegoz, C. Cassou, and V. Hanquiez, 2017: Impact of explosive volcanic eruptions on the main climate variability modes. *Global Planet. Change*, **150**, 24–45, https://doi.org/10.1016/j.gloplacha.2017.01.006.

Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and internal twentieth-century SST trends in the North Atlantic. *J. Climate*, **22**, 1469–1481, https://doi.org/10.1175/2008JCLI2561.1.

——, ——, and C. Li, 2014: North Atlantic multidecadal SST oscillation: External forcing versus internal variability. *J. Mar. Syst.*, **133**, 27–38, https://doi.org/10.1016/j.jmarsys.2013.07.006.

Toll, V., M. Christensen, J. Quaas, and N. Bellouin, 2019: Weak average liquid-cloud-water response to anthropogenic aerosols. *Nature*, **572**, 51–55, https://doi.org/10.1038/s41586-019-1423-9.

Trenberth, K. E., and D. J. Shea, 2006: Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.*, **33**, L12704, https://doi.org/10.1029/2006GL026894.

Undorf, S., M. A. Bollasina, B. 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. *Geophys. Res. Lett.*, **45**, 11 930–11 940, https://doi.org/10.1029/2018GL079970.

Waite, A. J., J. M. Klavans, A. C. Clement, L. N. Murphy, V. Liebetrau, A. Eisenhauer, R. J. Weger, and P. K. Swart, 2020: Observational and model evidence for an important role for volcanic forcing driving Atlantic multidecadal variability over the last 600 years. *Geophys. Res. Lett.*, **47**, e2020GL089428, https://doi.org/10.1029/2020GL089428.

Wang, J., B. Yang, F. C. Ljungqvist, J. Luterbacher, T. J. Osborn, K. R. Briffa, and E. Zorita, 2017: Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years. *Nat. Geosci.*, **10**, 512–517, https://doi.org/10.1038/ngeo2962.

Wang, X., J. Li, C. Sun, and T. Liu, 2017: NAO and its relationship with the Northern Hemisphere mean surface temperature in CMIP5 simulations. *J. Geophys. Res. Atmos.*, **122**, 4202–4227, https://doi.org/10.1002/2016JD025979.

Watanabe, M., and H. Tatebe, 2019: Reconciling roles of sulphate aerosol forcing and internal variability in Atlantic multidecadal climate changes. *Climate Dyn.*, **53**, 4651–4665, https://doi.org/10.1007/s00382-019-04811-3.

Yan, X., R. Zhang, and T. R. Knutson, 2019: A multivariate AMV index and associated discrepancies between observed and CMIP5 externally forced AMV. *Geophys. Res. Lett.*, **46**, 4421–4431, https://doi.org/10.1029/2019GL082787.

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. *Rev. Geophys.*, **57**, 316–375, https://doi.org/10.1029/2019RG000644.

Zhang, W., B. Kirtman, L. Siqueira, A. Clement, and J. Xia, 2021: Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model. *Climate Dyn.*, **56**, 2895–2913, https://doi.org/10.1007/s00382-020-05621-8.