Low-Frequency North Atlantic Climate Variability in the Community Earth System Model Large Ensemble

WHO M. KIM AND STEPHEN YEAGER

National Center for Atmospheric Research, Boulder, Colorado

PING CHANG

Department of Oceanography, Texas A&M University, College Station, Texas, and Ocean University of China, Qingdao Collaborative Innovation Center of Marine Science and Technology, Qingdao, China

GOKHAN DANABASOGLU

National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT

There is observational and modeling evidence that low-frequency variability in the North Atlantic has significant implications for the global climate, particularly for the climate of the Northern Hemisphere. This study explores the representation of low-frequency variability in the Atlantic region in historical large ensemble and preindustrial control simulations performed with the Community Earth System Model (CESM). Compared to available observational estimates, it is found that the simulated variability in Atlantic meridional overturning circulation (AMOC), North Atlantic sea surface temperature (NASST), and Sahel rainfall is underestimated on multidecadal time scales but comparable on interannual to decadal time scales. The weak multidecadal North Atlantic variability appears to be closely related to weaker-than-observed multidecadal variations in the simulated North Atlantic Oscillation (NAO), as the AMOC and consequent NASST variability is impacted, to a great degree, by the NAO. Possible reasons for this weak multidecadal NAO variability are explored with reference to solutions from two atmosphere-only simulations with different lower boundary conditions and vertical resolution. Both simulations consistently reveal weaker-than-observed multidecadal NAO variability despite more realistic boundary conditions and better resolved dynamics than coupled simulations. The authors thus conjecture that the weak multidecadal NAO variability in CESM is likely due to deficiencies in air–sea coupling, resulting from shortcomings in the atmospheric model or coupling details.

1. Introduction

The Atlantic Ocean plays a unique role in the climate system. The meridional heat transport in the Atlantic is cross-equatorial and northward at all latitudes due to the existence of the basinwide, deep-reaching Atlantic meridional overturning circulation (AMOC). The direct estimates of its transport at 26.5°N since 2004 by the RAPID array show substantial variability on subseasonal to interannual time scales (Cunningham et al. 2007; Kanzow et al. 2010; McCarthy et al. 2012). While the observed record is still far too short to reveal low-frequency (i.e., decadal to multidecadal) variability of

AMOC, numerous coupled climate model simulations have shown rich low-frequency AMOC variability, with the dominant time scales varying substantially across models [see, e.g., Danabasoglu (2008) for a review]. The simulated low-frequency AMOC variability is shown to play an important role in modulating surface climate in broad regions of the Northern Hemisphere due to associated changes in northward heat transport[1] (e.g., Knight et al. 2005; Dong and Sutton 2005; Danabasoglu et al. 2012; Delworth and Zeng 2016).

---

[1] More precisely, because of changes in the convergence of the northward heat transport in the North Atlantic (Zhang and Zhang 2015).

Corresponding author: W. M. Kim, whokim@ucar.edu

One prominent feature associated with the low-frequency AMOC variability in coupled models is low-frequency, basinwide fluctuations in sea surface temperature (SST) in the North Atlantic (e.g., Delworth et al. 1993; Knight et al. 2005; Danabasoglu et al. 2012; Ba et al. 2014; Tandon and Kushner 2015). Observations also show basinwide fluctuations in the North Atlantic SST (NASST) on multidecadal time scales, often referred to as the Atlantic multidecadal oscillation or variability (AMV). The AMV has been linked to many low-frequency climate fluctuations in the Northern Hemisphere, such as Sahel and Northern Brazilian rainfall, Atlantic hurricane activity, North American and European temperatures, and Arctic sea ice extent (e.g., Folland et al. 1986; Goldenberg et al. 2001; Enfield et al. 2001; Sutton and Hodson 2005; Zhang and Delworth 2006; Knight et al. 2006; Day et al. 2012; Zhang 2015).

The AMOC changes during the twentieth century are often framed as externally forced, represented by multimodel or ensemble means of climate model simulations. Radiative external forcings cause an overall decline of AMOC strength during the late twentieth century in most coupled models (Cheng et al. 2013) as a precursor of the further substantial decrease in the AMOC in future projections due to anthropogenic warming (Weaver et al. 2012; Cheng et al. 2013). In addition to the long-term decline, external forcings give rise to multidecadal fluctuations in AMOC strength between preindustrial and present-day time periods in some models, with this variability attributed to the combined influence of anthropogenic aerosols and greenhouse gases (Cheng et al. 2013), volcanic forcing (Swingedouw et al. 2015), and changes in solar irradiance (Menary and Scaife 2014). However, the radiatively forced multidecadal AMOC variability during the historical period seems to be model dependent and weak compared to the magnitude of internal AMOC variability (Swingedouw et al. 2015).

The low-frequency AMOC variability associated with variations in the large-scale atmospheric circulation is highlighted in studies utilizing ocean–sea ice simulations forced with historical atmospheric state reanalyses (e.g., Robson et al. 2012; Yeager and Danabasoglu 2014; Danabasoglu et al. 2016). This low-frequency AMOC variability is consistently characterized by an overall increase from the 1970s to the mid-1990s and a decrease thereafter (Eden and Jung 2001; Böning et al. 2006; Biastoch et al. 2008; Robson et al. 2012; Yeager and Danabasoglu 2014; Danabasoglu et al. 2016) and has been attributed to low-frequency variations in the North Atlantic Oscillation (NAO), the leading mode of atmospheric variability over the North Atlantic (Hurrell 1995). The mechanism of the AMOC–NAO relationship involves deep-water formation in the Labrador Sea, driven largely by surface heat loss (Yashayaev 2007; Yashayaev and Loder 2016). Under positive NAO conditions, in general, enhanced westerlies carry cold air from North America to the Labrador Sea (Kim et al. 2016), promoting the formation of Labrador Seawater (LSW) and leading to an increase in the AMOC with some delay.

The link between low-frequency variations in NASST, AMOC, and NAO is also evident in some coupled models (e.g., Dong and Sutton 2005; Danabasoglu et al. 2012). However, it has recently been claimed that similar low-frequency NASST variability can be obtained in coupled simulations that use a slab-ocean model, instead of a fully active, dynamical ocean model (Clement et al. 2015). As such, Clement et al. (2015) argue that the observed AMV can be driven by stochastic atmospheric forcing alone. This claim is disputed by recent studies (Zhang et al. 2016; O'Reilly et al. 2016; Delworth et al. 2017) that show that the simulations with slab-ocean models cannot possibly reproduce the mechanisms maintaining the low-frequency NASST variability identified in fully coupled models. Nevertheless, the results of Clement et al. (2015) raise an intriguing question as to why the NASST power spectra in the coupled simulations are indistinguishable from those of the slab-ocean simulations. Indeed, it has been reported in previous studies that low-frequency NASST variability in many models participating in phase 3 and phase 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5) is less pronounced, and the decorrelation time scale is much faster, than in observations (e.g., Ting et al. 2011; Medhaug and Furevik 2011; Zhang and Wang 2013; Kavvada et al. 2013; Frankcombe et al. 2015; Peings et al. 2016).

A clue can be found in a recent study by Delworth and Zeng (2016), who explore the response of the AMOC and Northern Hemisphere climate to additional NAO-induced surface heat flux imposed in the ocean component of a coupled model. Using experiments in which the imposed periodic NAO-related heat flux anomalies have frequencies ranging from 2 to 100 years, they find that the frequency of the AMOC response matches that of the imposed NAO forcing almost linearly (i.e., a 50-yr NAO forcing frequency produces a 50-yr cycle in AMOC). In addition, using an ocean-only simulation forced with synthetic stochastic NAO forcing, Mecking et al. (2014) find enhanced low-frequency spectral power in AMOC and in the subpolar North Atlantic (SPNA) SST when stochastic forcing exhibits an enhanced power in a similar low-frequency band during the integration. Therefore, these studies suggest that there is an almost linear frequency relationship between the low-frequency variability of the NAO, AMOC, and NASST.

Furthermore, recent studies show that the simulated NAO in CMIP5 models lacks low-frequency variance compared to observations (Kravtsov 2017; Wang et al. 2017). Therefore, there is reason to believe that weak (indistinguishable from red noise) low-frequency NASST variability in the coupled models utilized by Clement et al. (2015) may be related to the weak low-frequency variations in the NAO simulated in these models. However, none of the aforementioned studies has directly related the weak low-frequency NASST (or other relevant climate variables) to the weak low-frequency NAO.

In the present study, we investigate the representation of low-frequency variability in the North Atlantic during the 1920–2009 historical period primarily in a set of Large Ensemble (LE) simulations performed with the Community Earth System Model (CESM; Kay et al. 2015), in comparison with available observational estimates. With its large ensemble size, LE samples a wide range of internal variability and allows for a robust estimation of externally forced signals. As will be shown below, the LE simulations exhibit a clear NAO–AMOC–NASST link, but substantially weaker NASST low-frequency spectral power than observed. We will show that the CESM is characterized by weak internal low-frequency variability in a set of North Atlantic fields closely related to NASST, including AMOC. We argue that a key aspect of the model AMV bias is a deficiency in simulating low-frequency NAO variability, which appears to be more of a cause than a symptom of anemic NASST variability on multidecadal time scales.

The paper is organized as follows. In section 2, we briefly describe the model and observational data used in our study along with analysis methods. Section 3 compares the low-frequency variability in selected variables from LE to observational estimates of low-frequency variability and makes the case that the weaker-than-observed simulated multidecadal variability in the North Atlantic is possibly related to weak variations in the simulated NAO. Possible explanations for the weak simulated NAO variability are also discussed. Section 4 provides a summary and concluding remarks on the implications of our findings for understanding low-frequency climate variability in the North Atlantic.

2. Model simulations, observational data, and analysis methods

a. Coupled simulations

The LE simulations use the CESM version 1 with active biogeochemistry and carbon cycle and with the Community Atmospheric Model version 5 (CESM1-CAM5; Kay et al. 2015). These simulations are forced with historical, observation-based natural, and anthropogenic forcings for the 1920–2005 period and with the representative concentration pathway 8.5 (RCP8.5) forcings for the 2006–2100 period, following the CMIP5 protocol (Taylor et al. 2012). The ensemble size of the LE has been increased from 30 to 40 since the publication of Kay et al. (2015). Here, we utilize the first 35 ensemble members and analyze the historical period 1920–2009 to match the end year of our forced ocean simulation described below.

The LE members are generated by perturbing the initial atmospheric temperature by round-off level changes (Kay et al. 2015) but using the same ocean initial conditions (OICs). Therefore, the ensemble variance associated with uncertainties in ocean initial conditions is possibly undersampled. To partially address this issue, we have run an additional 10-member ensemble with a different ocean initial state. The spread of this new ensemble is generated identically as in the original LE. Hereafter, we refer to this new ensemble as LE-OIC and note that it is used to supplement the analysis of LE. Further details of the LE-OIC along with a discussion of how its solutions differ from those of the original LE simulations are given in appendix A.

We also use the last 1400 years of the 2200-yr CESM LE preindustrial control simulation (CTRL)—avoiding the initial transients and drifts—to estimate and compare purely internal low-frequency variability to low-frequency variability obtained in LE. All time series from CTRL are linearly detrended prior to analyses.

b. Atmosphere-only simulations

In addition to the coupled model simulations, we make use of two 10-member ensembles of historical atmosphere-only simulations to investigate the impacts of boundary conditions on low-frequency NAO characteristics. The first ensemble, referred to as LT (low top), uses the same CAM5 and external forcings as in the LE and LE-OIC but employs as its surface boundary conditions the monthly Extended Reconstructed SST (ERSST) version 4 (Smith et al. 2008) between 28°S and 28°N and the monthly ERSST climatology poleward of 35° with prescribed climatological sea ice conditions. As the SSTs are fixed to climatology poleward of 35°, the LT ensemble allows us to isolate the influence of tropical–subtropical SST on low-frequency NAO variability. This ensemble is available for the 1880–2014 period.

The second atmosphere-only ensemble, referred to as MT (middle top), also employs CAM5, but with 46 vertical layers and an extended model top at 0.3 hPa (Richter et al. 2015), in contrast to the 30 levels and model top at ~2 hPa that is used in LE, LE-OIC, and LT. The lower boundary conditions of MT are the

monthly SST and sea ice conditions from Hurrell et al. (2008) and, in contrast to LT, vary interannually everywhere. In comparison with LT, MT allows us to examine the influence of extratropical SST and better resolved stratospheric dynamics on low-frequency NAO variability. This ensemble spans the 1953–2015 period, with RCP4.5 forcings used for the 2006–15 period.

c. Forced ocean–sea ice simulation

In the absence of any long, continuous observations of AMOC to quantify variability on decadal to multidecadal time scales, an alternative is to obtain an estimate of historical AMOC variability from a forced ocean–sea ice simulation (FO). In FO, the ocean and sea ice models are identical to those of CESM1-CAM5: the Parallel Ocean Program version 2 (POP2; Smith et al. 2010) and Community Ice Code version 4 (CICE4; Hunke and Lipscomb 2008). FO is forced with 6-hourly atmospheric state variables, daily radiative fluxes, and monthly precipitation from the Coordinated Ocean-Ice Reference Experiments (CORE) interannually varying atmospheric datasets (Large and Yeager 2009) and run for five forcing cycles (1948–2007) to obtain a cyclically quasi-steady state. Then, the last cycle is extended to 2009 and the 1958–2009 period is used for our analysis. For further details of the CORE datasets and forcing protocol, readers are referred to Large and Yeager (2009), Griffies et al. (2012), and Danabasoglu et al. (2014).

A key consideration supporting the use of the FO simulation is that it is able to reproduce important aspects of variability in the North Atlantic, compared to available observations. In particular, the simulated low-frequency variability in Labrador Sea hydrographic properties (Yeager and Danabasoglu 2014; Danabasoglu et al. 2016) and mixed layer depth (Kim et al. 2016) show good agreement with in situ observations. Such good agreement between the simulated and observed data in the particular fields that are known to impact AMOC variability on decadal and longer time scales (e.g., Yeager and Danabasoglu 2014) gives us confidence that the low-frequency thermohaline variability simulated in FO represents a reliable estimate of that of the real ocean.

d. Observational data

We use several observational datasets to evaluate the simulated variability in the North Atlantic. For the NASST and SPNA SST, we employ the observed SST data from the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST; Rayner et al. 2003) version 1.1. For Sahel rainfall, the precipitation data from Climate Research Unit, University of East Anglia

(CRU), version 3.23 (Harris et al. 2014) is used. For the evaluation of the simulated NAO variability, the winter [December–March (DJFM)] NAO index based on the sea level pressure difference between Lisbon, Portugal, and Reykjavik, Iceland (Hurrell et al. 2017), is employed. For the spatial pattern of the NAO as well as other leading atmospheric modes, the NOAA Twentieth Century Reanalysis version 2c (20CRv2c; Compo et al. 2011) is used. The temporal coverage of the above datasets varies, but all of them cover the entire twentieth century. Although we show the full length of the data in the time series plots, we only use the 1920–2009 period for our analyses in order to match the time period of LE. In addition, we utilize a recent reconstruction of winter NAO computed from multiple proxy records that is available at annual resolution from 1049 to 1969 (Ortega et al. 2015).

e. Methods

We estimate the low-frequency North Atlantic variability using two methods. First, we conduct a standard spectral analysis to obtain the spectral power of selected variables. However, the historical records are short relative to the low-frequency periods of up to 60 years that we are interested in (i.e., the spectral peak of the observed AMV) (e.g., Peings et al. 2016). Thus, to supplement the spectral analysis, we compute moving trends with window lengths of 5, 15, and 30 years and examine the distributions of the trends. We chose these trend lengths somewhat arbitrarily to quantify the amplitude of the variability on decadal to multidecadal time scales. The results from the moving trend analysis are not very sensitive to the precise choice of window lengths.

