Investigating the Roles of External Forcing and Ocean Circulation on the Atlantic Multidecadal SST Variability in a Large Ensemble Climate Model Hierarchy

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

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

(Manuscript received 12 March 2020, in final form 23 February 2021)

ABSTRACT: This paper attempts to enhance our understanding of the causes of Atlantic multidecadal variability (AMV). Following the literature, we define the AMV as the SST averaged over the North Atlantic basin, linearly detrended and low-pass filtered. There is an ongoing debate about the drivers of the AMV, which include internal variability generated from the ocean or atmosphere (or both) and external radiative forcing. We test the role of these factors in explaining the time history, variance, and spatial pattern of the AMV using a 41-member ensemble from a fully coupled version of CESM and a 10-member ensemble of the CESM atmosphere coupled to a slab ocean. The large ensemble allows us to isolate the role of external forcing versus internal variability, and the model differences allow us to isolate the role of coupled ocean circulation. Both with and without coupled ocean circulation, external forcing explains more than half of the variance of the observed AMV time series, indicating its important role in simulating the twentieth-century AMV phases. In this model the net effect of ocean processes is to reduce the variance of the AMV. Dynamical ocean coupling also reduces the ability of the model to simulate the characteristic spatial pattern of the AMV, but forcing has little impact on the pattern. Historical forcing improves the time history and variance of the AMV simulation, while the more realistic ocean representation reduces the variance below that observed and lowers the correlation with observations.

KEYWORDS: Atlantic Ocean; Ocean circulation; Sea surface temperature; General circulation models; Climate variability; Multidecadal variability

1. Introduction

Atlantic multidecadal variability (AMV; Kerr 2000) has a significant regional and hemispheric impact on climate, notably the number of Atlantic hurricanes and precipitation in the Sahel, northeast Brazil, the central United States, and northern Europe (Folland et al. 2001; Goldenberg et al. 2001; Knight et al. 2006; Latif et al. 2007; Sutton and Hodson 2005; Zhang and Delworth 2006). There is ongoing debate about its causes, which generally fall into two points of view. One, articulated most recently in a review paper by Zhang et al. (2019) and the large number of papers cited therein, is that changes in the strength of the Atlantic meridional overturning circulation (AMOC) cause changes in the upper-ocean temperatures. Another is that the changes observed in the twentieth century are mostly driven by changes in external forcing (greenhouse gases, anthropogenic aerosols, and volcanic eruptions), which set the pace and amplitude of the twentieth-century AMV (Otterå et al. 2010; Booth et al. 2012; Murphy et al. 2017; Bellucci et al. 2017; Bellomo et al. 2018; Birkel et al. 2018; Undorf et al. 2018; Watanabe and Tatebe 2019; Mann et al. 2020). This does not discount atmospheric circulation variability and the role of stochastic forcing by the dominant mode of variability in the North Atlantic, the North Atlantic Oscillation (NAO), on the ocean. The AMV has been described as both an upper-ocean thermal response to NAO forcing, a mechanism that does not require a dynamic ocean (Clement et al. 2015; Cane et al. 2017), and, in other accounts, as a lagged response to AMOC variability driven by the NAO (McCarthy et al. 2015; Delworth and Zeng 2016; Delworth et al. 2017; Mecking et al. 2014; Wills et al. 2019; O'Reilly et al. 2019), a process that requires dynamic coupling between the atmosphere and ocean. Klavans et al. (2019) recently argued that while the latter relationship is evident over a limited region (i.e., the subpolar gyre) in coupled model simulations with constant forcing, introducing variable external forcing obscures this relationship. In addition, modeling studies have demonstrated that the AMOC might be forced (Cheng et al. 2013; Tandon and Kushner 2015; Undorf et al. 2018; Menary et al. 2020).

Most of the arguments for these mechanisms rely on climate model simulations. Attribution studies are fairly straightforward in the case of external forcing: compare a simulation of Atlantic climate without changes in external forcing (preindustrial control runs) to those with a best estimate of radiative forcing over the industrialized period (historical simulations). The difference is attributed to forcing, but the forced simulations will also have an unforced component. A more decisive method for quantifying the role of forcing is with a large ensemble (Knutson et al. 2013; Kay et al. 2015). In this approach, a large number of simulations are performed

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

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-20-0167.s1.

Corresponding author: Lisa N. Murphy, lmurphy@rsmas.miami.edu

DOI: 10.1175/JCLI-D-20-0167.1

with a climate model that is forced with identical external forcing, but with changes in the initial conditions (Deser et al. 2020). The features of the climate simulation that are common to all ensemble members (typically as measured by the ensemble mean) can be attributed to forcing with reasonably high confidence, and differences among ensemble members are due to internal climate variability. Using these methods, previous studies have shown that in CMIP5 models the timing of the forced shifts in the AMV are consistent with observations and that there is more variance at multidecadal time scales with forcing than without (Clement et al. 2015; Murphy et al. 2017; Bellucci et al. 2017; Bellomo et al. 2018)—with the important caveat that almost all existing model simulations underestimate the AMV magnitude. Models may also respond too strongly to external forcing (Kim et al. 2018) leading to an underestimate of the role of internal variability in the AMV. Other methods of estimating the relative roles of external forcing and internal variability from a multimodel ensemble have found that there is an important contribution from internally generated variability to the AMV (Ting et al. 2009, 2015; Yan et al. 2019), indicating that the specific methods for extracting the externally forced signal can affect the interpretation.

Attributing AMV-related changes is much more difficult for the AMOC. Continuous observations of the AMOC strength only extend back to 2004, and come from the RAPID array across 24°N (McCarthy et al. 2015). However, it is unclear to what extent the AMOC signal represents local wind-driven ocean circulation variability (Zhao and Johns 2014) rather than internal interhemispheric AMOC variations in heat transport (Larson et al. 2019). So attributing surface temperature variability to AMOC requires looking for the signatures of the AMOC in climate models. Some of these include spatial patterns in surface and subsurface temperature (Frankignoul et al. 2015; Wills et al. 2016; Zhang 2008), correlations between AMOC indices and SST (Delworth and Mann 2000; Danabasoglu et al. 2012; Tandon and Kushner 2015), or correlations between SST and surface fluxes and salinity (Gulev et al. 2013; O'Reilly et al. 2016; Zhang 2017; Yan et al. 2019). Many of these relationships have been interpreted to imply an active role for the ocean circulation, which is then taken as evidence that the ocean circulation is the main driver or a necessary component to get the correct SST response to the driver. However, much of this work is done in coupled models where determining causality can be problematic, especially when model data are heavily filtered to extract the low-frequency variability, and this filtering can impact the interpretation of leads and lags in both models and observations (Trenary and DelSole 2016; Cane et al. 2017). Using a simplified model, where causality can be determined by design, Li et al. (2020) showed that ocean dynamics are required to produce multidecadal SST variability consistent with observations, but this analysis was restricted to the extratropical North Atlantic, and also did not attribute the ocean's role to the AMOC.

Herein, we present an experiment designed to isolate the roles of both external forcing and ocean dynamics on the AMV. The tools we use to address the causality conundrum are an existing ensemble of historically forced simulations with a fully coupled model, the CESM, for the period 1920–2005 (LENS-FC), an existing ensemble of millennium long simulations with the same model (LENS-LME), and simulations for 1920–2005 with a new 10-member ensemble of the same atmospheric general circulation model coupled to a motionless slab ocean (LENS-SOM). Our strategy is to compare the characteristics of the AMV in the fully coupled model with those in the slab ocean model (SOM) to see how the addition of the ocean changes the model performance as measured by comparisons with observations. When we began this investigation, we anticipated answering the following questions:

1) How much of the observed AMV is forced by historical forcing and how much is internal to the climate system?
2) How is the historically forced response made more realistic by adding a fully coupled ocean?
3) How much more internal multidecadal variance is added by including an ocean?

Questions 2 and 3 are premised on the notion that the fully coupled model simulations will be more realistic than those from the SOM. The SOM lacks ocean dynamics, mixed layer physics, and a treatment of salinity. Insofar as these are essential for the AMV (as defined by SST) then the SOM simulations must fall short of the fully coupled version. By construction, the SOM has far less realistic physics so we expected its simulations to be inferior. This expectation was not borne out.

a. Methodology

In this paper, we use the term AMV to denote low-frequency Atlantic SSTs, regardless of whether it is internal or forced, since both are present in observations. There are different ways to calculate the AMV in the literature (e.g., Yan et al. 2019; Frajka-Williams et al. 2017; Frankignoul et al. 2017; Knight 2009; Ting et al. 2009). All apply a low-pass filter, although the nature of the filter is not always specified. Most are detrended. The two most common methods used are those of Trenberth and Shea (2006) or of Enfield et al. (2001). The former removes the global SST average from 60°S to 60°N while the latter removes a simple linear trend in SST. The definition used is important since the Trenberth and Shea (2006) method shows a transition to a negative AMV just after 2010 that has been theorized to be due to a weakening of the AMOC (Smeed et al. 2014; Robson et al. 2016). The Enfield definition shows the AMV is still positive (Frajka-Williams et al. 2017). In Murphy et al. (2017) we found that the Trenberth and Shea (2006) method removed not just the linear part of the greenhouse gas forcing but also the changes associated with aerosols and volcanic eruptions. Here we are interested in the broader drivers of AMV, both natural and externally forced, so we adopt the Enfield definition in our analysis.

b. Model simulations

We perform climate model experiments following the methodology of the National Center for Atmospheric Research (NCAR) Community Earth System Model (CESM) Large Ensemble Project (LENS-Fully Coupled or LENS-FC;

Kay et al. 2015). LENS-FC consists of a 41-member ensemble of fully coupled climate model simulations that covers the historical period (1920–2005). Version 1 of CESM includes the Community Atmosphere Model version 5 (CAM5), which has updates in the ways it represents aerosol–cloud interactions. In addition to aerosol direct and semidirect effects, CAM5 simulates aerosol indirect effects, which was not done in earlier versions of the model (Hurrell et al. 2013). Other studies (Booth et al. 2012; Bellucci et al. 2017) have argued that aerosol indirect effects are important for explaining the time history of the AMV in other models. This experimental design has the utility of being able to test the impact of internal variability on the results, which is especially important in the study of multidecadal variability where observational records only contain a few phases.