We also perform standard regression and correlation analyses when examining the relationships between two variables, and empirical orthogonal function (EOF) analysis to capture the dominant modes of variability. Because we are interested in low-frequency variability, we first smooth all time series using a 15-yr Butterworth low-pass filter before computing regressions or correlations. EOFs are computed using unsmoothed annual time series and the principal component (PC) time series are normalized to have unit variance. Thus, the magnitudes of the EOF spatial patterns correspond to one standard deviation change in the PC time series. For all ensemble simulations, we apply all statistical analyses to each ensemble member separately and then compute the ensemble mean, if possible, with ensemble spread. When necessary, we test the statistical significance of the regressions and correlations at the 95% confidence level using a two-sided Student's t test with the effective degrees of freedom
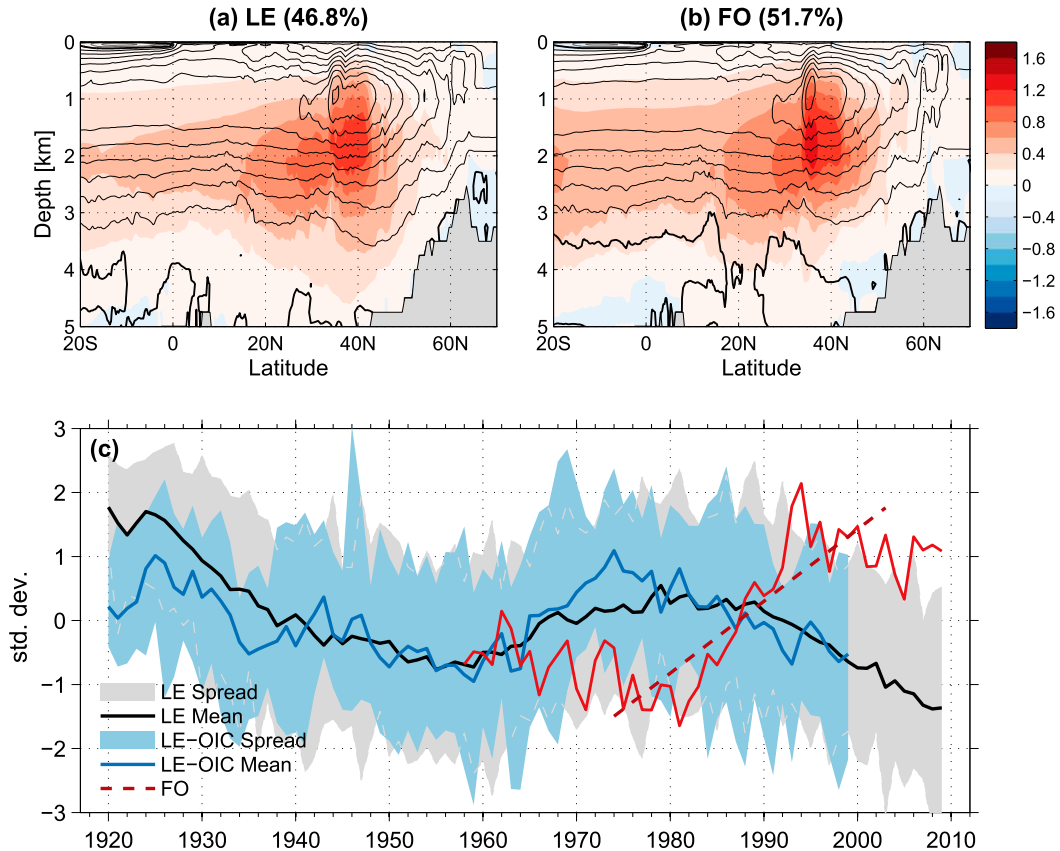
FIG. 1. Time-mean (contours) and EOF1 of AMOC (colors) from (a) LE and (b) FO, with the fraction of the total variance accounted for by EOF1 given in parentheses, and (c) PC1 time series from LE (black), LE-OIC (blue), and FO (red). In (a), the time mean, EOF1, and explained variance of AMOC are the ensemble mean of LE. In (c), the full ranges of ensemble spread of PC1 from LE and LE-OIC are indicated by light gray and blue shading, respectively, and the black and blue lines indicate the respective ensemble means. In (a) and (b), the unit is Sv (1 Sv $\equiv$ 10^6 m^3 s^{-1}), and the light black contour lines indicate positive (clockwise) circulation with a contour interval of 3 Sv; the zero contour level is denoted with a thick black line. The EOF analysis is for the 1920–2009 period for LE and for the 1958–2009 period for FO. The maroon dashed line in (c) indicates the period with the maximum 30-yr trend in FO.

computed based on the method of Bretherton et al. (1999).

3. Results

a. Low-frequency AMOC variability

We first present the time-mean and the first EOF (EOF1) of AMOC along with the associated PC time series (PC1) from LE and FO, as well as PC1 from LE-OIC, in Fig. 1. The AMOC time-mean and EOF1 are very similar between LE and FO. In both, the time-mean AMOC has a maximum overturning strength of about 25 Sv (1 Sv $\equiv$ 10^6 m^3 s^{-1}) centered at around 35°N, but the upper cell penetrates slightly deeper in LE than in FO. The EOF1, accounting for about 47% and 52% of the respective total AMOC variance in LE (ensemble

mean) and FO, reflects a basin-scale fluctuation of AMOC of order 1 Sv per unit standard deviation of PC1, in both LE and FO.

The ensemble-mean PC1 from LE exhibits a multi-decadal variability (Fig. 1c), showing three stages: a weakening from 1920 until about 1960, a subsequent increase until the late 1970s, and then a decline until the end of our analysis period. We note that the magnitude of the initial weakening trend is considerably smaller in LE-OIC than in LE, suggesting an influence of ocean initial conditions as discussed in appendix A. The peak-to-peak change of the forced AMOC variability in midlatitudes is about 3 Sv. This forced multidecadal variability is similar to that found in some CMIP5 models, the cause of which has been attributed to a combined influence of anthropogenic aerosols and greenhouse gases (Cheng et al. 2013). The forced signal
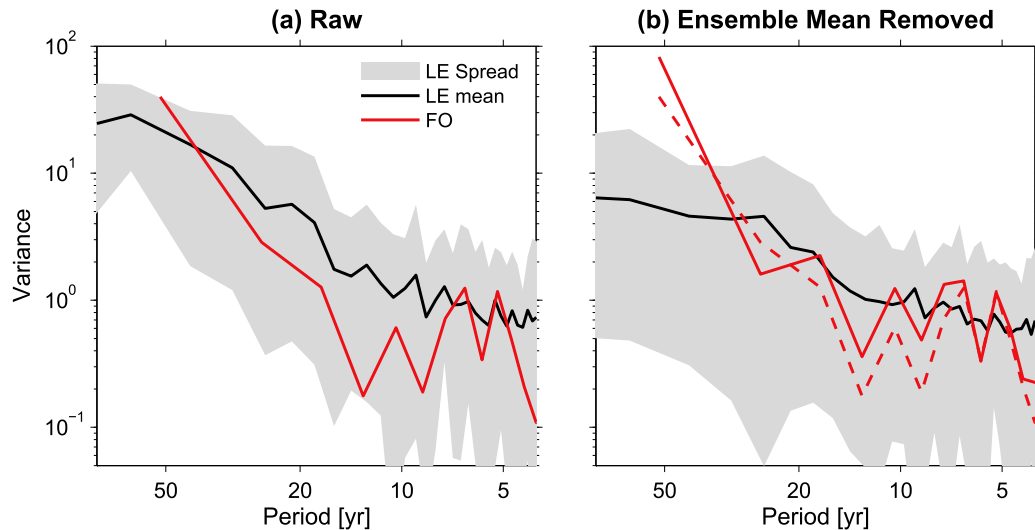
FIG. 2. Power spectra of AMOC PC1 time series from LE (ensemble mean; black) and FO (red) with the ensemble spread (5th–99th percentile) of LE (gray shading): (a) from the raw AMOC PC1 and (b) after removing the ensemble mean PC1 of LE. The dashed red line in (b) is the raw PC1 spectrum of FO shown in (a).

seen in LE and LE-OIC during the second half of the twentieth century stands in stark contrast with the variability diagnosed from FO. In particular, the FO PC1 shows a large upward trend during the 1980s and early 1990s, with a slight decrease only thereafter. This pronounced upward trend is also found in other AMOC indices based on the maximum overturning circulation at fixed latitudes (not shown). It is also a common feature in many other ocean hindcast simulations forced with atmospheric state reanalyses (e.g., Biastoch et al. 2008; Danabasoglu et al. 2016) and in some ocean reanalysis products (e.g., Pohlmann et al. 2013).

At first glance, the AMOC variability in FO appears consistent with the LE and LE-OIC ensemble spread, with only slight excursions outside of the shaded region in Fig. 1c. The power spectrum of the FO PC1 is also within the ensemble envelope of the LE PC1 power spectra in all frequency bands (Fig. 2a). However, this interpretation can be misleading as the multidecadal variability in the individual members of LE appears to be dominated by the externally forced signal. Figure 3 shows the 30-yr moving trends of the AMOC PC1 from all ensemble members of LE along with those from the ensemble mean PC1 (bottom row) and FO (top row). The 30-yr trends from the individual members of LE are largely consistent with those of the ensemble mean: negative trends in the earlier and later periods and positive trends in between (roughly from the mid-1940s to the mid-1960s). The dominant influence of external forcings on multidecadal AMOC variability in LE is also suggested in Fig. 2a, as the envelope of the LE spread becomes narrow toward multidecadal frequency bands.

In sharp contrast, the strong positive trend in FO takes place when the majority of the LE members show negative trends in the later period. Low-frequency variability in NAO-induced buoyancy forcing, which appears to be largely internal (Gillett and Fyfe 2013), is commonly invoked as the mechanistic explanation for the upward trend in AMOC during the late twentieth century (e.g., Latif et al. 2006; Biastoch et al. 2008; Yeager and Danabasoglu 2014). The dominant influence of external forcings on multidecadal AMOC variability in LE during this period therefore casts doubt on the fidelity of the representation of internal variability in LE.

Given the strong signature of forced multidecadal AMOC variability in LE, we recompute the power spectra of the AMOC PC1 from LE after subtracting the ensemble mean PC1 from both LE and FO (Fig. 2b). As expected, the variance in the low-frequency bands (time scales longer than ~20 yr) is substantially reduced in LE, while the variance of relatively high-frequency variability remains largely unaffected. As a result, the multidecadal variance of FO is now outside the LE range. Because the forced signal of LE is almost out of phase with that of the FO PC1 (Fig. 1c), the removal of the ensemble mean of LE amplifies the multidecadal variability in FO. In other words, to the extent that the LE ensemble mean would accurately reflect the forced AMOC variability of the real world, the internal multidecadal variability in FO would be even stronger (and even further outside the LE range).

Figure 4 presents an extended moving trend analysis, considering 5-, 15-, and 30-yr segments for AMOC PC1
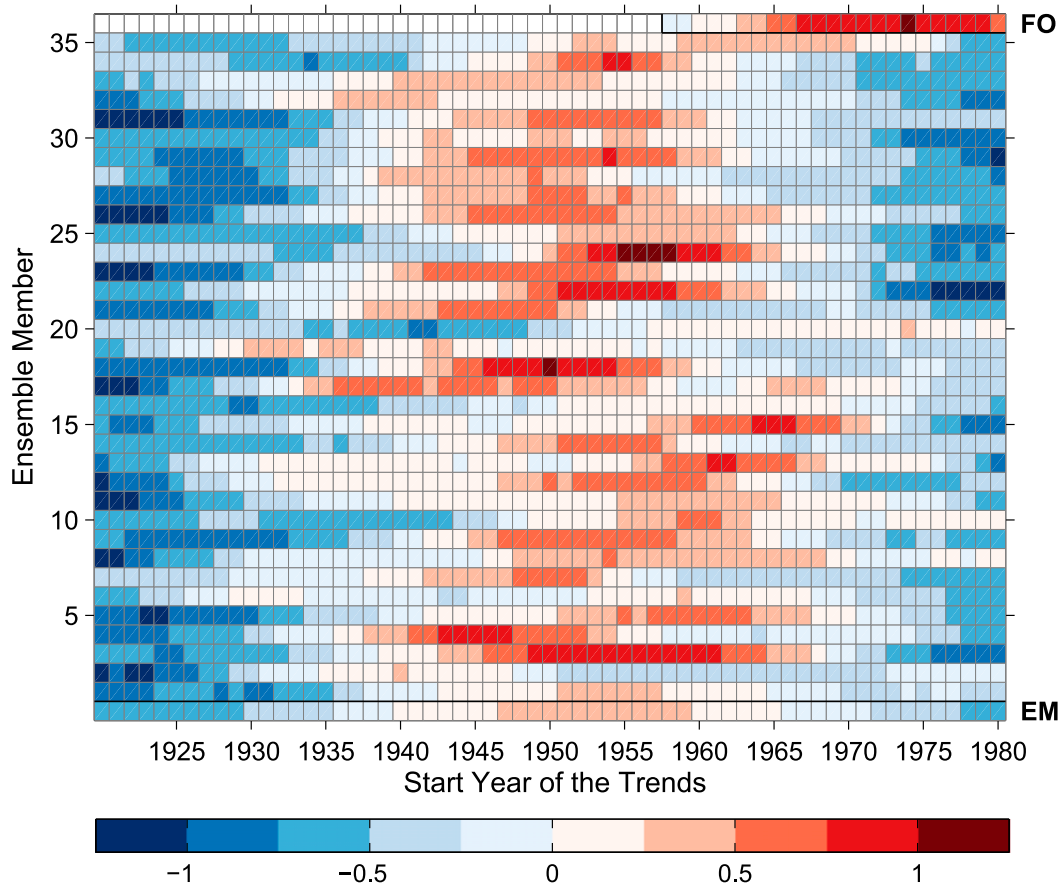
FIG. 3. 30-yr moving trends in the AMOC PC1 time series for each individual member of LE. 30-yr moving trends for the ensemble-mean AMOC PC1 of LE (EM; bottom row) and for the AMOC PC1 of the forced ocean–sea ice simulation (FO; top row) are also shown. All trends are normalized by the maximum 30-yr trend in FO [3.4 std dev (30 yr) $^{-1}$]. The trend values are plotted at the start year of the trend range.

from LE and CTRL without (top) and after (bottom) subtracting the LE ensemble mean. All trends are normalized to the corresponding maximum (absolute) trend in the raw FO PC1 (e.g., the 1974–2003 trend shown in Fig. 1c for the 30-yr trends). In general, while the upper limit (99th percentile) of the LE distribution is close to the maximum trend of FO for the 5-yr trends, regardless of whether the ensemble mean is removed or not, it tends to become lower as the trend length increases. Without removing the forced signal, there are still a few instances where the 30-yr moving trends from LE exceed the maximum trend of FO, and the distribution of LE 30-yr trends is wider than that of CTRL. However, when the ensemble mean is removed, none of the 30-yr trends exceeds the maximum trend of FO, and the distribution of LE becomes comparable to CTRL. This indicates that the forced signal significantly enhances the multidecadal AMOC variability in LE (the variance of 30-yr moving trends is roughly 40% greater

with the ensemble mean left in). In contrast, if the ensemble mean of LE is removed from FO, the maximum 30-yr trend increases by about 40%, consistent with the results from spectral analysis. Therefore, these results demonstrate that the presumably realistic multidecadal AMOC trend in FO falls either outside or in the extreme tail of the LE distribution, suggesting that the internal multidecadal variability of the AMOC in LE is too weak.