In our new simulations we replace the dynamical ocean model with a slab ocean model (LENS-SOM), where the ocean heat transport convergence is prescribed as a q-flux. The q-flux is calculated using output from the last 100 years of a 2200-yr-long CESM fully coupled preindustrial (PI) control simulation. The mixed-layer depth is calculated as the average boundary layer depth from the same PI control run. It varies spatially but not seasonally. LENS-SOM uses the same dynamic-thermodynamic sea ice model that is used in the fully coupled model (Bitz et al. 2012; described in more detail at https://www.cesm.ucar.edu/models/cesm1.1/data8/doc/SOM.pdf). As in the LENS-FC each ensemble member is forced with the same observed estimates of historical forcing (i.e., greenhouse gases, tropospheric and stratospheric aerosols, land use changes, etc.). Each of the 10 members of the LENS-SOM ensemble is created by perturbing each member's initial atmospheric state (i.e., temperature is perturbed at the level of round-off error), as was done in LENS-FC. This 10-member LENS-SOM ensemble is then compared to the 41-member ensemble from LENS-FC. A comparison of these two ensembles allows us to isolate the influence of internal ocean dynamical changes on the AMV.

Note that since all of the members of the LENS-FC ensemble were forced with the same ocean initial conditions, internal oceanic variability does not result from differing ocean states. To explore this effect, we examine the Last Millennium Ensemble (LENS-LME; Otto-Bliesner et al. 2016), which is a fully coupled 13-member ensemble of CESM at 2° resolution that was initialized in 850 CE. Each LME ensemble member was forced with the same time variant changes in observed historical forcing, but with small changes in the atmospheric initial conditions. We also analyze the last 901 years of the 1000-yr-long preindustrial (PI) SOM simulation (PI-SOM) and the last 1801 years of the 2200-yr-long PI-FC simulation.

c. AMV index

We define the AMV in the classic manner of Enfield et al. (2001) and Kerr (2000), as the linearly detrended, area-average SST anomalies over the North Atlantic Ocean. Those early AMV papers were interested in the AMV for its influence on climate, including hurricane frequency and landfall, and rainfall in North America. Our focus on SST is similarly motivated. Since we are interested in multidecadal variability, we apply a 10-yr low-pass (LP) Butterworth filter to the data, but we also show the unfiltered data in our figures below. We also tested the impact of the filter length on our results. When we speak of "variance" or "correlation" it should be understood to be for the low-pass data unless it is explicitly stated otherwise. The SST anomalies are calculated by removing the monthly mean SST and then averaging over each annual cycle. We define AMV indices as the average SST in three regions: the AMV (0°–60°N, 80°W–0°), the AMVmid (40°–60°N, 80°W–0°), and the AMVtrop (0°–20°N, 80°W–0°). This is done for each individual ensemble member of the LENS-FC 41-member ensemble, the LENS-SOM 10-member ensemble, the LENS-LME 13-member ensemble, and observations. The ensemble mean (EM) is calculated by first averaging the SST data in each grid cell over all members of each ensemble and then averaging over the appropriate regions to obtain the AMV, AMVmid, and AMVtrop indices. Since they are linear operations, one gets the same answer if we first average each ensemble member spatially and then average over the ensemble members. The EM is the best estimate of the response to the common external forcing. We estimate the internal, or unforced, AMV by removing the EM AMV from each individual member. The difference in the internal AMV in SOM versus the FC model is due to internal ocean dynamics that are absent in SOM. The difference between the FC and SOM EM values is attributable to the ocean's response to forcing, which is absent in SOM. We will also have need of the mean variance over all ensemble members, denoted the mean over the ensemble (MOE) variance. The MOE variance includes the variance due to internal variability (the departures from the ensemble mean) while the EM variance does not; it is based on the ensemble mean value (i.e., the forced response).

In the results below, we test the null hypothesis that the reported correlation coefficients are equal to zero. When calculating the degrees of freedom for this test we account for the autocorrelation of the time series; if the time series has N values, we reduce the degrees of freedom to $N/(2L)$, where L is the integral (sum) of the autocorrelation coefficients over 50 lags (Anderson 1942; Dawdy and Matalas 1964; Leith 1973; Zar 1999). We chose 50 lags because this value is well beyond the first zero crossing of the autocorrelation in all cases. The p values and the reduced degrees of freedom (Neff) reported for the EM correlation coefficients are based on this test. In addition, we perform two-sample F tests for equal variances on the time series of the AMV.

d. Estimating variance

We make unbiased estimates of the forced variance σ_F^2 and the mean internal variance σ_I^2:

$$\sigma_F^2 = \sigma_{\mathrm{EM}}^2 - \frac{(\sigma_{\mathrm{MOE}}^2 - \sigma_{\mathrm{EM}}^2)}{(J-1)},$$

$$\sigma_I^2 = (\sigma_{\mathrm{MOE}}^2 - \sigma_{\mathrm{EM}}^2) + \frac{(\sigma_{\mathrm{MOE}}^2 - \sigma_{\mathrm{EM}}^2)}{(J-1)},$$

where J is the ensemble size, σ_{MOE}^2 is the mean variance over all ensemble members, and σ_{EM}^2 is the variance of the ensemble mean (Table 1). These formulas are standard, though usually phrased in terms of signal and noise rather than forced and

TABLE 1. Statistics of LP filtered Atlantic SSTs from 1930 to 2005 for LENS-SOM, LENS-FC, and observations. Correlations are Pearson correlation coefficients of the model ensemble with observations. The EM denotes the ensemble mean; it is the best estimate of the model response to the imposed historical forcing. The label MOE denotes the mean correlation or variance over all ensemble members. The EM (adjusted) and internal (adjusted) are unbiased estimates of the variances as described in section 1a and derived in the online supplemental material. The observed values are based on ERSSTv5. The cells that are italicized are our best estimates of the partitioning between forced and internal variance in the observations, and are uncertain. The statistics are for area averages over AMV (0°–60°N), AMVmid (40°–60°N), and AMVtrop (0°–20°N).

	AMV (0°–60°N)			AMVmid (40°–60°)			AMVtrop (0°–20°N)		
	SOM	FC	OBS	SOM	FC	OBS	SOM	FC	OBS
Correlations									
MOE	0.67	0.59		0.55	0.50		0.41	0.29	
EM	0.80	0.75		0.79	0.74		0.50	0.48	
Variance explained (EM)	64%	56%		62%	55%		25%	23%	
Variances ($°C^2$)									
MOE	0.038	0.011	0.026	0.120	0.027	0.065	0.028	0.015	0.023
EM (adjusted)	0.025	0.007	*0.017*	0.051	0.012	*0.040*	0.018	0.005	*0.006*
Internal (adjusted)	0.0133	0.004	*0.010*	0.069	0.015	*0.024*	0.010	0.009	*0.017*
Percentage forced	66%	60%	*64%*	42%	44%	*62%*	64%	36%	*25%*
Percentage internal	34%	40%	*36%*	58%	56%	*38%*	36%	64%	*75%*

internal. A short derivation is included in the online supplemental material.

e. *Observational data*

Observed SST is from the National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperature version 5 (ERSSTv5) (Huang et al. 2017). ERSSTv5 extends from 1854 to the present; however, we use the time period 1920–2005 to be consistent with the LENS simulations. All model data are bilinearly interpolated to the observational 2° × 2° grid. SST variance in ERSSTv5 is compared to version 4 of the same dataset (ERSSTv4; Huang et al. 2015), COBE SST2 data (Hirahara et al. 2014), and the Hadley Centre Sea Ice and SST dataset version 1.1 (HadISST; Rayner et al. 2003) for the same time period. We found minimal differences between versions 4 and 5 of the ERSST SST variance and thus we do not include ERSSTv4 results in the graphics. Additionally, the observed, unfiltered variances are similar among the observed datasets and thus we only show the ERSSTv5 variance values. Atmospheric data including sea level pressure (SLP) and surface winds are taken from the NOAA CIRES twentieth-century reanalysis version 2c, which includes monthly sea level pressure (SLP) and 1000-hPa wind data (Compo et al. 2011). We linearly detrend, remove the long-term mean, and low-pass filter the data (the same postprocessing that was applied to the SST data) before calculating the regression coefficients (coefficients are shown in Fig. 5). As discussed in Deser and Phillips (2017), observations in the North Atlantic suffer from incomplete and uneven coverage especially prior to 1970, which may impact our results.

2. Results

Figure 1 shows the LP filtered AMV for the three different regions analyzed. The LENS-FC fails to reproduce the AMV cool phase evident in the observations in the first decade from 1920 to 1930. LENS-SOM is closer to the observed values especially in the AMVmid region. We note that the ocean heat transport in LENS-SOM was calculated based on the 1850 PI control simulation, which causes the SOM to start out at a cooler state. Not only does LENS-SOM start out cooler but it remains cooler over the length of the simulation (see Fig. S1 in the online supplemental material). All simulations fail to capture the 1940s warming in one of the observational datasets (ERSSTv5) but this simulated–observed discrepancy is not evident in the other observational datasets (COBE SST2 and HadISST). The absence of volcanic aerosol forcing and increasing trend in both CO_2 and solar forcing from the late 1920s to 1960 is thought to contribute to the AMV warm period (Birkel et al. 2018). Following this relatively clean (i.e., low aerosol) period and warm phase of the AMV, a series of strong volcanic eruptions beginning with Mount Agung in 1963 and higher anthropogenic aerosol loading are thought to contribute to the AMV cool phase that lasted from 1960 to the mid-1990s (Otterå et al. 2010; Wang et al. 2017; Birkel et al. 2018; Booth et al. 2012; Bellomo et al. 2018; Bellucci et al. 2017; Undorf et al. 2018; Watanabe and Tatebe 2019; Qin et al. 2020). In the model simulations the AMV cool period occurs a decade earlier (in the 1960s) compared to the observed AMV minima found in the 1970s. These results are similar to the CMIP5 models (Murphy et al. 2017). The post-1990s warming is better represented in the LENS-SOM ensemble mean compared to the LENS-FC and LENS-LME, which typically show a leveling off of the warming or even a cooling beginning near year 2000, whereas in observations the trend continues upward.

a. *AMV phasing*

Figure 2 shows the Pearson correlation coefficient between the observations and each ensemble member (open marker) and the EM (enlarged and filled in marker; this is the forced AMV)

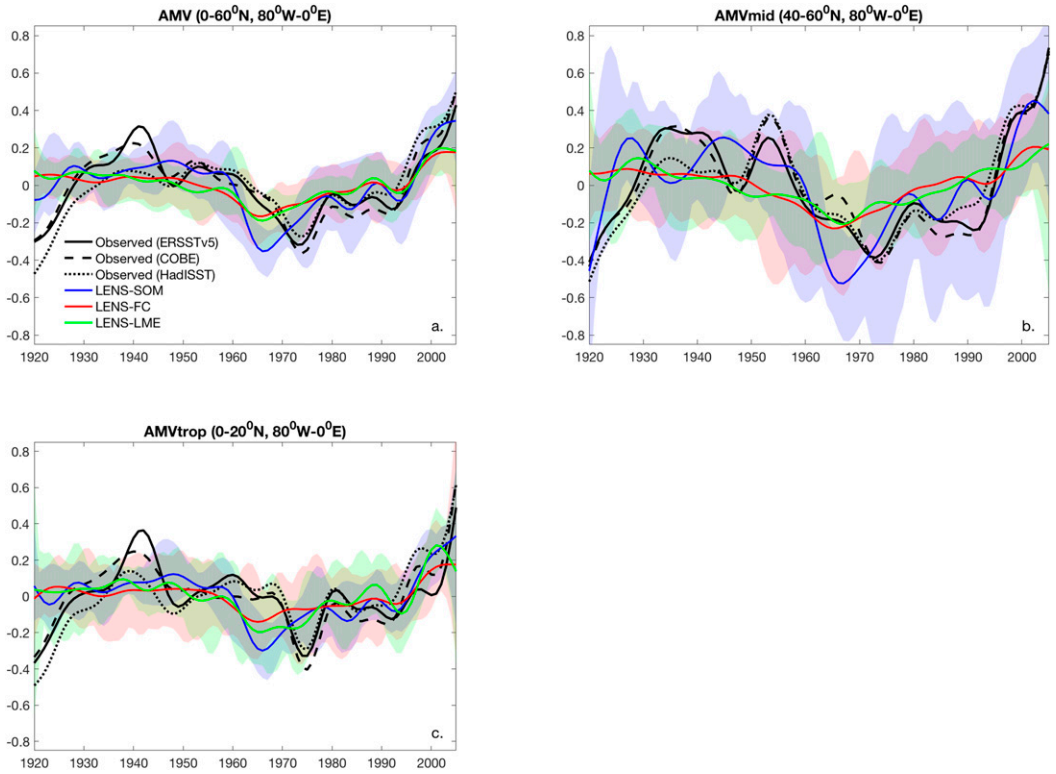

FIG. 1. Time series of 10-yr low-pass filtered and detrended (a) AMV (0°–60°N, 80°W–0°), (b) AMVmid (40°–60°N, 80°W–0°), and (c) AMVtrop (0°–20°N, 80°W–0°) over the historical period from 1920 to 2005. The solid red line shows the LENS-FC ensemble mean and the LENS-FC ensemble spread is depicted with light red shading. The solid blue line shows the LENS-SOM ensemble mean and the LENS-SOM ensemble spread is depicted with light blue shading. The solid green line shows the LENS-LME ensemble mean and the LENS-LME ensemble spread is depicted with light green shading. The observed ERSSTv5 (solid black line), observed COBE SST2 (dashed black line), and observed HadISST (dotted black line) AMV time series are also drawn.

for the periods 1920–2005, 1930–2005, 1940–2005, 1950–2005, and 1960–2005 for both the unfiltered and 10-yr LP filtered AMV index. While the LENS-FC correlations are typically lower than LENS-SOM for the full period, they become closer if the first decade, which is too warm in the LENS-FC, is removed. For 1930–2005 the LP filtered EM correlation coefficients increase to 0.75 in LENS-FC ($p < 0.05$, Neff = 5) and 0.80 in LENS-SOM ($p < 0.02$, Neff = 6). The unfiltered correlation coefficients for the EM are also similar between the models (0.60 and 0.62 for the FC and SOM, respectively, $p < 0.10$). Thus, adding the more realistic physics of an active ocean that is able to respond to external forcing makes the model's simulated time history slightly less like the observed AMV time series. As expected, we found increasing correlation coefficients for longer filter lengths and, in general, over time. The LENS-SOM values remained slightly higher than those for LENS-FC.

Historical forcing makes the time history of the simulated AMV more realistic in both ensembles, as evidenced by large filtered correlations of observations with the EM compared to the lower correlations generally found in each member. While the mean correlation coefficients over all ensemble members are 0.59 for LENS-FC and 0.67 for LENS-SOM, respectively, the mean correlations between observations and the implied internal AMV (i.e., the AMV in each ensemble member minus the EM) are near zero for both models (not shown) as well as for the LME. This was expected, since on average random noise is uncorrelated with any signal. We found similar results in the PI simulations when we calculated the correlation coefficients between the observed 1930–2005 LP filtered AMV and 76-yr chunks of the 1801-yr-long PI-FC simulation (123 r values) and the 901-yr-long PI-SOM simulation (58 r values) separated by 14-yr intervals to account for autocorrelation. The mean correlation coefficients are -0.011 ($2\sigma = 0.621$) and -0.002 ($2\sigma = 0.533$), and the maximum correlation coefficients are 0.62 and 0.47, in PI-SOM and PI-FC, respectively. Thus, the chance that the correlation coefficients in the PI simulations are equal or greater than the EM is less than 1 in 123 in the FC and less than 1 in 58 in the SOM. This indicates that internal variability alone is very unlikely to produce a time history consistent with observations [as previously argued in Murphy et al. (2017)]. Correlation coefficients are roughly as high in AMVmid as they are for AMV (Fig. 2b), but are much lower for AMVtrop (Fig. 2c). It is well known that this region is influenced from outside the Atlantic, especially by ENSO. Even models that simulate these physics cannot be expected

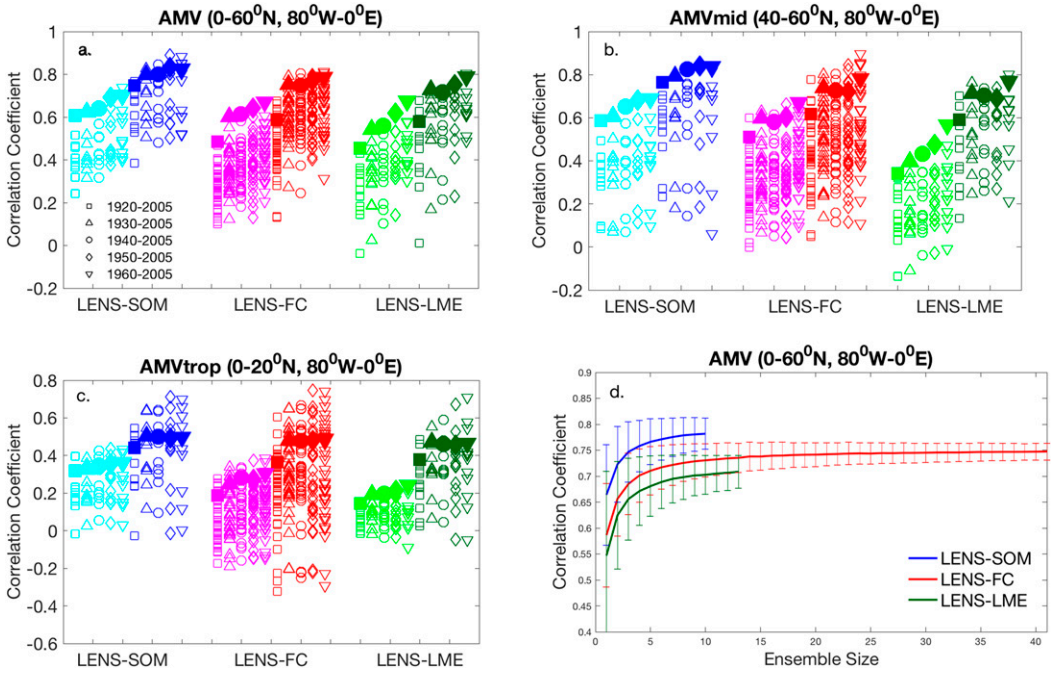

FIG. 2. Correlation coefficients between LENS-SOM (left group in each panel), LENS-FC (center group), and LENS-LME (right group) and the observed 1920–2005 (squares), 1930–2005 (upward-pointing triangles), 1940–2005 (circles), 1950–2005 (diamonds), and 1960–2005 (downward-pointing triangles) for the unfiltered and detrended (cyan, magenta, and green) and 10-yr low-pass filtered and detrended (blue, red, dark green) (a) AMV (0°–60°N, 80°W–0°), (b) AMVmid (40°–60°N, 80°W–0°), and (c) AMVtrop (0°–20°N, 80°W–0°). Each individual member's correlation coefficient is shown as an open marker and the ensemble mean (EM) value is shown as an enlarged and filled in marker. Note that the *y*-axis limit is different in (c). (d) Correlation coefficients vs ensemble size $n = 1{:}N$, where $N = 10$ in LENS-SOM, $N = 41$ in LENS-FC, and $N = 13$ in LENS-LME over the period from 1930–2005. The solid line indicates the mean correlation coefficient and the error bars their standard deviation calculated from 10 000 random samples of size n.

to duplicate the phasing of the observations. And ENSO-like low-frequency variability can arise in slab models as well (Dommenget 2010; Clement et al. 2011).