b. Climate impacts of the simulated multidecadal AMOC variability

Figures 5a and 5b show the SST spatial pattern associated with the AMV for the ensemble mean of LE and that of the HadISST. These patterns are obtained as the simultaneous regressions of the annual-mean SST time series onto the respective AMV indices, defined as the SST averaged over 0°–60°N, 75°–7°W and 15-yr low-pass filtered. To isolate internal variability, the ensemble mean
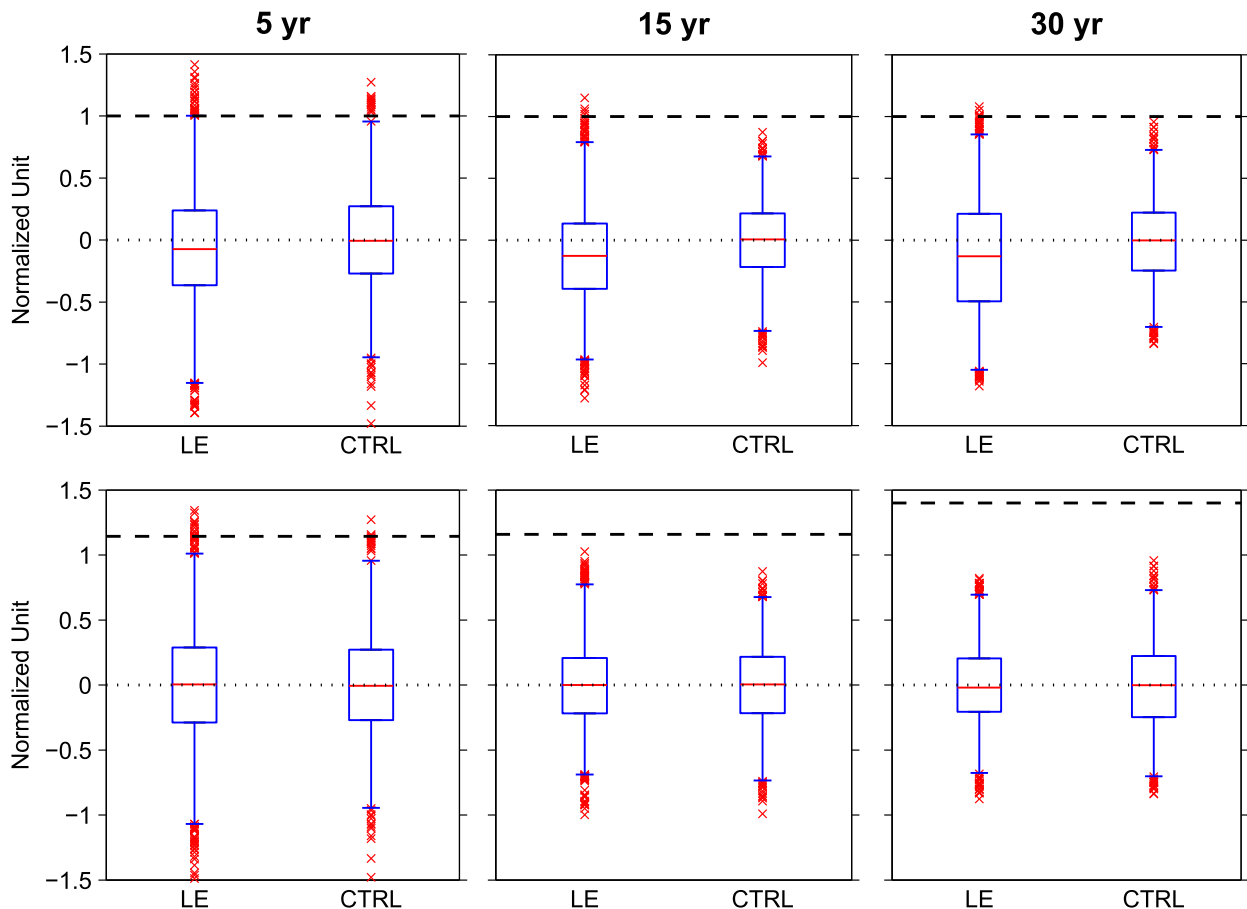
FIG. 4. Box plots of 5-, 15-, and 30-yr moving trends from LE and CTRL (top) for the raw AMOC PC1 time series and (bottom) after removing the ensemble mean AMOC PC1 time series from each LE member (CTRL is the same in top and bottom panels). The moving trends from the individual ensemble members of LE are binned into a single box plot. The sample sizes are 3010 and 1397 (5 yr), 2660 and 1387 (15 yr), 2135 and 1372 (30 yr) for LE and CTRL, respectively. The box plots show the median (red center line), interquartile range (box height), 1st–99th percentile range (whiskers), and outliers (red crosses). All trends are normalized to the corresponding maximum trend of the raw AMOC PC1 time series from FO with the normalized maximum FO trend indicated by the dashed black lines in each panel.

AMV index is subtracted from each individual AMV index in LE, and a linear trend is subtracted from the observed AMV index prior to the regressions.[2] The AMV patterns from observations and LE show broad agreement, although LE has a wider cold anomaly in the western midlatitude North Atlantic, with a canonical AMV shape (e.g., Ting et al. 2011): a broad warming in the entire North Atlantic with the maximum anomaly in the SPNA and a limb extending along the eastern part of the basin into the subtropics and tropics. However, the

magnitude of the AMV anomalies in LE is actually much weaker than in observations because the standard deviation of the AMV index in LE—ranging from 0.04° to 0.09°C with an ensemble mean of 0.06°C—is substantially weaker than the observed value of 0.14°C (note that the AMV regressions are per unit °C in both Figs. 5a and 5b).

Recent modeling studies suggest that the SPNA component of the AMV results from the AMOC-driven meridional heat transport convergence in the subpolar region while the subtropical to tropical extension is largely due to an atmospheric response to the SPNA SST (Zhang and Zhang 2015; Brown et al. 2016b). The lead–lag correlations between the AMOC PC1 and SPNA SST (averaged over the boxed region indicated in Fig. 5a) time series in LE and CTRL support such a relationship between AMOC and SPNA SST (Fig. 5c). The correlations are maximized and statistically significant, as indicated in

---

[2] Removing the forced component from observations remains an area of ongoing research, and removing a linear trend is likely suboptimal. However, we note that a similar regression pattern for observations is obtained with a somewhat weaker amplitude in the SPNA when the ensemble mean AMV index of LE is removed from the observed AMV index.
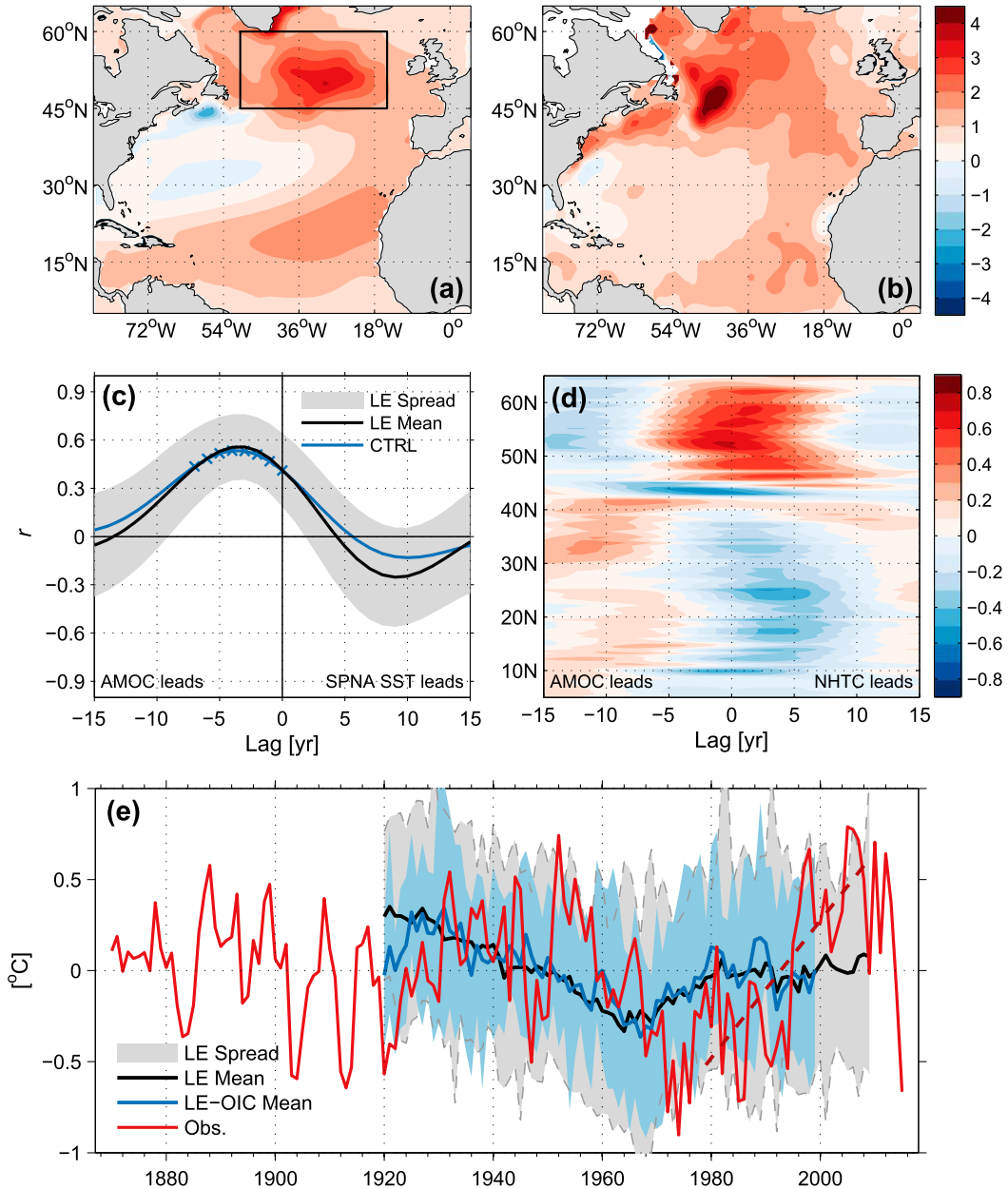
FIG. 5. SST regressions onto respective AMV indices from (a) LE (ensemble mean) and (b) HadISST for the 1920–2009 period. (c) Lead–lag correlations between the AMOC PC1 and SPNA SST time series from LE (gray shading for one standard deviation of ensemble spread and black line for the ensemble mean) and CTRL (blue). (d) Ensemble mean lead–lag correlations between the AMOC PC1 and northward heat transport convergence (NHTC) time series as a function of latitude from LE. (e) Annual SPNA SST anomaly time series (relative to the 1920–2009 mean) from LE (ensemble mean; black), LE-OIC (ensemble mean; blue), and HadISST (red). The boxed area in (a) shows the region used to calculate SPNA SST. In (e), the shading indicates the ensemble spread of LE and LE-OIC, as in Fig. 1. In (a), (c), and (d), the ensemble mean is removed for LE prior to the regression and correlation analyses, while the linear trend is removed from the observations in (b). In (c), statistically significant correlations in CTRL at the 95% confidence level are indicated by crosses. The maroon dashed line in (e) indicates the period with the maximum 30-yr trend in the observed SPNA SST. The unit is °C °C^{-1} in both (a) and (b).
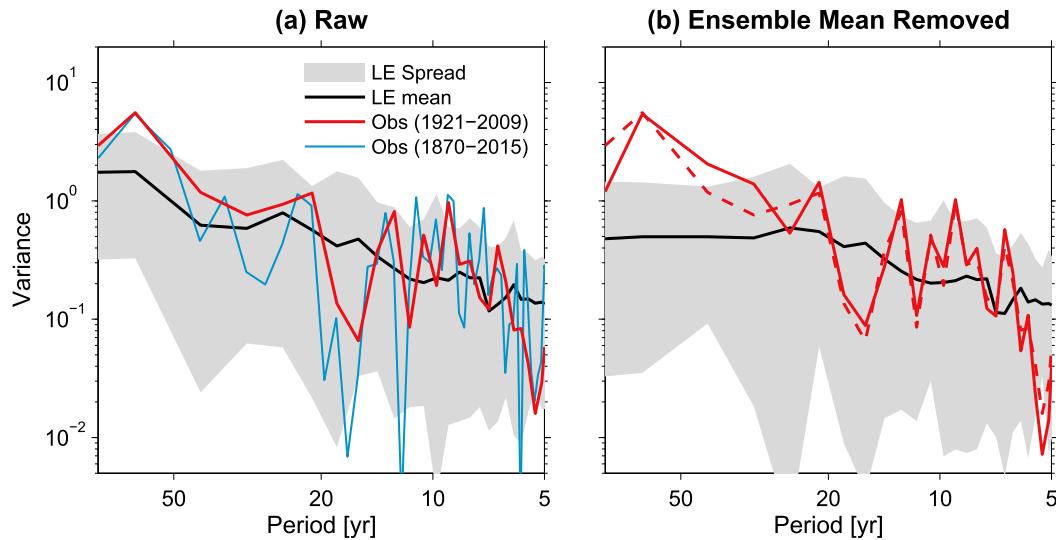
FIG. 6. As in Fig. 2, but for the SPNA SST. The thin blue line in (a) is the spectrum from observations (HadISST) for the longer (1870–2015) period.

CTRL, when the AMOC PC1 leads the SPNA SST by about 3 years in both LE and CTRL. Consistent with the results of Brown et al. (2016b), the northward heat transport associated with the low-frequency AMOC variability shows convergence in the SPNA, as demonstrated by the lead–lag correlations between the time series of the AMOC PC1 and the meridional convergence of the northward heat transport from LE (Fig. 5d).

Because of its direct connection with AMOC in the model, SPNA SST may serve as a proxy for AMOC-related variability both in models and observations. The SPNA SST from observations shows pronounced multidecadal fluctuations (Fig. 5e), similar to that of the traditional AMV. During the 1920–2009 period that overlaps with LE, the peak-to-peak amplitude of the multidecadal variability is >1.5°C, although the multidecadal variability is relatively weak and interannual to decadal variability is relatively strong during the earlier period (1870–1920). The ensemble mean SPNA SST of LE shows a forced multidecadal signal with a peak-to-peak amplitude of about 0.5°C. The very early warm state of the ensemble mean in LE is partly related to the ocean initial conditions, as LE-OIC shows a somewhat colder state, but the two ensemble means quickly converge within a few years. The ensemble ranges of LE and LE-OIC appear to encompass the observed SPNA SST range but, as was the case for AMOC, this interpretation is misleading when multidecadal variability is considered.

Figure 6a shows the power spectra of the annual-mean SPNA SST time series from LE and observations. The LE ensemble encompasses the power spectrum of the observed SPNA SST in almost all frequency bands, but the observed variance around the 60-yr period is slightly

above the upper limit of the LE distribution. This ~60-yr peak is also evident when the full record of the observed SPNA SST is used (i.e., 1870–2015). As in the case of AMOC, the multidecadal SPNA SST variability in LE appears to be significantly enhanced by external forcings. The subtraction of the ensemble mean from LE yields much weaker variance in multidecadal frequency bands (Fig. 6b), increasing the difference between LE and observations. Removing the ensemble mean of LE from observations does not affect the observed spectrum much, because the amplitude of the former is much weaker than that of the latter and they are largely in quadrature (Fig. 5e).

Figure 7 shows the 30-yr moving trends of the SPNA SST from the individual ensemble members of LE, along with those from the ensemble mean (bottom row) and observations (top row). As was the case for AMOC, they tend to cluster around the same sign of the trends in the ensemble mean, supporting the significant influence of the forced signal on multidecadal SPNA SST variability in LE found in the spectral analysis above. Although there are some ensemble members where the timing of the 30-yr moving trends shows some similarity to the observations (e.g., ensemble members 6, 7, 20, 32, and 35), the amplitudes are much weaker than observed.

Box plots of the moving trends further support the weaker internal multidecadal variability of SPNA SST in LE relative to that of observations (Fig. 8). The top panels of Fig. 8 show the 5-, 15-, and 30-yr moving trends of SPNA SST from LE, CTRL, and observations. Here, the ensemble mean has been removed from both LE and observations. While the box plots show relatively
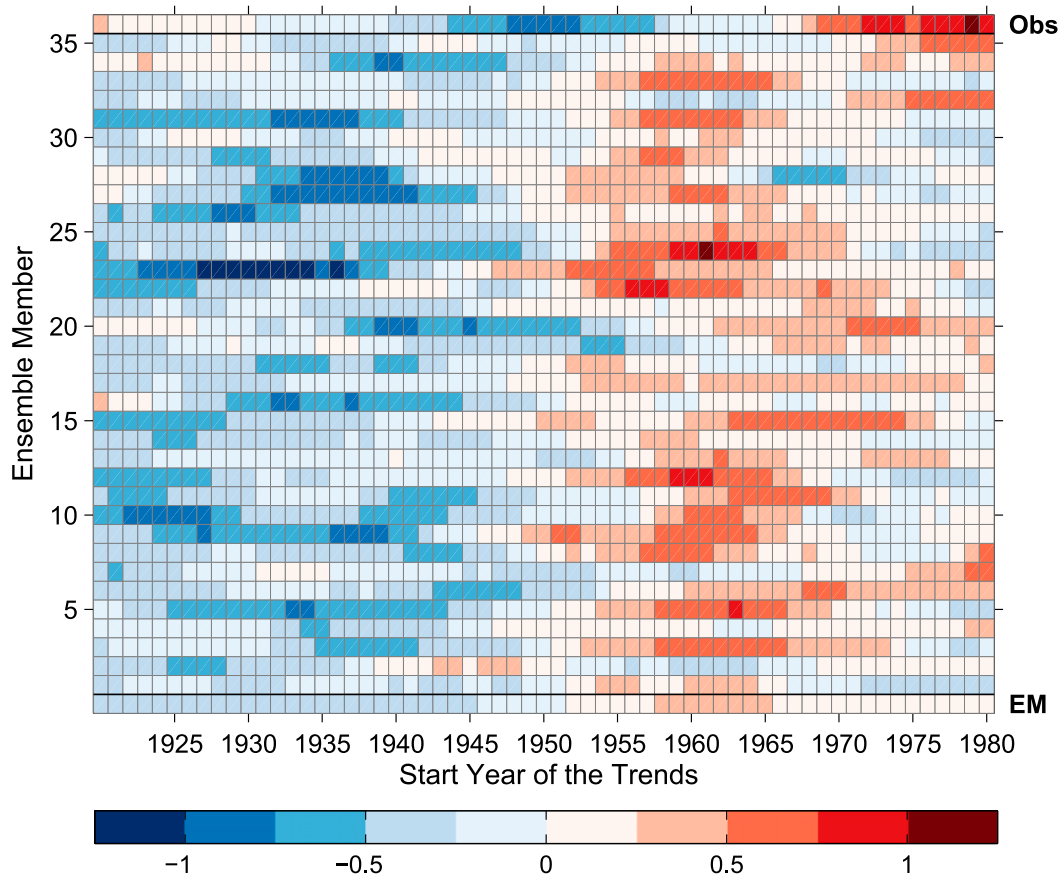
FIG. 7. As in Fig. 3, but for the SPNA SST. The top row shows the 30-yr trends from observations (HadISST). All trends are normalized by the maximum 30-yr trend in observations [$1.1°C \, (30 \, yr)^{-1}$].

similar distributions of LE and CTRL for the 5-yr trends, relative to observations, they become narrower when the trend length increases (top row). However, given the short length of the observations relative to the 30-yr moving trend window, comparing the distribution from a single observed realization to that from 35 simulated realizations may give biased results. Therefore, we also compare the distribution of the observed 30-yr moving trends to those from all individual ensemble members of LE (Fig. 8, bottom row). Although the distributions of the 30-year trends from LE vary quite substantially across the ensemble members, all distributions from LE are narrower than observed, and none of the members simulates a positive trend as large as that observed between 1979 and 2008.