If we include the entire simulation (1920–2005), the ensemble mean correlation with observations in the LENS-FC goes down from 0.75 to 0.59 ($p < 0.20$, Neff = 5). Our initial hypothesis, which we subsequently rejected, was that errors in the ocean initial conditions (which are the same across all ensemble members in LENS-FC) degrade the correlation, but that this effect is essentially removed after discarding the initial 10 years as model spinup. We rejected it on the basis of the 13-member Last Millennium Ensemble (LENS-LME; green markers) that covers the period from 850 to 2005. The range of correlation values for 1920–2005 within these alternate ensembles are similar to LENS-FC and they all exhibit the same the same warm bias around 1920 (green and red lines in Fig. 1). This suggests that it was not the particular initialization of the LENS-FC historical runs, but instead that the coupled model is too warm in 1920. It is an open question on what caused the early-twentieth-century cool phase of the AMV. It has been argued that this unusual cold coincided with higher volcanic activity in the late nineteenth and early twentieth centuries (Birkel et al. 2018), and it is possible that the forcing imposed in the model underestimates the true forcing. It is also possible that the cold state in the early twentieth century was produced by internal processes in the climate system, such as a strong positive phase of the NAO (Visbeck et al. 2001; O'Reilly et al. 2019). However, Bellomo et al. (2018) showed that the none of the LME members produce the persistent cold phase in the early twentieth century, suggesting that this cold phase is unlikely to arise from internal processes. However, there is the possibility that internal variability in this model is too weak.

An ensemble size of 10 appears to be large enough to diagnose the role of external forcing on AMV phasing. In Fig. 2d, we analyze the impact of different ensemble sizes in LENS-SOM, LENS-FC, and LENS-LME on the EM LP filtered correlation coefficients for the period 1930–2005. For each ensemble size (n) we choose 10 000 random draws with replacement of n ensemble members to create an ensemble mean. We do this for 1 to 10 members of LENS-SOM (blue line), for 1 to 41 members of LENS-FC (red line), and for 1 to 13 members of LENS-LME (green line). The error bars represent the standard deviation of each ensemble mean value over the 10 000 random samples. Initially, there is a

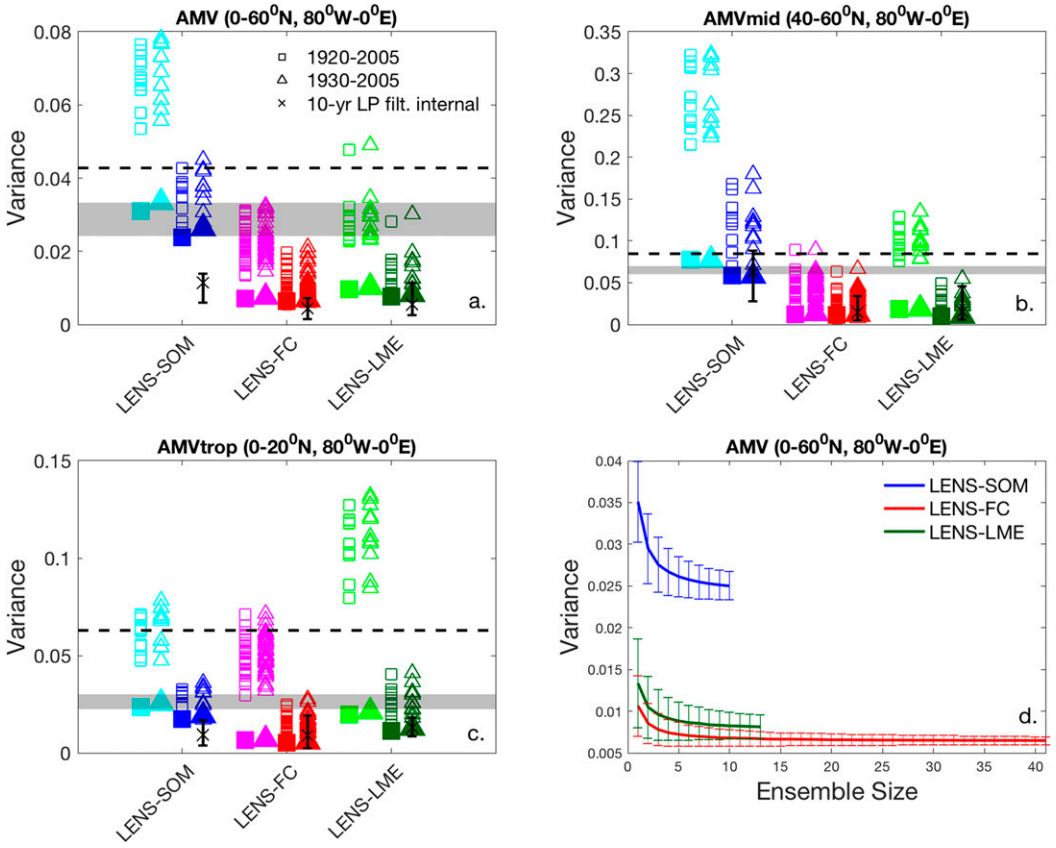

FIG. 3. Unfiltered and detrended (cyan, magenta, green) and 10-yr low-pass filtered and detrended (blue, red, dark green) (a) AMV (0°–60°N, 80°W–0°), (b) AMVmid (40°–60°N, 80°W–0°), and (c) AMVtrop (0°–20°N, 80°W–0°) variance for the LENS-SOM (left group in each panel), LENS-FC (center group), and LENS-LME (right group) over the period 1920–2005 (squares) and 1930–2005 (triangles). Each individual member's variance is shown as an open marker and the ensemble mean (EM) value is shown as an enlarged and filled in marker. The black × and whiskers denotes the 10-yr low-pass filtered and detrended variance of the internal AMV. The internal AMV is calculated by subtracting the ensemble mean (EM) AMV from each individual member. The observed 10-yr low-pass filtered variance over the 1930–2005 period is represented by the gray horizontal bar, which represents the variance calculated from three datasets: the ERSSTv5, COBE SST2, and HadISST. The unfiltered ERSSTv5 variance (1930–2005) is indicated by the horizontal, dashed black line. Note that the *y*-axis limits vary among (a)–(c). (d) Variance vs ensemble size $n = 1{:}N$, where $N = 10$ in LENS-SOM, $N = 41$ in LENS-FC, and $N = 13$ in LENS-LME over the period from 1930–2005. The solid line indicates the mean correlation coefficient and the error bars their standard deviation calculated from 10 000 random samples of size n.

rapid increase in the ensemble mean correlation coefficient. More than 90% of the total enhancement in correlation coefficient occurs within five members of the full ensemble. In the 41-member ensemble of LENS-FC, the correlation coefficient between the EM and the observed AMV is 0.75 (±0.016). In the 10-member ensemble of LENS-SOM, the average EM correlation coefficient is 0.78 (±0.031). To better compare the ensemble values, we examine 10 000 random selections of 10-member subsets within LENS-FC and LENS-LME. The average correlation coefficient of these 10 000 10-member ensembles of LENS-FC is 0.73 (±0.033) and in LENS-LME is 0.70 (±0.038). This shows that after neglecting the first decade in LENS-FC, the forced responses in all the ensembles are not distinguishable at about the 95% level. Using the 1920–2005 period, the mean correlation of our 10 000 10-member ensembles is 0.57 (±0.044) in LENS-FC and 0.56 (±0.042) in LENS-LME. LENS-SOM has a higher correlation of 0.73 (±0.032).

b. *AMV magnitude*

We find that the AMV amplitude is reduced when we add an active ocean. Figure 3 shows that in all three averaging regions, at all time scales, the North Atlantic SST variance is larger in LENS-SOM than in LENS-FC.

First, we will discuss the unfiltered data: Outside of the tropics, in almost all individual cases, the LENS-SOM unfiltered SST variance is typically higher than the unfiltered observed (black dashed line; the *F* test shows significant difference at the 95% confidence level), while the LENS-FC unfiltered SST variance is lower than the unfiltered observed

(with the exception of one member in the AMVmid region, again significantly different at the 95% confidence level) (Fig. 3). In the tropical region LENS-FC is likely strengthened by its active dynamics, including influences from ENSO, while LENS-SOM variance is diminished there by the lack of an active ocean, which may help to make the two ensembles more similar to each other. There are some differences between LENS-LME and LENS-FC variance, particularly in the tropics. These are the same model version but LENS-LME is at coarser resolution.

It has been argued that that on decadal and longer time scales, SST variability is driven through changes in oceanic heat transport (Bjerknes 1964; Buckley et al. 2014; Zhang 2017). Thus, one might expect that with an active ocean, LENS-FC would have greater low-frequency variability than the LENS-SOM. This did not occur. On multidecadal time scales, the LENS-SOM ensemble spread typically overlaps with the observed SST variance (Fig. 3, gray bar), although the mean variance over all ensemble members (MOE variance) is clearly higher for AMV and for AMVmid and only slightly higher for AMVtrop (see Table 1 for numerical values). In contrast, the LENS-FC MOE variance is far weaker than observations in all three regions (Fig. 3 and Table 1). An F test shows that at the 95% confidence level the LENS-SOM MOE variance is indistinguishable from the observed variance in the AMV and AMVtrop but unequal to the observed in the AMVmid, whereas the LENS-FC MOE variance is unequal to the observed variance in all three regions (all at the 95% confidence level). As is true for individual ensemble members, the EM variance (i.e., the externally forced variance) is far smaller in LENS-FC than in LENS-SOM; the active ocean damps the forced response. The mean internal multidecadal variability (calculated as the MOE variance minus the EM variance) is also weaker for LENS-FC compared to LENS-SOM in AMV and AMVmid and about the same in AMVtrop (values in Table 1). However, as a percentage of the total variance the internal variance is higher in LENS-FC than in LENS-SOM in AMV and AMVtrop (see Table 1). Thus, while the comparison of the two ensembles indicates that the active ocean damps both forced and internal variability, this damping is relatively greater for the forced component than for the internal component. Since the slab ocean is too variable compared with observations some damping from an active ocean is clearly required, but LENS-FC greatly overdoes it: for AMV and AMVmid the average LENS-FC ensemble member has less than half of the observed variance. All of the individual members show less variance than the observed AMV. The PI AMV variances are both lower than in the corresponding historically forced runs (which average 0.014 and 0.006 in the PI-SOM and PI-FC ensembles, respectively). This indicates that historical forcing enhances the variance consistent with Murphy et al. (2017). As in Fig. 2d, in Fig. 3d we show that the LENS-SOM ensemble is large enough to evaluate the ensemble mean variance relative to the 41-member LENS-FC, so the difference in variance is due to model physics, and not the different ensemble size.