Both observations (e.g., Ting et al. 2011) and model simulations (e.g., Zhang and Delworth 2006) suggest that the low-frequency variability of rainfall in the Sahel region, particularly during the summer months [June–September (JJAS)], is associated with the AMV. In Fig. 9, we show the regression distributions of JJAS precipitation over land onto the annual AMV index

(15-yr low-pass filtered) from LE and observations. In both, an increase in rainfall in the Sahel region associated with a positive AMV phase is clearly seen, although it is constrained to a relatively small region near the west coast of Africa between 10° and 20°N in LE in contrast to a continent-wide strip in observations. This increase in rainfall is due to the northward shift of the intertropical convergence zone, associated with the AMV (Folland et al. 1986).

To compare the rainfall variability between observations and LE, we obtain the area-averaged JJAS precipitation time series for a region in the western part of the Sahel (shown in Fig. 9a). The JJAS Sahel precipitation shows, overall, a gradual decrease in both LE ensemble mean and observations (Fig. 9c), consistent with previous studies, suggesting a decreasing trend in Sahel precipitation in response to global warming (e.g., Held et al. 2005). Superimposed on this slow decrease, the observed Sahel rainfall time series also exhibits large-amplitude multidecadal variability. In particular, a downward trend during the 1950s to 1970s and an upward trend thereafter are consistent with the multidecadal
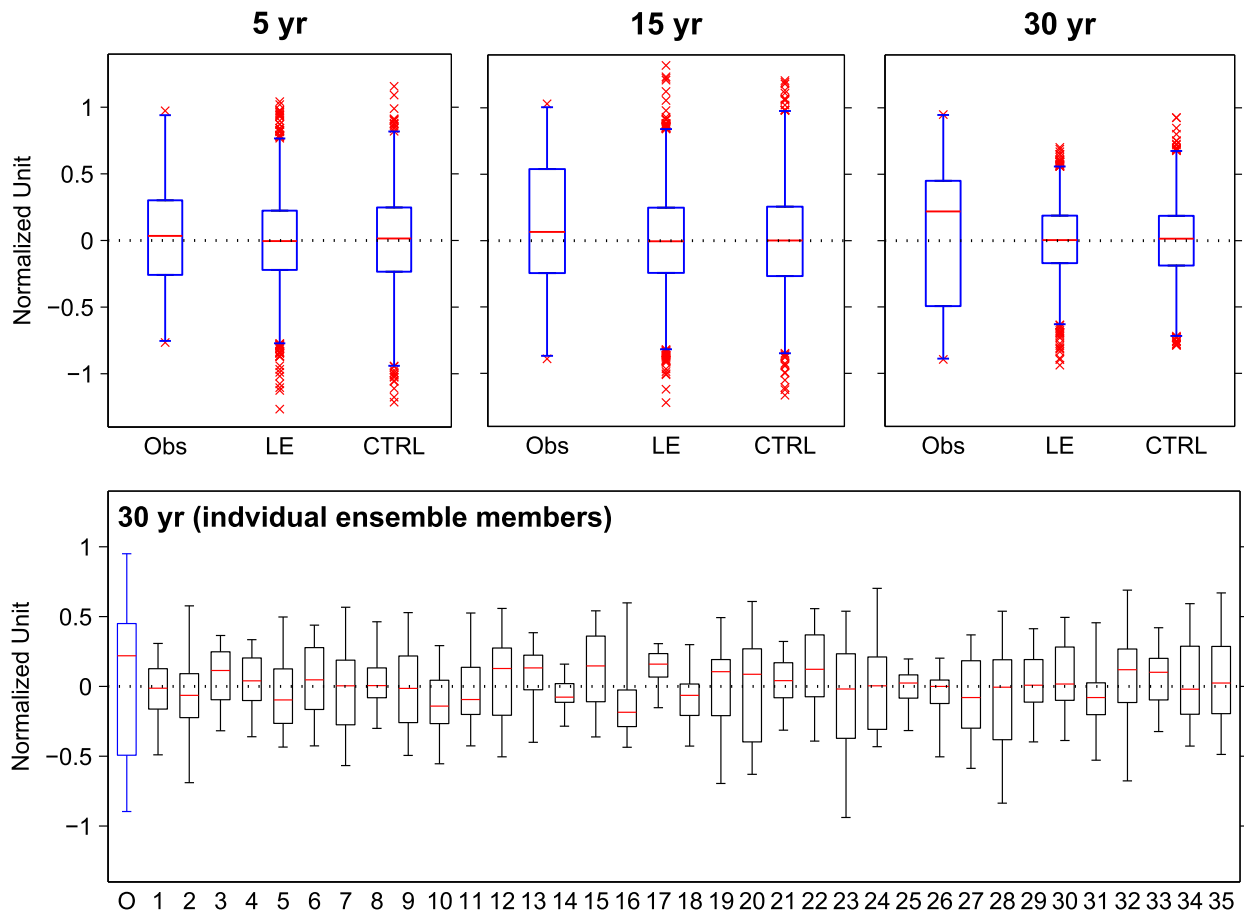
FIG. 8. (top) Box plots of 5-, 15-, and 30-yr moving trends of SPNA SST from observations (HadISST), LE, and CTRL after removing the ensemble mean SPNA SST from both observations and LE. (bottom) The box plots of the same 30-yr moving trends, but for individual ensemble members. All moving trends are normalized to the corresponding maximum trend from the raw observed SPNA SST. The sample sizes for the observations in the top panels are 86, 76, and 61 for 5-, 15-, and 30-yr trends, respectively; otherwise as in Fig. 4. Whiskers in the bottom panel indicate the maximum range in contrast to the 1st–99th percentile range in the top panels. The box plot labeled as O in the bottom panel is from the observations and same as in the top-right panel.

variability of the observed SPNA SST (Fig. 5e) as well as NASST (not shown), although the strongest trends occur during different periods.

The observed Sahel rainfall variability is again within the variability envelope of LE. However, consistent with the other variables discussed above, the spectral analysis shows that the simulated Sahel rainfall lacks the variance in multidecadal frequency bands, whether the ensemble mean is removed from LE or not and whether the ensemble mean of LE is removed from observations or not (Figs. 10a,b). The distributions of the 30-yr trends from the individual ensemble members of LE are much narrower than those from observations, further confirming that multidecadal Sahel rainfall variability in LE is weaker than in observations (Fig. 10c). Hence, these results further demonstrate the overall weakness of multidecadal North Atlantic

climate variability in LE, which is likely associated with the weak multidecadal AMOC variability.

c. Low-frequency NAO

Because the NAO is one of the primary drivers of variability in the North Atlantic, including that of AMOC (e.g., Visbeck et al. 2003; Delworth and Zeng 2016), we next investigate the variability characteristics of the simulated NAO in LE in comparison with that of observations. We start with the spatial structures of the NAO from LE and observations, obtained as the EOF1 of the wintertime (DJFM) sea level pressure (SLP) over the North Atlantic (Figs. 11a,b). In general, the ensemble mean NAO pattern and its explained variance in LE are in good agreement with observations, but with a noticeably weaker southern lobe in LE. We note that the pattern and explained
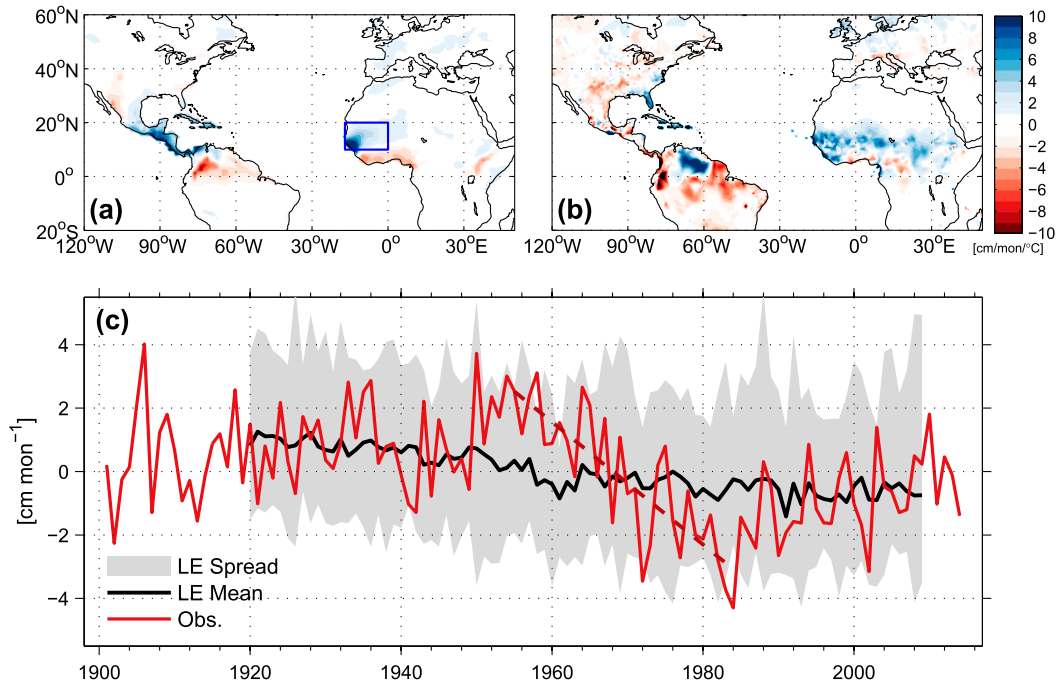
FIG. 9. JJAS land precipitation regression onto the AMV index from (a) LE (ensemble mean) and (b) observations (CRU) for the same 1920–2009 period. (c) JJAS rainfall anomaly (relative to the 1920–2009 mean) in the Sahel region, averaged over the boxed region indicated in (a), from the ensemble mean of LE (black), with the maximum ensemble range (gray shading), and CRU (light red). The maximum (negative) 30-yr trend in observations is indicated by the dashed maroon line.

variance vary considerably across ensemble members (Deser et al. 2017).

The NAO index time series are presented in Fig. 11c. For consistency with the station-based NAO index (Hurrell et al. 2017), we use the difference of normalized SLP between Lisbon, Portugal, and Reykjavik, Iceland (denoted as boxed regions in Fig. 11a), to obtain NAO indices from LE. The LE ensemble-mean NAO index shows a very weak upward trend, consistent with the results of a study analyzing CMIP5 simulations suggesting that there is a weak upward trend in the NAO in response to increasing greenhouse gases (e.g., Gillett and Fyfe 2013). In addition to strong interannual variability, the observed NAO index shows discernible low-frequency variations, and the upward trend during the early 1960s to the mid-1990s is particularly pronounced. As discussed earlier, this trend has been linked to the positive AMOC trend simulated in forced ocean simulations over a comparable time span (Fig. 1c).

Figure 12 shows the lead–lag correlations between the low-pass filtered NAO index and AMOC PC1 from LE and CTRL. The correlation functions are largely consistent between LE ensemble mean and CTRL with the maximum correlations (r ~0.4) occurring when the NAO leads the AMOC by about 4–5 years. As indicated for CTRL, the correlations are statistically significant

around these lags. Also, the maximum correlation becomes greater ($r > 0.6$) if an AMOC index in the subpolar region (e.g., 45°N) is used (not shown), suggesting that the NAO plays an important role in driving low-frequency AMOC variability, as in FO.

The spectral analysis of the NAO indices shows the range of the power spectra from LE mostly encompassing the power spectrum of the observed NAO on interannual to decadal time scales (Fig. 13a). However, while the ensemble mean LE NAO spectrum shows the characteristics of white noise, the observed spectral power on multidecadal time scales exceeds the upper limit of the LE range whether or not the ensemble mean is removed from LE (Fig. 13b). Also, the pronounced multidecadal power in the observed NAO appears to be robust regardless of whether the overlapping period with LE or the entire observational record (1864–2015) is used.

The box plots of the moving trends in the NAO are shown in Fig. 14. Here, all moving trends are computed without subtracting the ensemble mean of LE. Because of the weak forced signal (Fig. 11c), the results after removing the ensemble mean are almost identical. For the 5- and 15-yr trends, the binned distributions of the NAO from LE and CTRL are slightly narrower but comparable to those from observations (top row). However, for the 30-yr trends, the distributions from both LE and CTRL

(a) Raw

(b) Ensemble Mean Removed
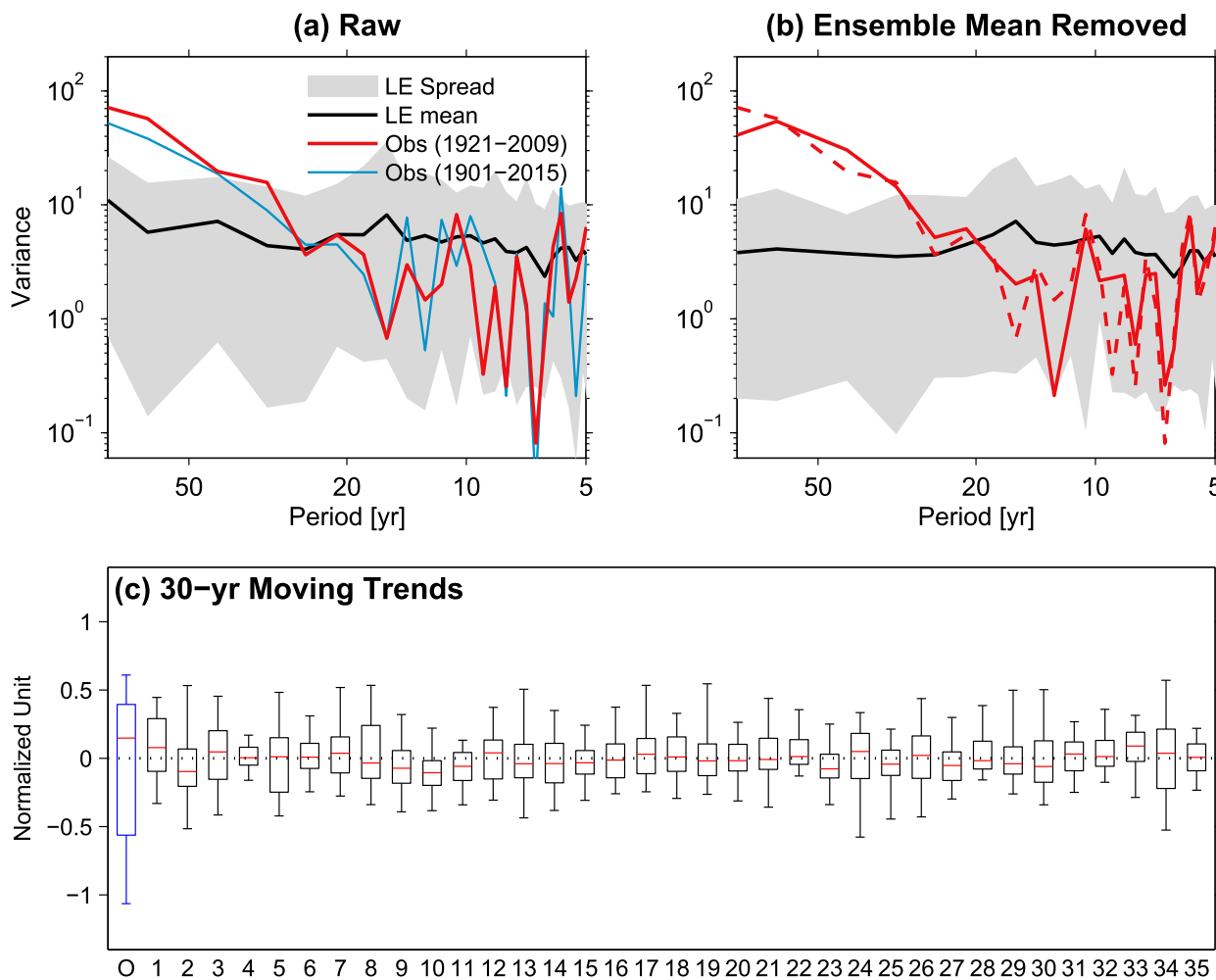

(c) 30-yr Moving Trends

FIG. 10. (a),(b) As in Fig. 6 and (c) as in the bottom panel of Fig. 8, but for the JJAS Sahel rainfall. The 30-yr moving trends from LE in (c) are normalized to the maximum (absolute) trend from the raw observed Sahel rainfall (CRU) denoted in Fig. 9c. The sample size for each box plot in the bottom panel is 61.

are substantially narrower than observed. The distributions of the 30-yr trends from the individual ensemble members of LE vary quite substantially across the ensemble members (bottom panel), as for SPNA SST, but are all much narrower than the observed distribution. Thus, these results suggest that the multidecadal variability of the simulated NAO in LE is weaker than that of the relatively short observational record, and are in line with previous studies reporting weaker-than-observed low-frequency NAO variability in CMIP5 models (Kravtsov et al. 2014; Kravtsov 2017; Wang et al. 2017).

d. Some remarks on multidecadal NAO variability

As shown above, the observed NAO exhibits more pronounced multidecadal variability than that simulated in CESM1-CAM5. Our analysis thus leads us to the following fundamental questions: what explains

the observed multidecadal variability in NAO and why is this variability deficient in CESM1-CAM5? On intraseasonal time scales, the NAO can be primarily explained by intrinsic atmospheric processes (e.g., Feldstein 2000), but its low-frequency variability likely arises from interactions with other climate components (Feldstein 2002; Czaja et al. 2003) or from external forcing (e.g., Gray et al. 2013). Multidecadal variability, in particular, is likely associated with the ocean due to its long thermal inertia and adjustment time scales. SST forcings from both the tropical Indo-Pacific Ocean (e.g., Hoerling et al. 2001; Bader and Latif 2003) and NASST (e.g., Peings and Magnusdottir 2014, 2016) have been proposed as a source of the low-frequency NAO variability. In particular, it has been shown that a negative NAO phase follows a positive AMV phase in both observations (Peings and

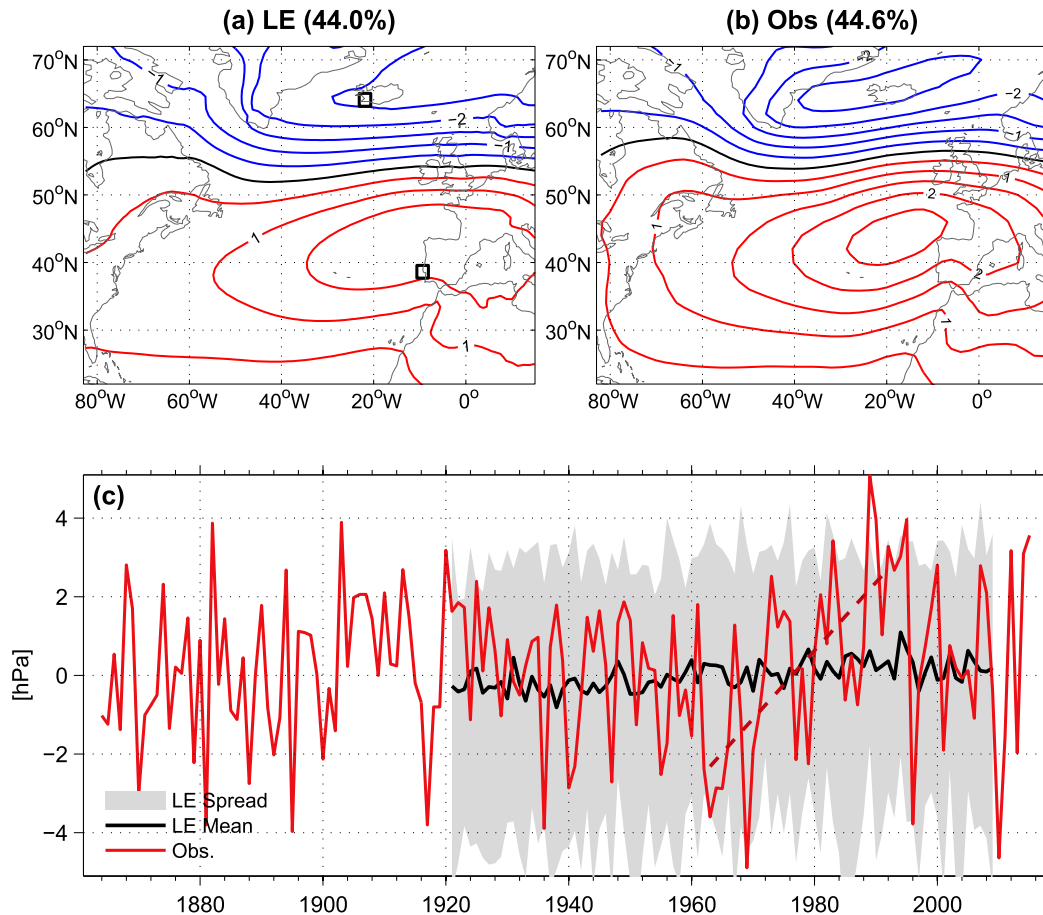**(a) LE (44.0%)**   **(b) Obs (44.6%)**

Fig. 11. EOF1 of the DJFM sea level pressure over the North Atlantic sector bounded by 20°–80°N, 90°W–40°E from (a) ensemble mean LE and (b) 20CRv2c, with the fraction of the total variance accounted by EOF1 given in parentheses, and (c) DJFM station-based NAO index anomaly (relative to the 1921–2009 mean) from the ensemble mean of LE (black) and observations (Hurrell et al. 2017; red). The shading shows the range of the LE ensemble. The maximum 30-yr trend in the observed NAO index is shown by the dashed maroon line. In (a) and (b), the unit is hPa and the contour interval is 0.5 hPa. In (a), the boxed areas indicate the locations of the sea level pressure values used in the station-based NAO index.

Magnusdottir 2014; Gastineau and Frankignoul 2015) and model simulations (Gastineau et al. 2013; Omrani et al. 2014; Peings and Magnusdottir 2016). This suggests that a positive feedback between the AMV and NAO may be implicated, with positive (negative) SST anomalies in the North Atlantic leading to a negative (positive) NAO phase that reinforces the positive (negative) AMV (e.g., Czaja et al. 2003; Farneti and Vallis 2011). This feedback, in turn, possibly enhances the low-frequency NAO variability.

The NAO-like atmospheric response to AMV, however, seems to be weaker or absent in models, compared to observations (Gastineau et al. 2013; Peings et al. 2016). This appears to be also the case in CESM1-CAM5. Figure 15 shows the lead–lag correlations between the low-pass filtered NAO and SPNA SST from CESM1-CAM5 and observations. The correlations are

substantially enhanced when the NAO leads the SPNA SST by around 10 years in both CESM1-CAM5 and observations,[3] indicative of the delayed SPNA SST response to the NAO via ocean dynamics (i.e., the AMOC), discussed in section 3b. The correlations are also negatively enhanced simultaneously and when the SPNA SST leads by few years, suggesting both the instantaneous SPNA SST response to the NAO and a NAO-like response to the SPNA SST. However, the negative correlations in CESM1-CAM5 decorrelate much faster than in

---

[3] The observed correlations are not statistically significant because of the small number of effective degrees of freedom estimated by the formula of Bretherton et al. (1999), but become significant if a less conservative estimate is used, such as one depending on low-pass filtering frequency (e.g., Trenberth 1984).
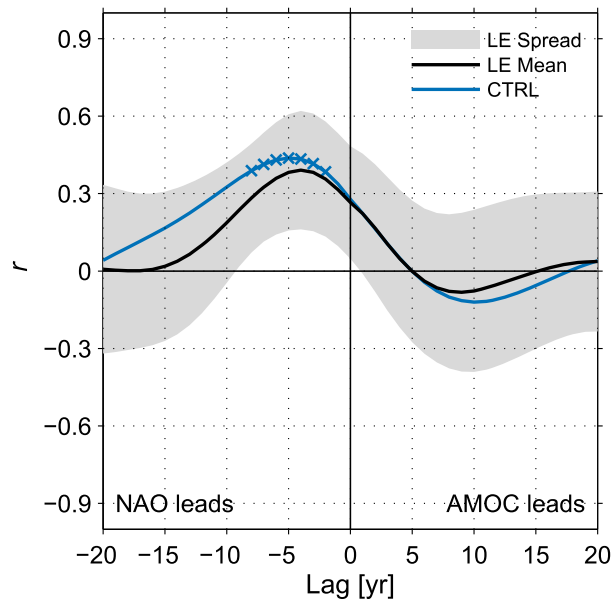
FIG. 12. Lead–lag correlations between the low-pass filtered NAO index and AMOC PC1 time series from LE (gray shading for the ensemble spread and black line for the ensemble mean) and CTRL (blue). The ensemble mean is removed from LE prior to the computation of the correlations. Statistically significant correlations at 95% confidence level are indicated by crosses in CTRL.

observations, suggesting less persistent coupling between the SPNA SST and NAO in CESM1-CAM5. We note that correlations are much lower with the low-pass filtered NASST (i.e., AMV) instead of the SPNA SST in CESM1-CAM5 (not shown), indicating that the NAO response to the tropical AMV (Peings and Magnusdottir 2016) is also not operating in CESM1-CAM5.

To some extent, the weak atmospheric response in CESM1-CAM5 appears to be due to the unrealistic spatial structure and/or amplitude of the simulated AMV. When the observed AMV pattern is applied as a forcing in CESM1-CAM5 ensembles (Ruprich-Robert et al. 2017) or prescribed in forced CAM5 simulations (Peings and Magnusdottir 2016), a NAO-like response is more robust. Another factor contributing to the weak response could be that the low top of CAM5 cannot represent stratospheric dynamics properly. Recent studies have suggested that a NAO-like atmospheric response to AMV requires a high-top atmospheric model (Omrani et al. 2014, 2016), although the influence of a high top is not clear in CAM5 (Peings and Magnusdottir 2016). In any case, the impact of realistic boundary conditions and better resolved stratospheric dynamics on low-frequency NAO variability has not been examined in previous studies.

The forced CAM5 simulations introduced in section 2 can help address this question. Figure 16 shows the

distributions of the 5-, 15-, and 30-yr moving NAO trends from these two simulations along with those from LE and CTRL (as in the top panels of Fig. 14). Also, the same distributions from a synthetic white noise time series (WN)[4] are displayed in Fig. 16. Despite different sample sizes due to both different ensemble sizes and time periods, the distributions are very similar between these two forced CAM5 simulations and LE across all trend lengths, which are in turn close to those from WN and CTRL. Despite the more realistic (SST and sea ice) boundary conditions and presumably better representation of stratospheric dynamics in LT and MT, there is no enhancement of multidecadal NAO variability, and as in CESM1-CAM5, the simulated NAO is not distinguishable from WN. A comparison of the distributions of the moving trends for the individual ensemble members of LT and MT with the observed NAO also confirms that the multidecadal NAO variability in these simulations is substantially weaker (not shown). These results suggest that realistic boundary conditions do not necessarily lead to a better representation of NAO variability on multidecadal time scales, and thus the weak multidecadal NAO variability in the hierarchy of simulations using CAM5 can probably be attributed to deficiencies in CAM5 itself, including horizontal and vertical resolution and parameterized physics, or air–sea coupling details.

With the observed NAO time series spanning only about 150 years, there are only a couple of independent cycles of the multidecadal variability of 60 years or so. Multidecadal NAO variability has also been examined using various proxy records (e.g., Cook et al. 1998; Olsen et al. 2012). Box plots of 30-yr moving trends from a NAO reconstruction (Ortega et al. 2015; see section 2d) show a distribution easily encompassing the observed one (Fig. 17). The spectral analysis of the NAO reconstruction also reveals enhanced power on broad multidecadal time scales (~55–100 yr; not shown). These results suggest that there has been prominent multidecadal variability in NAO during the last 1000 years and add another piece of evidence that both the CESM1-CAM5 and the stand-alone CAM5 simulations underestimate NAO variability on multidecadal time scales.

4. Summary and discussion

We have examined the low-frequency variability in the North Atlantic simulated in historical large ensemble and preindustrial control simulations using CESM1-CAM5,

[4] We randomly generate an 89-yr-long (the same length as the 1921–2009 segment of LE) white noise time series with zero mean and unit standard deviation and repeat this process 5000 times.
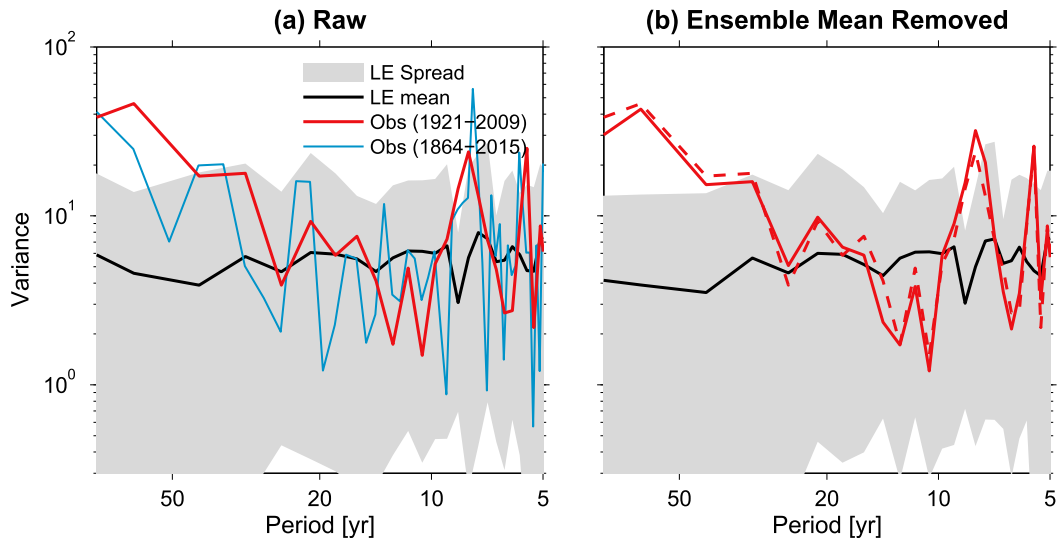
FIG. 13. As in Fig. 6, but for the DJFM NAO indices with observations from Hurrell's NAO index.

focusing on AMOC, SPNA SST, NAO, and Sahel rainfall, in comparison with available observations and a forced ocean–sea ice simulation. A key finding of our study is that all these variables exhibit substantially weaker multidecadal variability in CESM1-CAM5 than in observational estimates, while relatively high-frequency variability (i.e., interannual to decadal) is comparable.

We argue that this weak multidecadal North Atlantic variability in CESM1-CAM5 is mechanistically related to the underestimated multidecadal variability in the simulated NAO. The NAO controls deep-water formation in the Labrador Sea through its associated surface buoyancy fluxes and thus plays a major role in driving low-frequency variability in AMOC (e.g., Eden and Jung 2001; Yeager and Danabasoglu 2014). AMOC then conveys these signals to the ocean surface in the extratropics, and the atmospheric response to these SST anomalies further impacts climate around the North Atlantic (Brown et al. 2016b; Drews and Greatbatch 2017). An immediate ramification of the weak multidecadal AMOC variability is weak multidecadal variability in SPNA SST, which is connected to AMOC through the meridional heat transport convergence. As the SPNA SST signal contributes substantially to the basin-scale AMV, including perhaps as a forcing for low-frequency tropical signals, the model AMV is also too weak. Weak multidecadal variability is then also apparent in Sahel rainfall as it is closely linked to the AMV (e.g., Ting et al. 2011).