We wish to see how much of the total variability is attributable to multidecadal variability. To do so we examine the ratio of low-frequency (i.e., 10-yr LP filtered) to total (unfiltered) variance in our ensembles and in observations for the period 1930–2005 (although conclusions are similar to the full period), as well as for the unforced PI simulations. In observations, ERSSTv5 and HadISST produce similar ratios (0.62 and 0.64), while COBE SST2 gives a value of 0.72. The mean ratio over all ensemble members of multidecadal to total variance is 0.54 in LENS-SOM (with a range of: 0.50–0.61) and 0.48 in LENS-FC (with a range of 0.27–0.67). The observations have a higher percentage of low-frequency variability, but we cannot say if this results from a stronger response to external forcing or from greater multidecadal internal variability. If we remove the forced variability from the models by subtracting the EM, these ratios are comparable for the two models: 0.31 for the LENS-SOM (with a range of 0.22–0.41) and 0.28 for the LENS-FC (with a range of 0.14–0.43). The PI ratios are just slightly higher: 0.37 for PI-SOM and 0.33 for PI-FC. Thus, either without forcing or with the forced signal removed, the model with an active ocean has about the same percentage of its variability at low frequencies as the model without an active ocean, indicating that in this model, the ocean has a limited role in enhancing multidecadal variability of the AMV. An interpretation consistent with these findings is that the low-frequency variability is just the low-frequency component of the response to forcing from the NAO, which Wunsch (1999) and Stephenson et al. (2000) have shown to be indistinguishable from white noise. Moreover, we have seen that the AMV variance is much weaker in the model with an active ocean, indicating that the strongest effect of the model ocean is to dampen variability, whether it is forced, total internal, or low-frequency internal variability. The tendency for the coupled model to reduce multidecadal variability compared to the SOM, as evidenced by the wider range in ratios, was unexpected and difficult to reconcile with its more complete physics. Kim et al. (2018) proposed that the weak multidecadal variability in both the AMV and AMOC in CESM was due to weak multidecadal NAO variability driven by deficiencies in the atmospheric model or the coupling between the model components. However, the NAO power spectra are indistinguishable between LENS-SOM and LENS-FC, casting doubt that the atmospheric model is to blame (Fig. S2).

The averaged SST metrics in Fig. 3 show that the absence of ocean processes produces too much SST variability, while the net effect of inclusion of ocean processes in this model produces too little variability. If the model's AMOC time scale and magnitude are unlike the real world, this has the potential to dampen the AMV response since the active ocean will move warm water northward at the wrong time. This damping role for the ocean was also found to weaken AMV power in PI simulations of CESM (Garuba et al. 2018). A net damping may arise from vertical entrainment of water from beneath the thermocline (Frankignoul 1985). Efficient vertical turbulent mixing can thus modulate SSTs anomalies in LENS-FC while this mechanism, in addition to advection and diffusion, are all absent in LENS-SOM. However, there is considerable spatial structure to the differences in simulated variance between these two models (Fig. 4) and between each model and observations (Fig. S3). In particular, there are parts of the North

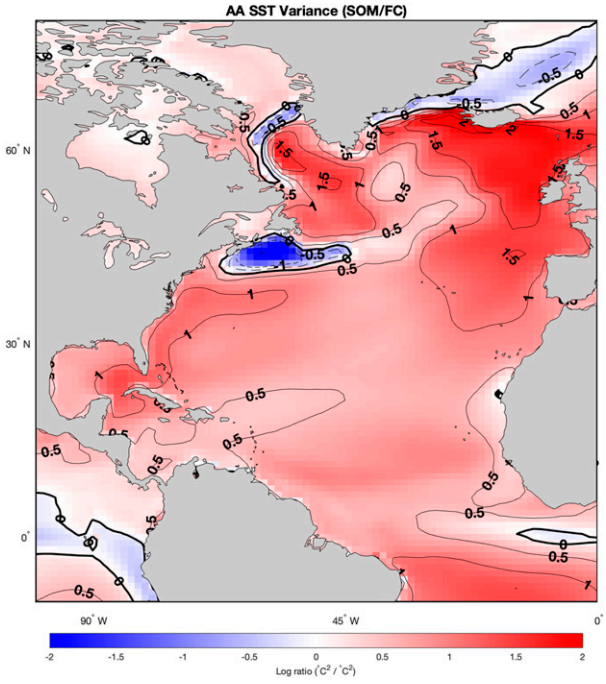

FIG. 4. Ratio of the log of the annual average (AA) SST variance between LENS-SOM and LENS-FC from 1930 to 2005. The map shows the log ratio of the average of all individual ensemble members, i.e., the value at each grid point is computed in each ensemble member before averaging. For reference, a log ratio value of 1 implies that the variance in LENS-SOM is 10 times that in LENS-FC. Likewise, a log ratio of −1 implies that the variance in LENS-FC is 10 times than in LENS-SOM.

Atlantic where LENS-FC does simulate more low-frequency variability than LENS-SOM. Zhang (2017) argued that in some regions, the effect of ocean dynamics is large enough to overcome the overall damping effect of ocean mixing. This effect is missed in metrics that average over large regions of the ocean.

It is difficult to discount the tropical variance in the coupled simulations as underrepresenting observed variability, as current models are known to be deficient in the representation of known local feedback mechanisms (e.g., Martin et al. 2014). These deficiencies include feedbacks between low clouds (Bellomo et al. 2016; Martin et al. 2014), dust aerosols (Evan et al. 2012; Yuan et al. 2016; Wang et al. 2012), and wind–evaporation–SST (WES; Amaya et al. 2017) feedback, which all may play a role by enhancing the model's temperature responses and may be resolution dependent. These deficiencies may prove less fatal in LENS-SOM due to the absence of ocean damping, whereas the perhaps realistic damping in LENS-FC may expose these biases.

c. *Partitioning of variance in the observations*

Knowing that the LENS-FC has the higher percentage of variance due to internal variability does not tell us which is closer to the observed, since we do not know how the observed variance is divided. We can, however, make an estimate for the percentage of observed variance that is forced. Suppose we regress the observed time series on the EM time series for LENS-SOM—the time series of the forced part of LENS-SOM. Since the correlation coefficients are 0.80, 0.79, and 0.50 for AMV, AMVmid, and AMVtrop, the observed variance accounted for is 64%, 62%, and 25%, respectively (based on r^2). This is for the period 1930–2005, a period with strong forcing from industrial aerosol and greenhouse gas emissions; other periods would have different percentages as the strength of external forcing varies in time. If we had regressed on the LENS-FC EM correlation coefficients (0.75, 0.74, and 0.48; from Table 1) the percentages would be lower: 56%, 55%, and 23%.

In the models averaging over the ensemble members strongly reduces the internal variability and gives a good estimate of the part attributable to forcing. However, we have only a single realization of the observed, so while the numbers we gave are best estimates of the correlation coefficients of the models with the observed, they are only estimates, which are subject to the caveat that internal variability will be different in each time series (whether observed or from the model). In Fig. 2 note the range of correlations with single ensemble members. The spread is larger for LENS-FC and LME than for LENS-SOM, reflecting the fact that for the latter a larger percentage of the variance is forced and so common to all ensemble members. But even for the LENS-SOM the range is substantial. These differences in ensemble spread introduce uncertainties, but they do not bias the estimates of the percentage of variability due to external forcing either high or low. On the other hand, these results are biased low because the observations and forcing are not known accurately and the model response to the forcing is not completely correct. Hence even if we had a large ensemble of nature's response to the forcing from 1930 to 2005, the mean of this ensemble would not correlate perfectly with the model mean so the explained variance would not be 100% of the true forced variance. In sum, we estimate that in the observations roughly two-thirds of the variance of the AMV is externally forced in recent decades. This estimate has high uncertainty, and is more likely to be too low than to be too high.

d. *Spatial pattern*

Here, we examine the AMV associated SST and atmospheric circulation patterns by plotting the regression of SST, SLP, and surface winds anomalies on the AMV over the period 1920–2005 in reanalysis/observational data (Fig. 5a), the LENS-SOM and LENS-FC MOE (Figs. 5b,e), EM (Figs. 5c,f), and the PI simulation (Figs. 5d,g). As before, MOE denotes the mean over all ensemble members. Observations include the impact of both internal and external forcing as do the MOE panels, while the EM panels show the modeled approximate response to external forcing alone, and the PI panels show the pattern associated with internal variability alone. In observations, when the AMV is positive there is a warming of the entire basin with a horseshoe-shaped pattern of warming that wraps around the eastern basin. SLP shows a negative NAO-like mode over the North Atlantic that involves a weakening of both the subtropical high (STH) and Icelandic low. Associated with these pressure changes are weaker trade winds and weaker westerlies. Calculating the regressions based on the