This weak multidecadal North Atlantic variability in LE is consistent with the findings of Frankcombe et al. (2015), Kravtsov (2017), and Wang et al. (2017), who analyzed a suite of historical CMIP5 simulations. Such consistently weak multidecadal North Atlantic variability found in CMIP5 models may help to explain the results of Clement et al. (2015), who showed that the power spectra of NASST from select CMIP5 models are not distinguishable from those generated by atmospheric noise alone. Indeed, the power spectra of the NASST simulated in CTRL and LE (after removing the ensemble mean) show no significant peaks in multidecadal bands relative to a red-noise null hypothesis (not shown). As demonstrated in this study, however, this is likely related to the overly "white" simulated NAO, in contrast to the real-world NAO, which shows substantial power in multidecadal frequency bands. Delworth and Zeng (2016) and Mecking et al. (2014) show that the response of the AMOC and hence the NASST is almost linear in the sense that their variance is enhanced in the frequency bands where the variance of the NAO is enhanced. Therefore, if the multidecadal spectral power of the NAO were more realistic in LE or in other CMIP5 models, the AMOC and NASST variability would likely be enhanced on multidecadal time scales, and hence more consistent with observations.

The findings of this study seem to be consistent with the analysis by Wang et al. (2017), who found an underestimated low-frequency variability of surface air temperature (SAT) in the Northern Hemisphere (NHSAT) in CMIP5 models compared to observations, due to the underestimated low-frequency NAO variability. NAO-induced surface heat flux forcing and subsequent AMOC response appear to explain a large fraction of the NHSAT changes during the twentieth century (Delworth et al. 2016). A key factor for this link
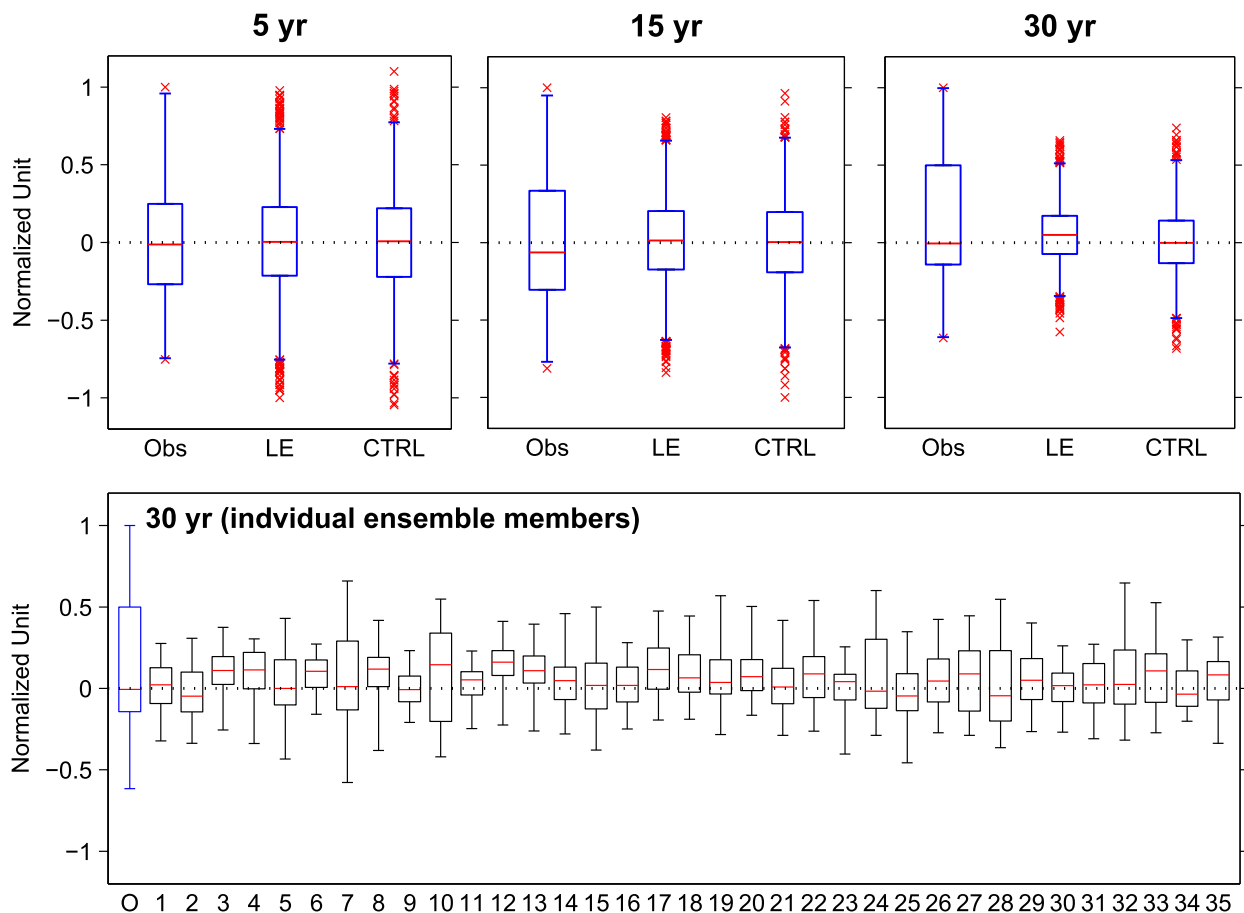
FIG. 14. As in Fig. 8, but for the raw annual NAO index. The moving trends from LE are normalized to the corresponding maximum trend from the station-based NAO index. The sample sizes are 85, 2975, and 1397 (5 yr); 75, 2625, and 1387 (15 yr); and 60, 2100, and 1372 (30 yr) for the observations, LE, and CTRL, respectively, in the top panel. The sample size for all box plots in the bottom panel is 60.

seems to be a sustained anomalous heat supply by the ocean circulation (i.e., AMV) and subsequent Arctic sea ice melting, which can impact SATs in the northern high to middle latitudes (Semenov et al. 2010; Delworth et al. 2016). In addition, surface heat flux differences in deep water formation and sea ice regions in the North Atlantic are suggested to explain the intermodel spread in the magnitude of the unforced global SAT changes in CMIP5 models (Brown et al. 2016a), which further underscores the importance of low-frequency North Atlantic variability in regulating low-frequency NHSAT and global SAT variability.

The NAO–AMOC–NASST coupled mechanism, however, may not be fully represented by every climate model. Some models have difficulty capturing the spatial structure of the NAO (Davini and Cagnazzo 2014). For example, models that exhibit large biases in the location and strength of the NAO pressure centers may not show a strong link between NAO and deep-water formation in the Labrador Sea. Some models

also fail to simulate realistic sea ice extent in the Labrador Sea, which is sometimes entirely covered with ice. In that case, deep-water formation is completely shut down in the Labrador Sea, breaking the mechanistic link that we highlight. Additionally, even if the Labrador Sea is ice free, stratification in the Labrador Sea can be too weak to allow for much variability in deep-water formation there (Danabasoglu et al. 2016). Therefore, it would seem that models must first and foremost have a credible representation of the North Atlantic mean climate for the link described above to be operational.

Low-frequency AMOC variability can also be controlled by mechanisms other than low-frequency NAO-induced buoyancy flux forcing. In some coupled models, the east Atlantic and Scandinavian patterns, the next leading atmospheric modes after the NAO, appear to be more actively involved in driving low-frequency AMOC variability (Msadek and Frankignoul 2009; Medhaug et al. 2012; Ruprich-Robert and Cassou 2015). In
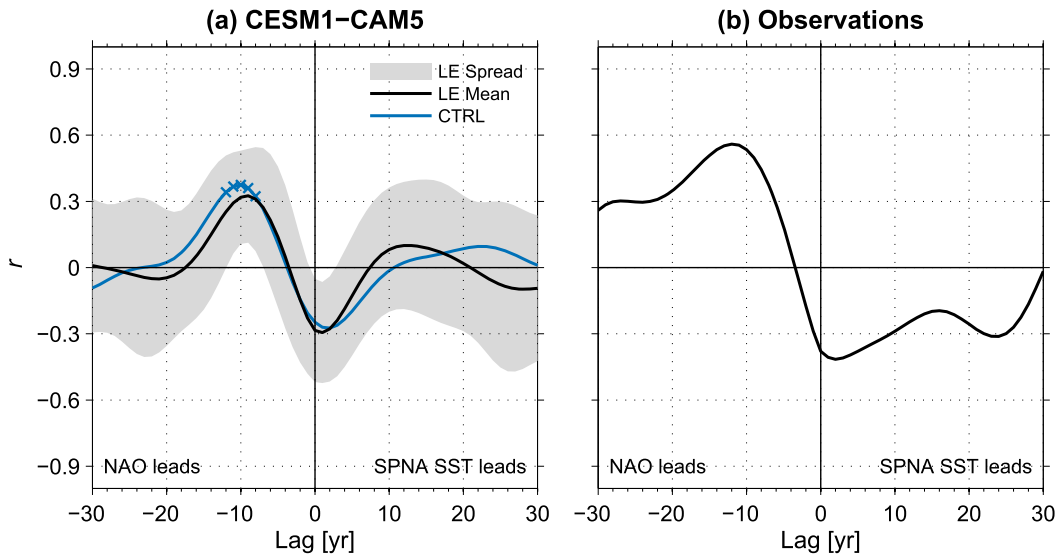
FIG. 15. As in Fig. 12, but between the low-pass filtered NAO index and SPNA SST time series from (a) CESM1-CAM5 (i.e., LE and CTRL) and (b) observations (between the station-based NAO index and HadISST). The considered time period in (b) is 1870–2015 and both time series are linearly detrended before computing the correlations.

CESM1-CAM5, a link between the east Atlantic pattern and SPNA SST is also found, but the correlations are weaker than those with NAO and not statistically significant. Furthermore, these modes in observations do not exhibit an enhanced variance on multidecadal time scales, and their spectral power is well within the spectral power range of the corresponding mode simulated in LE in all frequency bands. Therefore, these findings further support the notion that the NAO is a dominant player in the observed multidecadal North Atlantic variability and that the lack of an energetic multidecadal NAO variability in CESM1-CAM5 is

likely the primary reason for the weak multidecadal North Atlantic variability.

The strong influence of the forced signal in the multidecadal SPNA SST variability (Fig. 7) seems to be consistent with recent studies by Murphy et al. (2017) and Bellomo et al. (2017), which put forward external forcings as the primary factor in driving observed low-frequency NASST changes. The latter study is particularly relevant to the present work because they also use LE for their analysis and highlight the predominant forced signal in the SPNA SST. However, as shown in Fig. 5e, the amplitude of the forced SPNA SST signal in
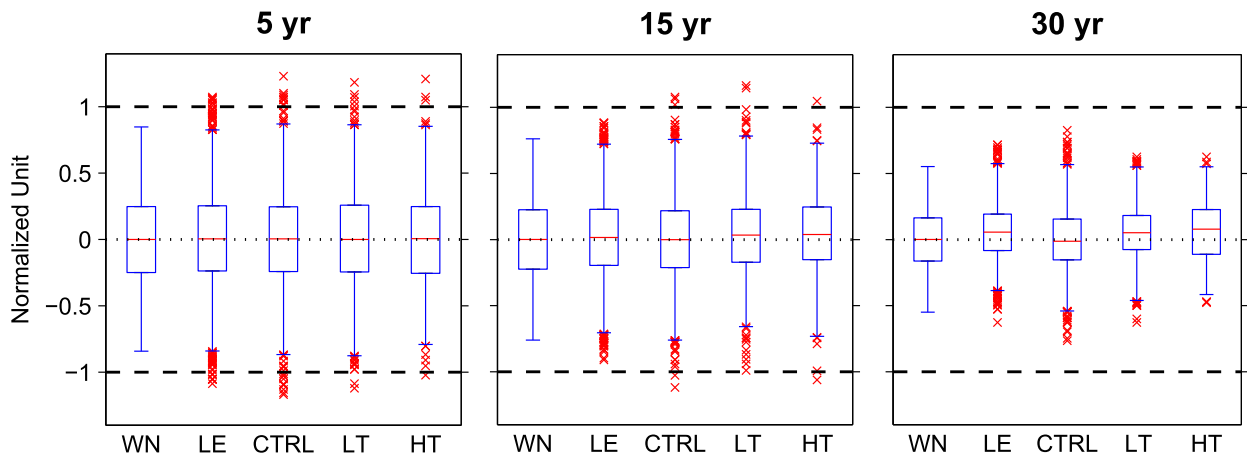

FIG. 16. As in the top panels of Fig. 14, with the inclusion of the distributions of the moving trends from WN, LT, and MT, but without those from observations. The sample sizes for WN, LT, and MT are 425 000, 1300, and 590 (5 yr); 375 000, 1200, and 490 (15 yr); and 300 000, 1050, and 340 (30 yr), respectively. Outliers are omitted for WN.
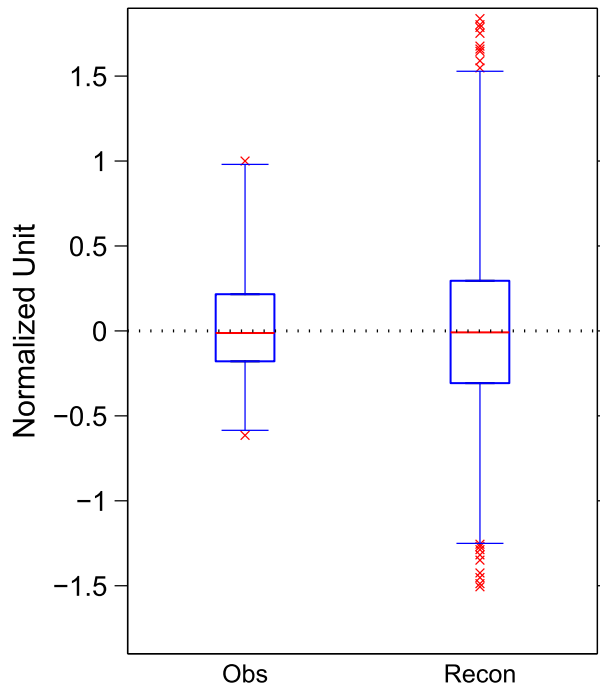
FIG. 17. Box plots of 30-yr moving trends from the NAO reconstruction (1049–1969) and the full record length (1864–2015) of observed NAO index. The sample sizes are 892 and 123 for the reconstruction and observations, respectively.

LE is substantially weaker than that of the observed multidecadal SPNA SST variability, even though the forced signal in LE appears to be stronger than in other CMIP5 models (M. Ting 2017, personal communication). Furthermore, the phase of the forced signal in LE is not aligned with the observed phase (in quadrature). In addition, CESM1 decadal prediction simulations show substantially higher skill at predicting SPNA SST changes at multiyear lead times than uninitialized historical simulations (Yeager et al. 2015). Therefore, while there may be some contributions from the external forcings, it seems more reasonable to view that internal variability plays a larger role in the observed low-frequency SPNA SST and NASST variability.

Although we stress in this study the primary role of low-frequency NAO variability in driving other low-frequency North Atlantic variability, the low-frequency NAO variability itself is likely caused by coupling between the ocean and atmosphere. In particular, the North Atlantic is suggested as an active region for this coupling, as discussed in section 3d. However, while we find conspicuous low-frequency variability in AMOC (as well as in some other fields in the SPNA; see Yeager et al. 2015), tightly associated with observed NAO, in the forced ocean simulation with observational

atmospheric states (i.e., FO) there is no enhancement of low-frequency NAO variability in CAM5 simulations forced with observed SST and sea ice. This leads us to conjecture that the simulated weak low-frequency NAO variability in CESM1-CAM5 is likely due to deficiencies in either CAM5 itself or air–sea coupling details rather than deficiencies in the ocean (or sea ice) component. Given the underestimated low-frequency NAO variability in most CMIP5 models (Wang et al. 2017), this problem may not be specific to CESM1-CAM5 but rather shared with other state-of-the-art coupled models.

The underestimated low-frequency NAO variability arising possibly from the lack of the coupled feedbacks has implications for the decadal prediction of North Atlantic climate. The NAO in current climate prediction systems is only predictable, at most, one to two years ahead (Dunstone et al. 2016). If the positive feedback between NASST and NAO is indeed a mechanism necessary to enhance low-frequency NAO variability, an improvement in the modeled representation of this feedback may lead to skillful prediction of NAO at longer lead times. This may in turn yield a more long-lasting prediction skill in NASST, which has been so far benefited mostly from the initialization of realistic ocean conditions, but not from air–sea heat exchanges (Yeager et al. 2012). Therefore, a better representation of this missing feedback in coupled systems may yield more skillful predictions of the North Atlantic climate.

We used the AMOC solutions from FO to compare with the simulated low-frequency AMOC variability in CESM1-CAM5. The estimated AMOC from FO is subject to uncertainties arising from, for example, the coarse resolution and inadequate representations of parameterized physics in the ocean model, along with uncertainties in the forcing fields. However, previous studies (Delworth and Greatbatch 2000; Zhu and Jungclaus 2008; Farneti and Vallis 2011) have shown that, when forced with surface fluxes taken from coupled simulations, the ocean components largely reproduce the phase and magnitude of the low-frequency AMOC variability in the coupled simulation. Therefore, as long as the forcing data are realistic, using simulated AMOC from a forced ocean simulation as a proxy for the observed estimate seems to be reasonable, particularly when the same ocean model is used for both forced ocean and coupled simulations as in this study, so that model biases remain the same. We also note that the FO simulation was used to initialize CESM1 decadal prediction simulations that exhibit very high skill at predicting SPNA SST changes at multiyear lead times (Yeager et al. 2015).
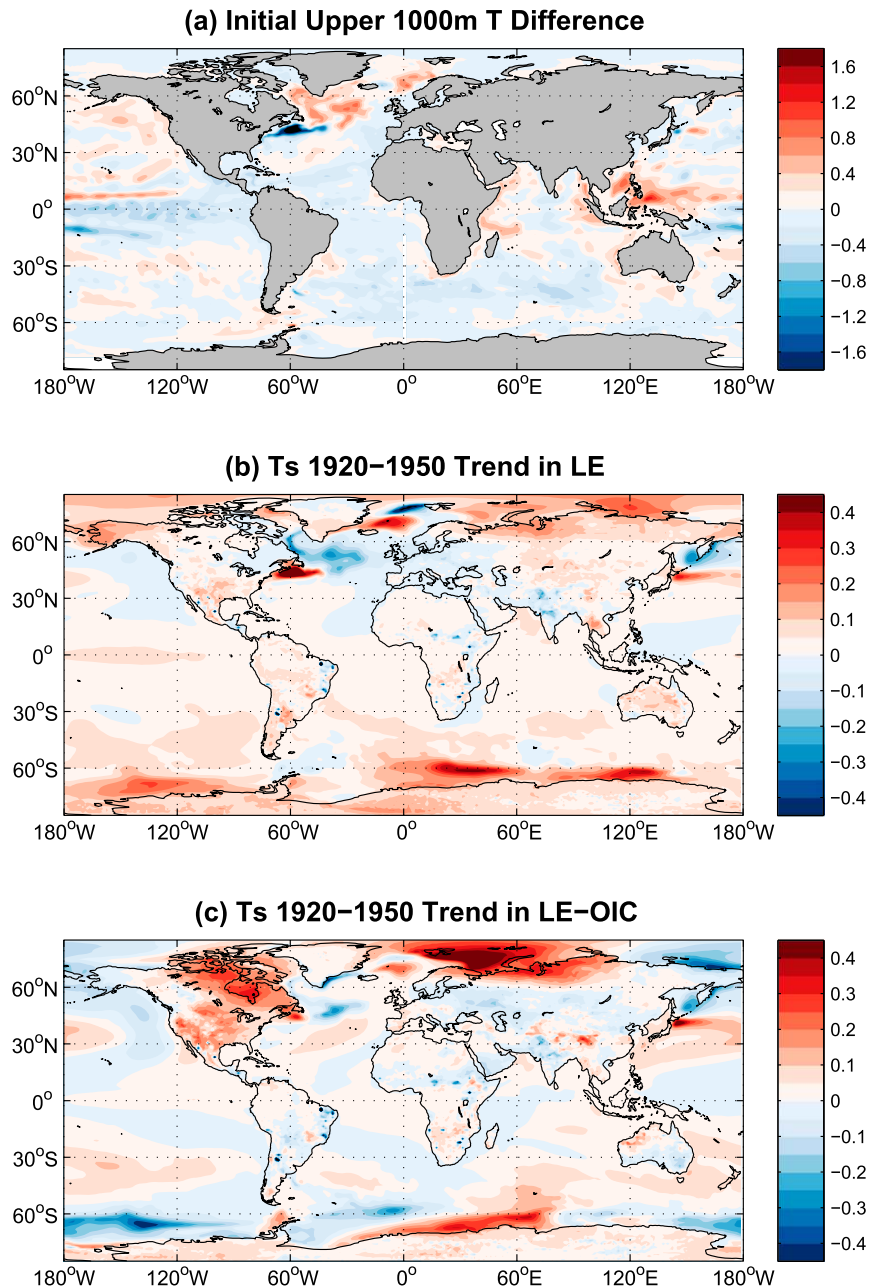
(a) Initial Upper 1000m T Difference

(b) Ts 1920−1950 Trend in LE

(c) Ts 1920−1950 Trend in LE−OIC

FIG. A1. (a) Differences in the upper 1000 m-mean potential temperature (*T*; LE minus LE-OIC) on 1 Jan 1920. First 31-yr trends (1920–1950) in the ensemble mean surface air temperature (Ts) from (b) LE and (c) LE-OIC. Units are in °C.

As is the case for many studies examining simulated multidecadal climate variability, comparing the simulated North Atlantic multidecadal variability with observational records is hampered by the limited independent cycles in relatively short observational records. It is thus open to question whether the observed multidecadal variability is statistically robust. However, as discussed in section 3d for NAO, the existence of persistent multidecadal variability in the North Atlantic is underpinned by numerous studies examining climate proxy records. In particular, evidence of multidecadal variability in NASST is found in tree-ring (Gray et al. 2004), coral (Kilbourne et al. 2008), ice-core, lacustrine, and marine proxy records (Knudsen et al. 2011, and references therein) around the North Atlantic.
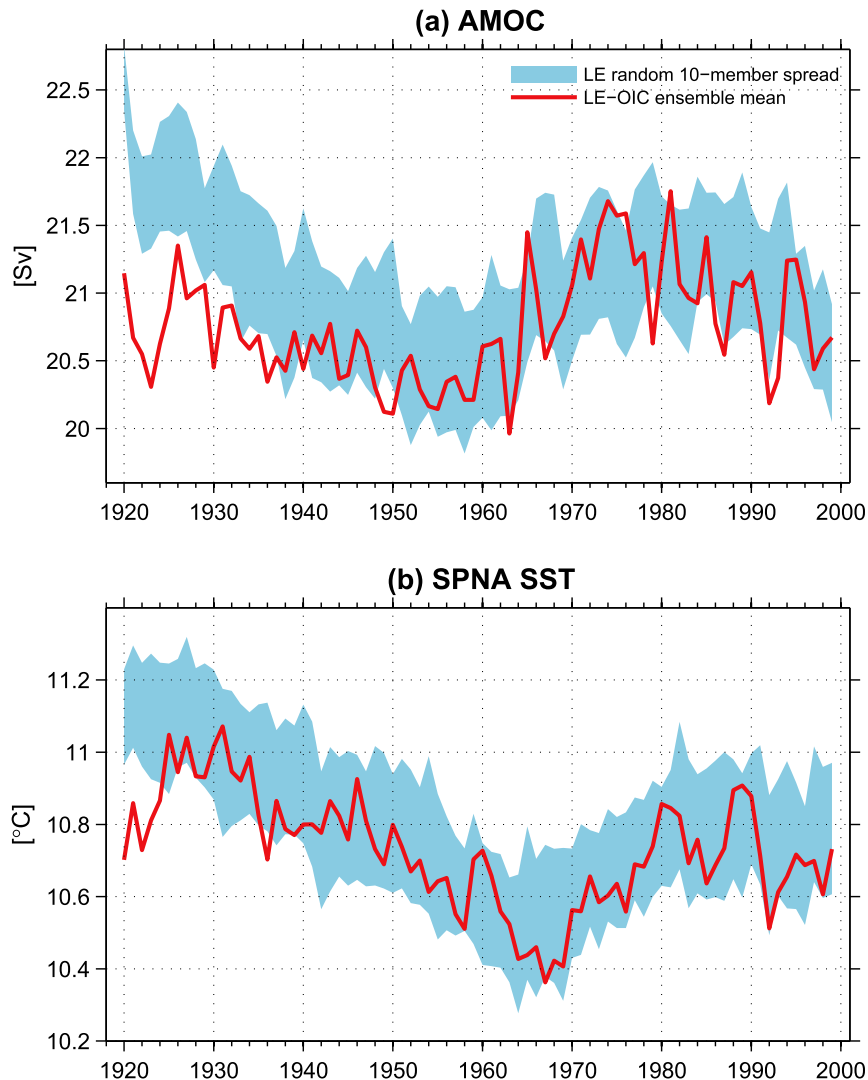
(a) AMOC

(b) SPNA SST

FIG. A2. Ensemble mean from LE-OIC (line) and 1st–99th percentile range of subsampled ensemble means from LE (shading) of (a) the maximum AMOC strength at 26.5°N and (b) SPNA SST.

An important advantage of using large ensembles such as LE is that they can help to quantify uncertainties arising from internal variability of the climate system (e.g., Swart et al. 2015; Deser et al. 2017). However, as shown above, multidecadal internal variability is probably underestimated in such simulations. Thus, care must be taken when examining uncertainties in climate change arising from multidecadal internal variability.

and Information Systems Laboratory, sponsored by the NSF.

APPENDIX

CESM LE with Different Ocean Initial Conditions

To sample internal variability arising from different ocean initial conditions, we performed additional historical simulations with 10 ensemble members using the identical model configuration as in LE (Kay et al. 2015). Obtaining a new set of ocean initial conditions is a nontrivial exercise. Specifically, as detailed in Kay et al. (2015), the starting point of the LE simulations is an 1850 preindustrial control simulation (CTRL). Although CTRL was eventually integrated for 2200 years, the first ensemble member started from 1 January 402 from CTRL and was integrated from 1850 to 1920. The subsequent ensemble members were all started from 1 January 1920 of this first ensemble member. All ensemble members used the same ocean initial conditions from the first ensemble member, and the ensemble spread was obtained by applying random round-off level perturbations to the air temperature in the atmospheric restart files. Thus, simply using sufficiently different ocean initial conditions in 1920 to start a new ensemble set is not easily achievable because we have only one ocean state in 1920 from the first ensemble member. Consequently, we adopted the following approach, using the state of AMOC as a key metric to subjectively obtain sufficiently different ocean initial conditions. As CTRL shows drifts throughout the 2200-yr simulation, a consideration was to use ocean states that do not deviate significantly from the mean ocean state of year 402. Thus, we performed four simulations for the 1850–1920 period, starting at years 374, 384, 466, and 496 of CTRL, sampling high, low, increasing, and decreasing AMOC, respectively. In 1920, only one of these simulations (start year 496) showed an AMOC state that was deemed sufficiently different than that of the original LE (see Fig. A2a). Thus, the ocean state from this first ensemble member was used as the ocean initial conditions for the remaining nine ensemble members in which the same round-off level perturbations in air temperature were used to create ensemble spread. This new ensemble set, referred to as LE-OIC, was integrated for the 1920–99 period. We note that our ensemble generation approach is essentially the same as discussed in Hawkins et al. (2016).

We present the differences in the upper 1000-m mean potential temperature distributions on 1 January 1920 between LE and LE-OIC in Fig. A1a, showing substantially warmer SPNA conditions in LE consistent with a stronger AMOC state in LE than in LE-OIC.

Also, there are differences in the western and eastern tropical Pacific temperatures, indicating that a different ENSO state is sampled. Although the ensemble means of most of the surface variables from LE and LE-OIC converge within a decade in much of the globe, climate transitions appear to be different for a long time due to the different ocean initial conditions. For example, Figs. A1b and A1c show the trends in surface air temperature for the first 31 years (1920–50) from LE and LE-OIC, respectively, and reveal substantially different trends in high-latitude regions surrounding the SPNA. Also, rather surprisingly, the sign of the trends is even opposite in a broad region in the Pacific sector of the Southern Ocean, which persists for an even longer period than depicted in Fig. A1.

In the main text, we show that the forced signal (i.e., ensemble mean) of the AMOC and SPNA SST in the earlier period of the historical simulations is quite different, most notably in the AMOC, between LE and LE-OIC. However, given the substantially different ensemble sizes of LE and LE-OIC (35 vs 10), it may be useful to test if the different ensemble means are statistically significant. We evaluate this using a Monte Carlo method: We randomly subsample 10-member ensembles from LE for 5000 times, and consider them to be significantly different when the ensemble mean of LE-OIC falls outside of the first to 99th percentile range of the subsampled ensemble means of LE. Indeed, the ensemble mean AMOC from LE-OIC is found to be below the range of LE for more than 15 years, and it stays near the lower limit for another 25 years (Fig. A2a). Because of this difference, if one measures the long-term trend in AMOC between 1920 and 1999, it is positive in LE-OIC while negative in LE, and they are statistically different (not shown). In contrast, the ensemble mean SPNA SST from LE-OIC is only significantly different for the first 5 years (Fig. A2b). However, the subsurface temperatures in the SPNA reveal a significantly different ensemble mean for a longer time span comparable to the AMOC (not shown).

REFERENCES

Ba, J., and Coauthors, 2014: A multi-model comparison of Atlantic multidecadal variability. *Climate Dyn.*, **43**, 2333–2348, https://doi.org/10.1007/s00382-014-2056-1.

Bader, J., and M. Latif, 2003: The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. *Geophys. Res. Lett.*, **30**, 2169, https://doi.org/10.1029/2003GL018426.

Bellomo, K., L. N. Murphy, M. A. Cane, A. C. Clement, and L. M. Polvani, 2017: Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble. *Climate Dyn.*, https://doi.org/10.1007/s00382-017-3834-3, in press.

Biastoch, A., C. W. Böning, J. Getzlaff, J.-M. Molines, and G. Madec, 2008: Causes of interannual–decadal variability in the meridional overturning circulation of the midlatitude North Atlantic Ocean. *J. Climate*, **21**, 6599–6615, https://doi.org/10.1175/2008JCLI2404.1.

Böning, C. W., M. Scheinert, J. Dengg, A. Biastoch, and A. Funk, 2006: Decadal variability of subpolar gyre transport and its reverberation in the North Atlantic overturning. *Geophys. Res. Lett.*, **33**, L21S01, https://doi.org/10.1029/2006GL026906.

Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009, https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2.

Brown, P. T., W. Li, J. H. Jiang, and H. Su, 2016a: Spread in the magnitude of climate model interdecadal global temperature variability traced to disagreements over high-latitude oceans. *Geophys. Res. Lett.*, **43**, 12 543–12 549, https://doi.org/10.1002/2016GL071442.

——, M. S. Lozier, R. Zhang, and W. Li, 2016b: The necessity of cloud feedback for a basin-scale Atlantic multidecadal oscillation. *Geophys. Res. Lett.*, **43**, 3955–3963, https://doi.org/10.1002/2016GL068303.

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., 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.

Cook, E. R., R. D. D'Arrigo, and K. R. Briffa, 1998: A reconstruction of the North Atlantic Oscillation using tree-ring chronologies from North America and Europe. *Holocene*, **8**, 9–17, https://doi.org/10.1191/095968398677793725.

Cunningham, S. A., and Coauthors, 2007: Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. *Science*, **317**, 935–938, https://doi.org/10.1126/science.1141304.

Czaja, A., A. W. Robertson, and T. Huck, 2003: The role of Atlantic Ocean–atmosphere coupling in affecting North Atlantic Oscillation variability. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr.*, Vol. 134, Amer. Geophys. Union, 147–172.

Danabasoglu, G., 2008: On multidecadal variability of the Atlantic meridional overturning circulation in the Community Climate System Model version 3. *J. Climate*, **21**, 5524–5544, https://doi.org/10.1175/2008JCLI2019.1; Corrigendum, **22**, 1586–1587, https://doi.org/10.1175/2008JCLI2986.1.

——, 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.

——, and Coauthors, 2014: North Atlantic simulations in Coordinated Ocean-Ice Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modell.*, **73**, 76–107, https://doi.org/10.1016/j.ocemod.2013.10.005.

——, and Coauthors, 2016: North Atlantic simulations in Coordinated Ocean-Ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability. *Ocean Modell.*, **97**, 65–90, https://doi.org/10.1016/j.ocemod.2015.11.007.

Davini, P., and C. Cagnazzo, 2014: On the misinterpretation of the North Atlantic Oscillation in CMIP5 models. *Climate Dyn.*, **43**, 1497–1511, https://doi.org/10.1007/s00382-013-1970-y.

Day, J. J., J. C. Hargreaves, J. D. Annan, and A. Abe-Ouchi, 2012: Sources of multi-decadal variability in Arctic sea ice extent. *Environ. Res. Lett.*, **7**, 034011, https://doi.org/10.1088/1748-9326/7/3/034011.

Delworth, T. L., and R. J. Greatbatch, 2000: Multidecadal thermohaline circulation variability driven by atmospheric surface flux forcing. *J. Climate*, **13**, 1481–1495, https://doi.org/10.1175/1520-0442(2000)013<1481:MTCVDB>2.0.CO;2.

——, 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.