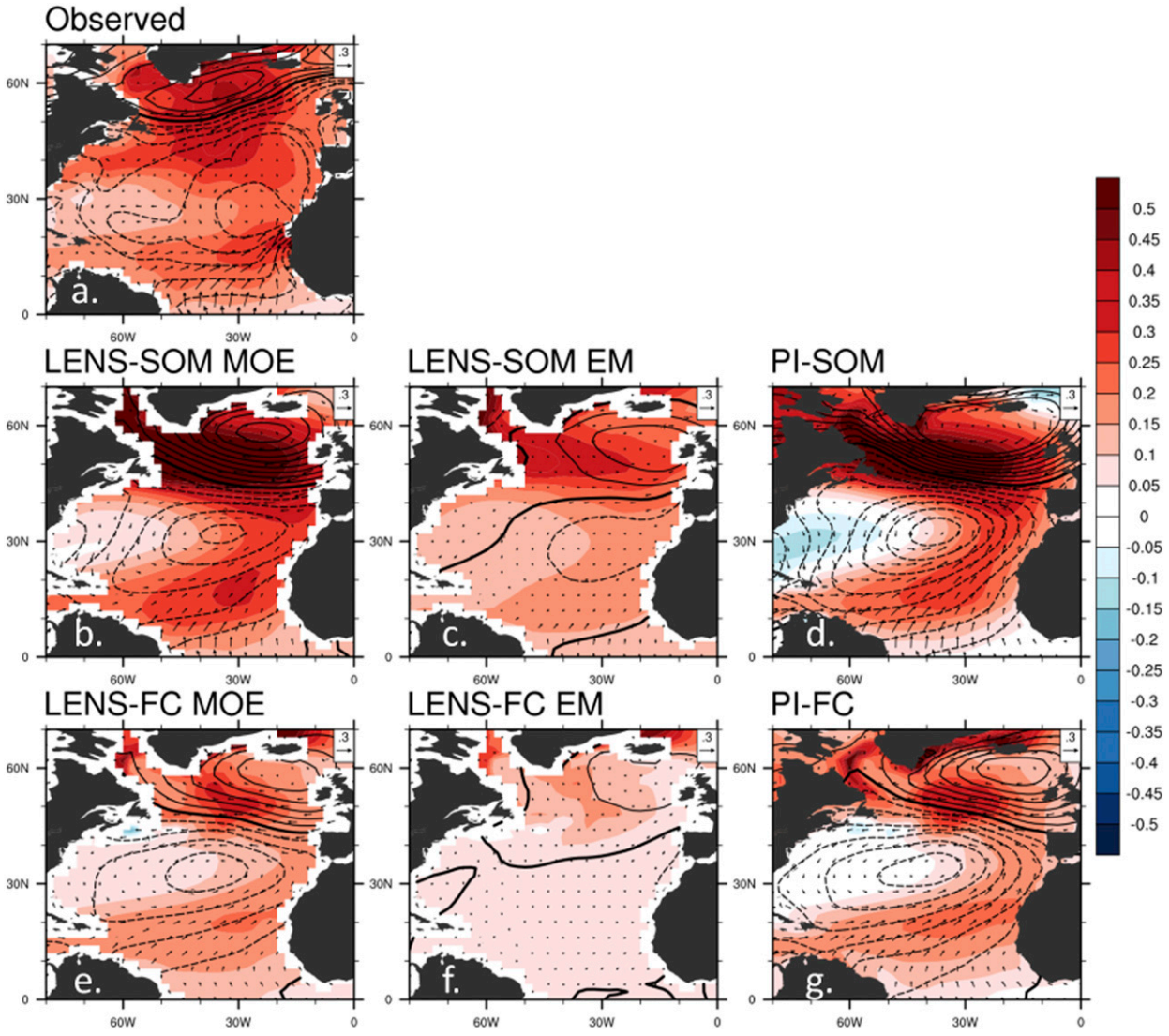

FIG. 5. (a) Regression of SST (shaded), SLP (contours), and surface winds (vectors) on the standardized AMV index from 1920–2005 in observations, the mean over all ensemble members (MOE) in (b) LENS-SOM and (e) LENS-FC, the ensemble mean (EM) in (c) LENS-SOM and (f) LENS-FC, and (d) PI-SOM and (g) PI-FC. All data are detrended and low-pass filtered with a 10-yr Lanczos filter. Units are K, hPa, and m s^{-1} per unit of standard deviation of the AMV index for SST, SLP, and wind, respectively. SLP is plotted from −0.4 to 0.4 hPa with intervals of 0.05 hPa. Negative pressures are shown as dashed contour lines. Surface winds were interpolated to 1000 hPa. The reference vector length of 0.3 m s^{-1} is plotted in the upper-right corner of each panel.

1930–2005 period does not change the simulated patterns but does lead to slight changes in the observed wind pattern, which is based on the twentieth-century reanalysis data, likely due to the large uncertainties in the early-twentieth-century wind observations (Wohland et al. 2019).

Since both the PI-SOM and PI-FC simulations show the characteristic spatial pattern of the AMV, this is an internal mode of variability that occurs even without interactive ocean circulation. It is clear that the SOM is able to simulate this spatial pattern, but this does not, on its own, exclude a role for the ocean (e.g., Zhang et al. 2016; Zhang 2017; O'Reilly et al. 2016). Booth et al. (2012) showed a horseshoe pattern of warming in their forced coupled model, although Watanabe and Tatebe (2019) showed just a basinwide warming with little structure. In CESM, the forced horseshoe pattern is evident in the EM LENS-SOM (Fig. 5c) but not in the EM LENS-FC (Fig. 5f). The EM LENS-SOM also shows a reduction in the North Atlantic subtropical high evident in the surface winds and SLP. The MOE, the mean over all ensemble members, includes both forced and internal variability. In both models the MOE show similar patterns although the atmospheric circulation anomalies are stronger in LENS-SOM compared to LENS-FC (cf. Figs. 5b and 5e). In this model, the forced response is much weaker when we include active ocean dynamics (cf. Figs. 5f and 5c). Air–sea feedbacks discussed earlier likely play a large role in the model's tropical AMV pattern.

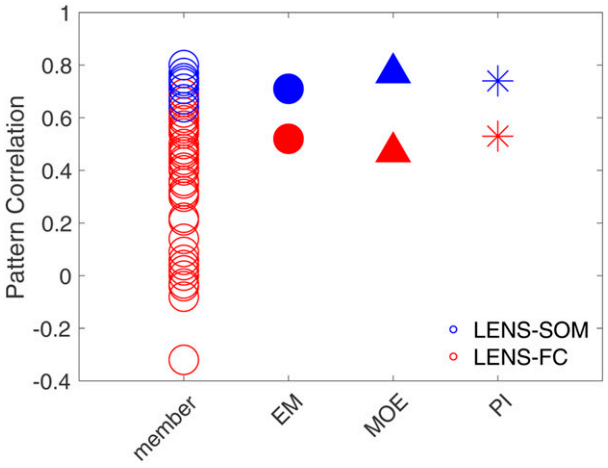

FIG. 6. Pattern correlation between the observed and simulated SST regression patterns (see Fig. 5) for each individual ensemble member (open circle), the ensemble mean (EM) (filled circle), mean of all ensemble members (MOE) (filled triangle), and PI simulation (asterisk) for the LENS-SOM (blue) and LENS-FC (red).

To quantify whether the addition of an ocean improves the spatial pattern of the modeled AMV, we calculate the Pearson product-moment coefficient of linear correlation (centered pattern correlation) between the observed (Fig. 5a) and simulated 10-yr LP filtered regression patterns (SST only) over the period 1920–2005 (Fig. 6). These results show that the pattern of SST changes associated with the AMV are better simulated in the LENS-SOM (blue markers) compared to LENS-FC (red markers). All but one member of the 41-member LEN-FC ensemble (range: −0.21 to 0.76) show spatial correlation values less than all the members of the LENS-SOM ensemble (range: 0.68 to 0.84). The MOE and EM spatial correlations (filled in triangle and circle, respectively) for LENS-SOM are 0.78 and 0.74; for the LENS-FC the corresponding values are 0.52 and 0.53. In addition, the PI spatial correlations are higher in SOM compared to the FC (0.74 vs 0.57). This indicates that the addition of ocean physics has made the spatial structure of the model's internal mode less realistic, with or without forcing. The difference in the AMV pattern in response to forcing is not an artifact of the ensemble size, since 10-member LENS-FC subsets results in the same pattern shown in Fig. 5f. These results are essentially the same whether we use the period 1920–2005 or 1930–2005 (not shown); unlike the time history, the spatial pattern is unaffected by the first decade of the simulation.

Overall, our interpretation is that the observed spatial pattern can be thought of as an internal mode of variability. External forcing has a large influence on its amplitude, but not its spatial structure except that the forcing for a positive AMV adds a warming almost everywhere. However, this should be tested in other models as it may be a model-specific result.

3. Discussion and conclusions

We compared existing ensembles of radiatively forced simulations with a fully coupled model (LENS-FC and LENS-LME) to those with the same atmospheric model coupled to a motionless slab ocean (LENS-SOM). Previous modeling studies have shown important lead–lag relationships between the AMOC and North Atlantic SST and therefore the LENS-FC configuration, with its more complete and realistic physics, was expected to outperform the LENS-SOM. On the contrary, we find that adding ocean dynamics, mixed layer physics, and a treatment of salinity (i.e., coupling the model) renders the model less able to reproduce the characteristics of the observed AMV. The slab ocean model better reproduces the observed historical (1920 or 1930 to 2005) AMV, according to the metric of correlations with observed SST time history and spatial pattern. In addition, while the slab ocean model SST variance is too high, the active ocean reduces it to be less than the observed variance. That is, a primary effect of the ocean in this model is to strongly damp the variance over most of the North Atlantic (Fig. 4). In addition, while the active model ocean damps both forced and internal variability, this damping is relatively greater for the forced component than for the internal component, resulting in relatively less low-frequency variability.

Historical forcing makes the simulated AMV more like the observed AMV. External forcing is not necessary to reproduce the AMV's characteristic horseshoe pattern of warming. It is generated by atmospheric internal variability as shown in PI simulations in Clement et al. (2015), and again here in Fig. 5. The model response to external forcing appears as an addition of an overall warming for positive AMV phases.