——, S. Manabe, and R. J. Stouffer, 1993: Interdecadal variations of the thermohaline circulation in a coupled ocean–atmosphere model. *J. Climate*, **6**, 1993–2011, https://doi.org/10.1175/1520-0442(1993)006<1993:IVOTTC>2.0.CO;2.

——, F. Zeng, G. A. Vecchi, X. Yang, L. Zhang, and R. Zhang, 2016: The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere. *Nat. Geosci.*, **9**, 509–512, https://doi.org/10.1038/ngeo2738.

——, ——, 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 Atlantic multidecadal oscillation. *J. Climate*, **30**, 3789–3805, https://doi.org/10.1175/JCLI-D-16-0358.1.

Deser, C., J. W. Hurrell, and A. S. Phillips, 2017: The role of the North Atlantic Oscillation in European climate projections. *Climate Dyn.*, **49**, 3141–3157, https://doi.org/10.1007/s00382-016-3502-z.

Dong, B., and R. T. Sutton, 2005: Mechanism of interdecadal thermohaline circulation variability in a coupled ocean–atmosphere GCM. *J. Climate*, **18**, 1117–1135, https://doi.org/10.1175/JCLI3328.1.

Drews, A., and R. J. Greatbatch, 2017: Evolution of the Atlantic multidecadal variability in a model with an improved North Atlantic Current. *J. Climate*, **30**, 5491–5512, https://doi.org/10.1175/JCLI-D-16-0790.1.

Dunstone, N., D. Smith, A. Scaife, L. Hermanson, R. Eade, N. Robinson, M. Andrews, and J. Knight, 2016: Skillful predictions of the winter North Atlantic Oscillation one year ahead. *Nat. Geosci.*, **9**, 809–814, https://doi.org/10.1038/ngeo2824.

Eden, C., and T. Jung, 2001: North Atlantic interdecadal variability: Oceanic response to the North Atlantic Oscillation (1865–1997). *J. Climate*, **14**, 676–691, https://doi.org/10.1175/1520-0442(2001)014<0676:NAIVOR>2.0.CO;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.

Farneti, R., and G. K. Vallis, 2011: Mechanisms of interdecadal climate variability and the role of ocean–atmosphere coupling. *Climate Dyn.*, **36**, 289–308, https://doi.org/10.1007/s00382-009-0674-9.

Feldstein, S. B., 2000: The timescale, power spectra, and climate noise properties of teleconnection patterns. *J. Climate*, **13**, 4430–4440, https://doi.org/10.1175/1520-0442(2000)013<4430:TTPSAC>2.0.CO;2.

——, 2002: The recent trend and variance increase of the annular mode. *J. Climate*, **15**, 88–94, https://doi.org/10.1175/1520-0442(2002)015<0088:TRTAVI>2.0.CO;2.

Folland, C. K., 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.

Frankcombe, L. M., M. H. England, M. E. Mann, and B. A. Steinman, 2015: Separating internal variability from the externally forced climate response. *J. Climate*, **28**, 8184–8202, https://doi.org/10.1175/JCLI-D-15-0069.1.

Gastineau, G., and C. Frankignoul, 2015: Influence of the North Atlantic SST variability on the atmospheric circulation during the twentieth century. *J. Climate*, **28**, 1396–1416, https://doi.org/10.1175/JCLI-D-14-00424.1.

——, F. D'Andrea, and C. Frankignoul, 2013: Atmospheric response to the North Atlantic Ocean variability on seasonal to decadal time scales. *Climate Dyn.*, **40**, 2311–2330, https://doi.org/10.1007/s00382-012-1333-0.

Gillett, N. P., and J. C. Fyfe, 2013: Annular mode changes in the CMIP5 simulations. *Geophys. Res. Lett.*, **40**, 1189–1193, https://doi.org/10.1002/grl.50249.

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, L. J., and Coauthors, 2013: A lagged response to the 11 year solar cycle in observed winter Atlantic/European weather patterns. *J. Geophys. Res. Atmos.*, **118**, 13 405–13 420, https://doi.org/10.1002/2013JD020062.

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.

Griffies, S. M., M. Winton, B. Samuels, G. Danabasoglu, S. Yeager, S. Marsland, H. Drange, and M. Bentsen, 2012: Datasets and protocol for the CLIVAR WGOMD Coordinated Ocean Sea-Ice Reference Experiments (COREs). WCRP Rep. 21/2012, 21 pp.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014: Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *Int. J. Climatol.*, **34**, 623–642, https://doi.org/10.1002/joc.3711.

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.

Held, I. M., T. L. Delworth, J. Lu, K. L. Findell, and T. R. Knutson, 2005: Simulation of Sahel drought in the 20th and 21st centuries. *Proc. Natl. Acad. Sci. USA*, **102**, 17 891–17 896, https://doi.org/10.1073/pnas.0509057102; Corrigendum, **103**, 1152–1153, https://doi.org/10.1073/pnas.0510844103.

Hoerling, M. P., J. W. Hurrell, and T. Xu, 2001: Tropical origins for recent North Atlantic climate change. *Science*, **292**, 90–92, https://doi.org/10.1126/science.1058582.

Hunke, E. C., and W. H. Lipscomb, 2008: CICE: The Los Alamos Sea Ice Model documentation and software user's manual, version 4.0. Los Alamos National Laboratory Tech. Rep. LA-CC-06-012, 76 pp.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**, 676–679, https://doi.org/10.1126/science.269.5224.676.

——, J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski, 2008: A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model. *J. Climate*, **21**, 5145–5153, https://doi.org/10.1175/2008JCLI2292.1.

——, and Coauthors, 2017: The climate data guide: Hurrell North Atlantic Oscillation (NAO) index (station-based). National Center for Atmospheric Research, accessed 7 March 2016, https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based.

Kanzow, T., and Coauthors, 2010: Seasonal variability of the Atlantic meridional overturning circulation at 26.5°N. *J. Climate*, **23**, 5678–5698, https://doi.org/10.1175/2010JCLI3389.1.

Kavvada, A., A. Ruiz-Barradas, and S. Nigam, 2013: AMO's structure and climate footprint in observations and IPCC AR5 climate simulations. *Climate Dyn.*, **41**, 1345–1364, https://doi.org/10.1007/s00382-013-1712-1.

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.

Kilbourne, K. H., T. M. Quinn, R. Webb, T. Guilderson, J. Nyberg, and A. Winter, 2008: Paleoclimate proxy perspective on Caribbean climate since the year 1751: Evidence of cooler temperatures and multidecadal variability. *Paleoceanography*, **23**, PA3220, https://doi.org/10.1029/2008PA001598.

Kim, W. M., S. Yeager, P. Chang, and G. Danabasoglu, 2016: Atmospheric conditions associated with Labrador Sea deep convection: New insights from a case study of the 2006/07 and 2007/08 winters. *J. Climate*, **29**, 5281–5297, https://doi.org/10.1175/JCLI-D-15-0527.1.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann, 2005: A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Res. Lett.*, **32**, L20708, https://doi.org/10.1029/2005GL024233.

——, 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., M.-S. Seidenkrantz, B. H. Jacobsen, and A. Kuijpers, 2011: Tracking the Atlantic multidecadal oscillation through the last 8,000 years. *Nat. Commun.*, **2**, 178, https://doi.org/10.1038/ncomms1186.

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.

——, M. G. Wyatt, J. A. Curry, and A. A. Tsonis, 2014: Two contrasting views of multidecadal climate variability in the twentieth century. *Geophys. Res. Lett.*, **41**, 6881–6888, https://doi.org/10.1002/2014GL061416.

Large, W. G., and S. G. Yeager, 2009: The global climatology of an interannually varying air–sea flux data set. *Climate Dyn.*, **33**, 341–364, https://doi.org/10.1007/s00382-008-0441-3.

Latif, M., C. Böning, J. Willebrand, A. Biastoch, J. Dengg, N. Keenlyside, U. Schweckendiek, and G. Madec, 2006: Is the thermohaline circulation changing? *J. Climate*, **19**, 4631–4637, https://doi.org/10.1175/JCLI3876.1.

McCarthy, G., and Coauthors, 2012: Observed interannual variability of the Atlantic meridional overturning circulation at 26.5°N. *Geophys. Res. Lett.*, **39**, L19609, https://doi.org/10.1029/2012GL052933.

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.

Medhaug, I., and T. Furevik, 2011: North Atlantic 20th century multidecadal variability in coupled climate models: Sea surface temperature and ocean overturning circulation. *Ocean Sci.*, **7**, 389–404, https://doi.org/10.5194/os-7-389-2011.

——, H. R. Langehaug, T. Eldevik, T. Furevik, and M. Bentsen, 2012: Mechanisms for decadal scale variability in a simulated Atlantic meridional overturning circulation. *Climate Dyn.*, **39**, 77–93, https://doi.org/10.1007/s00382-011-1124-z.

Menary, M. B., and A. A. Scaife, 2014: Naturally forced multidecadal variability of the Atlantic meridional overturning circulation. *Climate Dyn.*, **42**, 1347–1362, https://doi.org/10.1007/s00382-013-2028-x.

Msadek, R., and C. Frankignoul, 2009: Atlantic multidecadal oceanic variability and its influence on the atmosphere in a climate model. *Climate Dyn.*, **33**, 45–62, https://doi.org/10.1007/s00382-008-0452-0.

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.

Olsen, J., N. J. Anderson, and M. F. Knudsen, 2012: Variability of the North Atlantic Oscillation over the past 5,200 years. *Nat. Geosci.*, **5**, 808–812, https://doi.org/10.1038/ngeo1589.

Omrani, N.-E., N. S. Keenlyside, J. Bader, and E. Manzini, 2014: Stratosphere key for wintertime atmospheric response to warm Atlantic decadal conditions. *Climate Dyn.*, **42**, 649–663, https://doi.org/10.1007/s00382-013-1860-3.

——, J. Bader, N. S. Keenlyside, and E. Manzini, 2016: Troposphere–stratosphere response to large-scale North Atlantic Ocean variability in an atmosphere/ocean coupled model. *Climate Dyn.*, **46**, 1397–1415, https://doi.org/10.1007/s00382-015-2654-6.

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.

Ortega, P., F. Lehner, D. Swingedouw, V. Masson-Delmotte, C. C. Raible, M. Casado, and P. Yiou, 2015: A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature*, **523**, 71–74, https://doi.org/10.1038/nature14518.

Peings, Y., and G. Magnusdottir, 2014: Forcing of the wintertime atmospheric circulation by the multidecadal fluctuations of the North Atlantic Ocean. *Environ. Res. Lett.*, **9**, 034018, https://doi.org/10.1088/1748-9326/9/3/034018.

——, and ——, 2016: Wintertime atmospheric response to Atlantic multidecadal variability: Effect of stratospheric representation and ocean–atmosphere coupling. *Climate Dyn.*, **47**, 1029–1047, https://doi.org/10.1007/s00382-015-2887-4.

——, G. Simpkins, and G. Magnusdottir, 2016: Multidecadal fluctuations of the North Atlantic Ocean and feedback on the winter climate in CMIP5 control simulations. *J. Geophys. Res. Atmos.*, **121**, 2571–2592, https://doi.org/10.1002/2015JD024107.

Pohlmann, H., D. M. Smith, M. A. Balmaseda, N. S. Keenlyside, S. Masina, D. Matei, W. A. Müller, and P. Rogel, 2013: Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. *Climate Dyn.*, **41**, 775–785, https://doi.org/10.1007/s00382-013-1663-6.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 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.

Richter, J. H., C. Deser, and L. Sun, 2015: Effects of stratospheric variability on El Niño teleconnections. *Environ. Res. Lett.*, **10**, 124021, https://doi.org/10.1088/1748-9326/10/12/124021.

Robson, J., R. Sutton, K. Lohmann, D. Smith, and M. D. Palmer, 2012: Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s. *J. Climate*, **25**, 4116–4134, https://doi.org/10.1175/JCLI-D-11-00443.1.

Ruprich-Robert, Y., and C. Cassou, 2015: Combined influences of seasonal east Atlantic pattern and North Atlantic Oscillation to excite Atlantic multidecadal variability in a climate model. *Climate Dyn.*, **44**, 229–253, https://doi.org/10.1007/s00382-014-2176-7.

——, 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.

Semenov, V. A., M. Latif, D. Dommenget, N. S. Keenlyside, A. Strehz, T. Martin, and W. Park, 2010: The impact of North Atlantic–Arctic multidecadal variability on Northern Hemisphere surface air temperature. *J. Climate*, **23**, 5668–5677, https://doi.org/10.1175/2010JCLI3347.1.

Smith, R. D., and Coauthors, 2010: The Parallel Ocean Program (POP) reference manual, ocean component of the Community Climate System Model (CCSM) and Community Earth System Model (CESM). Los Alamos National Laboratory Tech. Rep. LAUR-10-01853, 141 pp., http://www.cesm.ucar.edu/models/cesm1.0/pop2/doc/sci/POPRefManual.pdf.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296, https://doi.org/10.1175/2007JCLI2100.1.

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.

Swart, N., J. C. Fyfe, E. Hawkins, J. E. Kay, and A. Jahn, 2015: Influence of internal variability on Arctic sea–ice trends. *Nat. Climate Change*, **5**, 86–89, https://doi.org/10.1038/nclimate2483.

Swingedouw, D., P. Ortega, J. Mignot, E. Guilyardi, V. Masson-Delmotte, P. G. Butler, M. Khodri, and R. Séférian, 2015: Bidecadal North Atlantic Ocean circulation variability controlled by timing of volcanic eruptions. *Nat. Commun.*, **6**, 6545, https://doi.org/10.1038/ncomms7545.

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

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Ting, M., Y. Kushnir, R. Seager, and C. Li, 2011: Robust features of the Atlantic multi-decadal variability and its climate impacts. *Geophys. Res. Lett.*, **38**, L17705, https://doi.org/10.1029/2011GL048712.

Trenberth, K. E., 1984: Some effects of finite sample size and persistence on meteorological statistics. Part I: Autocorrelations. *Mon. Wea. Rev.*, **112**, 2359–2368, https://doi.org/10.1175/1520-0493(1984)112<2359:SEOFSS>2.0.CO;2.

Visbeck, M., E. P. Chassignet, R. G. Curry, T. L. Delworth, R. R. Dickson, and G. Krahmann, 2003: The ocean's response to North Atlantic Oscillation variability. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr.*, Vol. 134, Amer. Geophys. Union, 113–145.

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.

Weaver, A. J., and Coauthors, 2012: Stability of the Atlantic meridional overturning circulation: A model intercomparison. *Geophys. Res. Lett.*, **39**, L20709, https://doi.org/10.1029/2012GL053763.

Yashayaev, I., 2007: Hydrographic changes in the Labrador Sea, 1960–2005. *Prog. Oceanogr.*, **73**, 242–276, https://doi.org/10.1016/j.pocean.2007.04.015.

——, and J. W. Loder, 2016: Recurrent replenishment of Labrador Sea Water and associated decadal-scale variability. *J. Geophys. Res. Oceans*, **121**, 8095–8114, https://doi.org/10.1002/2016JC012046.

Yeager, S. G., and G. Danabasoglu, 2014: The origins of late-twentieth-century variations in the large-scale North Atlantic circulation. *J. Climate*, **27**, 3222–3247, https://doi.org/10.1175/JCLI-D-13-00125.1.

——, A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng, 2012: A decadal prediction case study: Late twentieth-century North Atlantic Ocean heat content. *J. Climate*, **25**, 5173–5189, https://doi.org/10.1175/JCLI-D-11-00595.1.

——, A. R. Karspeck, and G. Danabasoglu, 2015: Predicted slowdown in the rate of Atlantic sea ice loss. *Geophys. Res. Lett.*, **42**, 10 704–10 713, https://doi.org/10.1002/2015GL065364.

Zhang, J., and R. Zhang, 2015: On the evolution of Atlantic meridional overturning circulation fingerprint and implications for decadal predictability in the North Atlantic. *Geophys. Res. Lett.*, **42**, 5419–5426, https://doi.org/10.1002/2015GL064596.

Zhang, L., and C. Wang, 2013: Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations. *J. Geophys. Res. Oceans*, **118**, 5772–5791, https://doi.org/10.1002/jgrc.20390.

Zhang, R., 2015: Mechanisms for low-frequency variability of summer Arctic sea ice extent. *Proc. Natl. Acad. Sci. USA*, **112**, 4570–4575, https://doi.org/10.1073/pnas.1422296112.

——, 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.

Zhu, X., and J. Jungclaus, 2008: Interdecadal variability of the meridional overturning circulation as on ocean internal mode. *Climate Dyn.*, **31**, 731–741, https://doi.org/10.1007/s00382-008-0383-9.