To summarize, we return to our original research questions:

- How much of the observed AMV is forced by historical forcing and how much is internal to the climate system?

We estimate that about 2/3 of the LP filtered variance of the AMV averaged over the basin can be explained by variable external forcing, leaving 1/3 for internal ocean and atmospheric variability. This is a best estimate within this model, but comes with sizable uncertainty.

- How is the historically forced response made more realistic by adding a fully coupled ocean?

Unfortunately, adding the more realistic physics of an active ocean does not improve this model's ability to simulate either the time history or the spatial pattern of the observed AMV. If anything, it shows larger biases compared to the same atmosphere coupled to a slab ocean.

- How much more internal multidecadal variance is added by including an ocean?

Inclusion of ocean processes lowers the overall variability below observations. Ocean dynamics does not increase the ratio of multidecadal to total variability, with or without forcing. Without forcing, LENS-FC and LENS-SOM both produce too little multidecadal variance, evidence that the observed AMV has a sizable forced component.

While the CESM LENS-FC and LENS-SOM simulations provide a clear framework for evaluating the roles of external forcing and ocean circulation on the AMV, there are some important caveats. There is the question of whether these

results are unique to CESM. In particular, would other climate models simulate a net damping role for the ocean? Clement et al. (2015) showed that for PI runs not all models simulate weaker variability in the coupled than in SOM versions. In the six CMIP3 models that had sufficiently long simulations with a SOM, only two (HadGEM and MRI) produced more low-frequency variability in the SOM than coupled, while two produced less variability in the SOM versions (GFDL CM2.1 and GISS ModelE) and two were about the same, one of which was NCAR's CCSM4. These were PI control simulations, so that analysis did not test the role of the ocean with evolving forcing. Anthropogenic aerosols may force the ocean circulation in climate models (Cheng et al. 2013; Menary et al. 2020), although in our analysis there is no evidence that this mechanism helped to improve LENS-FC relative to LENS-SOM. Perhaps this failure is specific to this model. There is also the question of how this particular experimental design in this model relates to other ways of assessing the relative roles of the ocean, atmosphere, and external forcing, such as those presented in Garuba et al. (2018), Li et al. (2020), and Patrizio and Thompson (2021). A more thorough review of these various methods is clearly called for. It is also important to relate the mechanisms of the AMV to the impacts. For instance, Klavans et al. (2020) showed that external forcing was required to explain the impact of AMV on summertime Florida rainfall, which was one of the first impacts that motivated the study of the AMV (Enfield et al. 2001). On the other hand, Qasmi et al. (2020) argued that ocean-driven turbulent heat fluxes associated with the AMV were important for simulating the climate impacts over Europe.

The SOM is not a perfect tool. In existing configurations of the SOM, an annual mean mixed layer depth is prescribed. Here it was based on an 1850 PI control simulation of the fully coupled model. This has the limitation that seasonal variations in the mixed layer depth, which are large in the North Atlantic and important for determining the persistence of SST and upper-ocean heat content variability (Buckley et al. 2019), are not included. There is a clear need for a better tool for evaluating the role of the ocean circulation in climate variability, for example, by having a version of the SOM that includes an entraining mixed layer. This would help in deducing the role of vertical mixing on controlling the magnitude of the AMV variance.

The forced signal in the LENS-SOM explains about 2/3 of the observed variance, leaving 1/3 unaccounted for. Some of that is likely due to imperfect knowledge of the forcing and error in the model response to the forcing. Some is due to internal, random, atmosphere and ocean "noise." It is far from clear how much of this noise is predictable. The challenge for decadal SST prediction is to 1) properly account for the role of external forcing in producing persistent changes (e.g., Karspeck et al. 2015), 2) determine how much of the internal signal comes from the atmosphere, which can introduce long persistence in regions of deep mixed layers (Buckley et al. 2019), and then 3) determine whether what is left offers meaningful predictability from initialized ocean states.

Acknowledgments. The authors thank the editor and three anonymous reviewers for their thorough comments on this manuscript. The 20th Century Reanalysis V2c data, NOAA ERSSTv4, NOAA ERSSTv5, and the COBE-SST2 data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at https://www.esrl.noaa.gov/psd/. Support for the Twentieth Century Reanalysis Project version 2c dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. We also acknowledge the CESM Large Ensemble Community Project and supercomputing resources provided by NSF/CISL/Yellowstone. The LENS-SOM historical ensembles were done with the support of NSF Grant 1735245 for which supercomputing resources were provided by NSF/CISL/Cheyenne. Data from LENS-SOM will be made available on supercomputing resources from NCAR. We gratefully acknowledge support for this work from NSF Climate and Large-Scale Dynamics Grants AGS 1735245 and AGS 1650209 and from NSF P2C2 Grant AGS 1703076.

REFERENCES

Amaya, D. J., M. J. DeFlorio, A. J. Miller, and S.-P. Xie, 2017: WES feedback and the Atlantic meridional mode: Observations and CMIP5 comparisons. *Climate Dyn.*, **49**, 1665–1679, https://doi.org/10.1007/s00382-016-3411-1.

Anderson, R. L., 1942: Distribution of the serial correlation coefficient. *Ann. Math. Stat.*, **13**, 1–13, https://doi.org/10.1214/aoms/1177731638.

Bellomo, K., A. C. Clement, L. N. Murphy, L. M. Polvani, and M. A. Cane, 2016: New observational evidence for a positive cloud feedback that amplifies the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, **43**, 9852–9859, https://doi.org/10.1002/2016GL069961.

——, 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.

Bitz, C. M., K. M. Shell, P. R. Gent, D. A. Bailey, G. Danabasoglu, K. C. Armour, M. M. Holland, and J. T. Kiehl, 2012: Climate sensitivity of the Community Climate System Model, version 4. *J. Climate*, **25**, 3053–3070, https://doi.org/10.1175/JCLI-D-11-00290.1.

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

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.

Buckley, M. W., 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.

——, T. DelSole, M. S. Lozier, and L. Li, 2019: Predictability of North Atlantic sea surface temperature and upper-ocean heat content. *J. Climate*, **32**, 3005–3023, https://doi.org/10.1175/JCLI-D-18-0509.1.

Cane, M., A. 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.

Cheng, W., J. C. H. Chiang, and D. Zhang, 2013: Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. *J. Climate*, **26**, 7187–7197, https://doi.org/10.1175/JCLI-D-12-00496.1.

Clement, A., P. DiNezio, and C. Deser, 2011: Rethinking the ocean's role in the Southern Oscillation. *J. Climate*, **24**, 4056–4072, https://doi.org/10.1175/2011JCLI3973.1.

——, K. Bellomo, L. N. Murphy, M. A. Cane, T. Mauritsen, G. Rädel, and B. Stevens, 2015: The Atlantic Multidecadal Oscillation without a role for ocean circulation. *Science*, **350**, 320–324, https://doi.org/10.1126/science.aab3980.

Compo G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, https://doi.org/10.1002/qj.776.

Danabasoglu, G., S. G. Yeager, Y.-O. Kwon, J. J . Tribbia, A. S. Phillips, and J. W. Hurrell, 2012. Variability of the Atlantic meridional overturning circulation in CCSM4. *J. Climate*, **25**, 5153–5172, https://doi.org/10.1175/JCLI-D-11-00463.1.

Dawdy, D. R., and N. C. Matalas, 1964: Statistical and probability analysis of hydrologic data, part 3: Analysis of variance, covariance and time series. *Handbook of Applied Hydrology: A Compendium of Water Resources Technology*, V. T. Chow, Ed., McGraw-Hill, 8.68–8.90.

Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dyn.*, **16**, 661–676, https://doi.org/10.1007/s003820000075.

——, and F. Zeng, 2016: The impact of the North Atlantic Oscillation on climate through its influence on the Atlantic meridional overturning circulation. *J. Climate*, **29**, 941–962, https://doi.org/10.1175/JCLI-D-15-0396.1.

——, ——, L. Zhang, and 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 A. Phillips, 2017: An overview of decadal-scale sea surface temperature variability in the observational record. *CLIVAR Exchanges*, No. 72, International CLIVAR Project Office, Southampton, United Kingdom, 5 pp., https://www.cgd.ucar.edu/staff/cdeser/docs/deser.decadal_sst_variability_obs.pgcm17.pdf.

——, 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.

Dommenget, D., 2010: The slab ocean El Niño. *Geophys. Res. Lett.*, **37**, L20701, https://doi.org/10.1029/2010GL044888.

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.

Evan, A. T., G. R. Foltz, and D. X. Zhang, 2012: Physical response of the tropical–subtropical North Atlantic Ocean to decadal–multidecadal forcing by African dust. *J. Climate*, **25**, 5817–5829, https://doi.org/10.1175/JCLI-D-11-00438.1.

Folland, C. K., A. W. Colman, D. P. Rowell, and M. K. Davey, 2001: Predictability of northeast Brazil rainfall and real-time forecast skill, 1987–98. *J. Climate*, **14**, 1937–1958, https://doi.org/10.1175/1520-0442(2001)014<1937:PONBRA>2.0.CO;2.

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.

Frankignoul, C., 1985: Sea surface temperature anomalies, planetary waves, and air–sea feedback in the middle latitudes. *Rev. Geophys.*, **23**, 357–390, https://doi.org/10.1029/RG023i004p00357.

——, G. Gastineau, and Y.-O. Kwon, 2015: Wintertime atmospheric response to North Atlantic Ocean circulation variability in a climate model. *J. Climate*, **28**, 7659–7677, https://doi.org/10.1175/JCLI-D-15-0007.1.

——, ——, and ——, 2017: Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic multidecadal oscillation and the Pacific decadal oscillation. *J. Climate*, **30**, 9871–9895, https://doi.org/10.1175/JCLI-D-17-0009.1.

Garuba, O. A., J. Lu, H. A. Singh, F. Liu, and P. Rasch, 2018: On the relative roles of the atmosphere and ocean in the Atlantic multidecadal variability. *Geophys. Res. Lett.*, **45**, 9186–9196, https://doi.org/10.1029/2018GL078882.

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.

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.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1.

Huang, B., and Coauthors, 2015: Extended reconstructed sea surface temperature version 4 (ERSST.v4): Part I. Upgrades and intercomparisons. *J. Climate*, **28**, 911–930, https://doi.org/10.1175/JCLI-D-14-00006.1.

——, 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.

Hurrell, J. W., M. M. Holland, and P. R. Gent, 2013: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteor. Soc.*, **94**, 1339–1360, https://doi.org/10.1175/BAMS-D-12-00121.1.

Karspeck, A., S. Yeager, G. Danabasoglu, and H. Teng, 2015: An evaluation of experimental decadal predictions using CCSM4. *Climate Dyn.*, **44**, 907–923, https://doi.org/10.1007/s00382-014-2212-7.

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.

Kerr, R., 2000: A North Atlantic climate pacemaker for the centuries. *Science*, **288**, 1984–1985, https://doi.org/10.1126/science.288.5473.1984.

Kim, W. M., S. Yeager, P. Chang, and G. Danabasoglu, 2018: 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.

Klavans, J. M., A. C. Clement, and M. Cane, 2019: Variable external forcing obscures the weak relationship between the NAO and North Atlantic multi-decadal SST variability. *J. Climate*, **32**, 3847–3864, https://doi.org/10.1175/JCLI-D-18-0409.1.

——, ——, 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.

Knight, J. R., 2009: The Atlantic multidecadal oscillation inferred from the forced climate response in coupled general circulation models. *J. Climate*, **22**, 1610–1625, https://doi.org/10.1175/2008JCLI2628.1.

——, 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.

Knutson, T. R., F. Zeng, and A. T. Wittenberg, 2013: Multimodel assessment of regional surface temperature trends: CMIP3 and CMIP5 twentieth-century simulations. *J. Climate*, **26**, 8709–8743, https://doi.org/10.1175/JCLI-D-12-00567.1.

Larson, S. M., M. W. Buckley, A. C. Clement, 2019: Extracting the buoyancy-driven Atlantic meridional overturning circulation. *J. Climate*, **33**, 4697–4714, https://doi.org/10.1175/JCLI-D-19-0590.1.

Latif, M., N. Keenlyside, and J. Bader, 2007: Tropical sea surface temperature, vertical wind shear, and hurricane development. *Geophys. Res. Lett.*, **34**, L01710, https://doi.org/10.1029/2006GL027969.

Leith, C. E., 1973: The standard error of time-average estimates of climatic means. *J. Appl. Meteor. Climatol.*, **12**, 1066–1069, https://doi.org/10.1175/1520-0450(1973)012<1066:TSEOTA>2.0.CO;2.

Li, L., M. S. Lozier, and M. W. Buckley, 2020: An investigation of the ocean's role in Atlantic multidecadal variability, **33**, 3019–3035, https://doi.org/10.1175/JCLI-D-19-0236.1.

Mann, M. E., 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.

Martin, E. R., C. Thorncroft, and B. B. B. Booth, 2014: The multidecadal Atlantic SST–Sahel rainfall teleconnection in CMIP5 simulations. *J. Climate*, **27**, 784–806, https://doi.org/10.1175/JCLI-D-13-00242.1.

McCarthy, G. D., and Coauthors, 2015: Measuring the Atlantic meridional overturning circulation at 26°N. *Prog. Oceanogr.*, **130**, 91–111, https://doi.org/10.1016/j.pocean.2014.10.006.

Mecking, J. V., N. S. Keenlyside, and R. J. Greatbatch, 2014: Stochastically-forced multidecadal variability in the North Atlantic: A model study. *Climate Dyn.*, **43**, 271–288, https://doi.org/10.1007/s00382-013-1930-6.

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.

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.

O'Reilly, C. H., M. Huber, T. Woollings, and L. Zanna, 2016: The signature of low-frequency oceanic forcing in the Atlantic multidecadal oscillation. *Geophys. Res. Lett.*, **43**, 2810–2818, https://doi.org/10.1002/2016GL067925.

——, T. Woollings, and L. Zanna, 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.

Patrizio, C. R., and D. W. J. Thompson, 2021: Quantifying the role of ocean dynamics in ocean mixed-layer temperature variability. *J. Climate*, **34**, 2567–2589, https://doi.org/10.1175/JCLI-D-20-0476.1.

Qasmi, S., C. Cassou, and J. Boé, 2020: Teleconnection processes linking the intensity of the Atlantic multidecadal variability to the climate impacts over Europe in boreal winter. *J. Climate*, **33**, 2681–2700, https://doi.org/10.1175/JCLI-D-19-0428.1.

Qin, M., A. Dai, and W. Hua, 2020: Aerosol-forced multidecadal variations across all ocean basins in models and observations since 1920. *Sci. Adv.*, **6**, eabb0425, https://doi.org/10.1126/sciadv.abb0425.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A., 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, https://doi.org/10.1029/2002JD002670.

Robson, J., P. Ortega, and R. Sutton, 2016: A reversal of climatic trends in the North Atlantic since 2005. *Nat. Geosci.*, **9**, 513–517, https://doi.org/10.1038/ngeo2727.

Smeed, D. A., and Coauthors, 2014: Observed decline of the Atlantic meridional overturning circulation 2004–2012. *Ocean Sci.*, **10**, 29–38, https://doi.org/10.5194/os-10-29-2014.

Stephenson, D. B., V. Pavan, and R. Bojariu, 2000: Is the North Atlantic oscillation a random walk? *Int. J. Climatol.*, **20**, 1–18, https://doi.org/10.1002/(SICI)1097-0088(200001)20:1<1::AID-JOC456>3.0.CO;2-P.

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.

Tandon, N. F., and P. J. Kushner, 2015: Does external forcing interfere with the AMOC's influence on North Atlantic sea surface temperature? *J. Climate*, **28**, 6309–6323, https://doi.org/10.1175/JCLI-D-14-00664.1.

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.

——, S. J. Camargo, C. Li, and Y. Kushner, 2015: Natural and forced North Atlantic hurricane potential intensity change in CMIP5 models. *J. Climate*, **28**, 3926–3942, https://doi.org/10.1175/JCLI-D-14-00520.1.

Trenary, L., and T. DelSole, 2016: Does the Atlantic multidecadal oscillation get its predictability from the Atlantic meridional overturning circulation? *J. Climate*, **29**, 5267–5280, https://doi.org/10.1175/JCLI-D-16-0030.1.

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. Bollasina, B. Booth, and G. 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.

Visbeck, M. H., J. W. Hurrell, L. Polvani, and H. M. Cullen, 2001: The North Atlantic Oscillation: Past, present, and future.

Proc. Natl. Acad. Sci. USA, **98**, 12 876–12 877, https://doi.org/10.1073/PNAS.231391598.

Wang, C., S. Dong, A. T. Evan, G. R. Foltz, and S. K. Lee, 2012: Multidecadal covariability of North Atlantic sea surface temperature, African dust, Sahel rainfall and Atlantic hurricanes. *J. Climate*, **25**, 5404–5415, https://doi.org/10.1175/JCLI-D-11-00413.1.

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.

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.

Wills, R. C. J., K. C. Armour, D. S. Battisti, and D. L. Hartman, 2019: Ocean–atmosphere dynamical coupling fundamental to the Atlantic multidecadal oscillation. *J. Climate*, **32**, 251–272, https://doi.org/10.1175/JCLI-D-18-0269.1.

Wills, S., D. W. J. Thompson, and L. M. Ciasto, 2016: On the observed relationships between variability in Gulf Stream sea surface temperatures and the atmospheric circulation over the North Atlantic. *J. Climate*, **29**, 3719–3730, https://doi.org/10.1175/JCLI-D-15-0820.1.

Wohland, J., N.-E. Omrani, D. Witthaut, and N. S. Keenlyside, 2019: Inconsistent wind speed trends in current twentieth century reanalyses. *J. Geophys. Res. Atmos.*, **124**, 1931–1940, https://doi.org/10.1029/2018JD030083.

Wunsch, C., 1999: The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillation. *Bull. Amer. Meteor. Soc.*, **80**, 245–255, https://doi.org/10.1175/1520-0477(1999)080<0245:TIOSCR>2.0.CO;2.

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.

Yuan, T. L., L. Oreopoulos, M. Zelinka, H. Yu, J. R. Norris, M. Chin, S. Platnick, and K. Meyer, 2016: Positive low cloud and dust feedbacks amplify tropical North Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, **43**, 1349–1356, https://doi.org/10.1002/2016GL067679.

Zar, J. H., 1999. *Biostatistical Analysis*. 4th ed. Prentice Hall, 944 pp.

Zhang, R., 2008: Coherent surface–subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.*, **35**, L20705, https://doi.org/10.1029/2008GL035463.

——, 2017: On the persistence and coherence of subpolar sea surface temperature and salinity anomalies associated with the Atlantic multidecadal variability. *Geophys. Res. Lett.*, **44**, 7865–7875, https://doi.org/10.1002/2017GL074342.

——, and T. L. Delworth, 2006: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.*, **33**, L17712, https://doi.org/10.1029/2006GL026267.

——, R. Sutton, G. Danabasoglu, T. L. Delworth, W. M. Kim, J. Robson, and S. G. Yeager, 2016: Comment on ''The Atlantic Multidecadal Oscillation without a role for ocean circulation.'' *Science*, **352**, 1527, https://doi.org/10.1126/science.aaf1660.

——, ——, ——, 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.

Zhao, J., and W. Johns, 2014: Wind-forced interannual variability of the Atlantic meridional overturning circulation at 26.5°N. *J. Geophys. Res. Oceans*, **119**, 2403–2419, https://doi.org/10.1002/2013JC009407.