Does External Forcing Interfere with the AMOC's Influence on North Atlantic Sea Surface Temperature?

NEIL F. TANDON AND PAUL J. KUSHNER

Department of Physics, University of Toronto, Toronto, Ontario, Canada

(Manuscript received 26 September 2014, in final form 27 February 2015)

ABSTRACT

Numerous studies have suggested that variations in the strength of the Atlantic meridional overturning circulation (AMOC) may drive predictable variations in North Atlantic sea surface temperature (NASST). However, two recent studies have presented results suggesting that coupled models disagree on both the sign and the phasing of the correlation between AMOC and NASST indices. These studies analyzed linearly detrended output from twentieth-century historical simulations in phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5). The present study argues that the apparent disagreement among models arises from a comingling of two processes: 1) a bottom-up effect in which unforced AMOC changes lead to NASST changes of the same sign and 2) a top-down effect in which forced NASST changes lead to AMOC changes of the opposite sign. Linear detrending is not appropriate for separating these two effects because the time scales of forced and unforced variations are not well separated. After forced variations are properly removed, the models come into much closer agreement with each other. This argument is supported by analysis of CMIP5 historical simulations, as well as preindustrial control simulations and a 29-member ensemble of the Community Earth System Model, version 1, covering the period 1920–2005. Additional analysis is presented suggesting that, even after the data are linearly detrended, a significant portion of observed NASST persistence may be externally forced.

1. Introduction

There is now great interest in predicting how climate will change on time scales of decades [see Meehl et al. (2009, 2014) for extensive reviews]. In this effort, the North Atlantic Ocean naturally assumes a prominent role. This is because sea surface temperature (SST) in the North Atlantic strongly feels the effects of the Atlantic meridional overturning circulation (AMOC), the strength of which may vary in predictable multidecadal cycles (Collins and Sinha 2003; Pohlmann et al. 2004; Msadek et al. 2010). Furthermore, these internally generated variations may act to amplify or cancel some of the warming due to increased greenhouse gases (e.g., Knight et al. 2005; Ting et al. 2009; Mann et al. 2014).

Herein lies a problem: observing systems for the AMOC have been in operation only since 2004 (Rayner et al. 2011), and attempts to reconstruct earlier AMOC variations have produced widely disparate results (Munoz et al. 2011). In light of this, various approaches have been taken to infer an influence of the ocean circulation on North Atlantic sea surface temperature (NASST). Kushnir (1994) found decades in the observed record during which the North Atlantic was anomalously warm. This warming had a basinwide structure that did not align with anomalies in the atmospheric circulation. Thus, Kushnir (1994) concluded that this warming must have been generated by the ocean circulation, in agreement with the hypothesis of Bjerknes (1964).

Delworth et al. (1993) demonstrated this AMOC–NASST connection in a coupled model integration. Specifically, they found that strengthening of the AMOC was associated with a 0.2°–0.5°C warming of NASST. Such a warming can arise because a stronger AMOC is associated with stronger poleward transport of heat in the upper ocean.

Numerous other modeling studies have shown an influence of the AMOC on NASST, but the associated time scales are heavily model dependent (Zhang 2008; Teng et al. 2011; Marini and Frankignoul 2014). More troubling, Medhaug and Furevik (2011) and Zhang and Wang (2013) have recently shown results suggesting that

Corresponding author address: Dr. Neil F. Tandon, Department of Physics, University of Toronto, 60 St. George St., Toronto, ON M5S 1A7, Canada.
E-mail: neil.tandon@utoronto.ca

DOI: 10.1175/JCLI-D-14-00664.1

models cannot agree on whether AMOC strength and NASST are positively correlated, or if AMOC variations lead or lag NASST variations. These studies analyzed linearly detrended historical simulations from models participating in phases 3 and 5 of the Coupled Model Intercomparison Project [CMIP3 and CMIP5; these intercomparisons are described in Meehl et al. (2007) and Taylor et al. (2012)]. These findings raise serious questions about the ability of models to produce realistic multidecadal variability. We address some of these questions in this study.

Specifically, we expand on the analysis of Zhang and Wang (2013) to include all available realizations from the CMIP5 historical simulations, as well as preindustrial control (PIC) simulations. We show that in the PIC simulations, all available models agree that AMOC and NASST anomalies are positively correlated, whereas this is not the case in the historical simulations. (Note that in the PIC simulations, year-to-year levels of prescribed ''external forcers''—such as greenhouse gases, aerosols and solar irradiance—are all held constant.) This suggests that variations due to prescribed external forcers, which we call ''forced variations,'' may be confounding the relationship between the AMOC and NASST. We discuss the physical mechanisms responsible for this confounding effect, and provide additional supporting evidence by analyzing a 29-member ensemble of integrations with the Community Earth System Model, version 1 (CESM1).

2. Method

We analyze output from coupled models participating in CMIP5. Data were obtained for historical simulations covering the period 1850–2005, as well as PIC simulations. Only the first 499 years of each PIC simulation were used; models with PIC simulations shorter than 499 years were not included in our analysis. Eight models had the SST and AMOC data required for our analysis: CanESM2, CCSM4, CNRM-CM5, INM-CM4.0, MPI-ESM-LR, MPI-ESM-MR, MPI-ESM-P, and MRI-CGCM3 (expansions of these model names are available online at http://www.ametsoc.org/PubsAcronymList). We have also obtained historical and PIC simulations of CESM1 directly from NCAR. In addition, we analyze output from the 29-member large ensemble (LE) of CESM1 covering the period 1920–2005. Each of these realizations is driven by the same exact external forcings, but with slightly different atmospheric initial conditions (Kay et al. 2015). (The ocean initial conditions are identical in all realizations, a matter we consider in more detail below.)

Much of our analysis requires computing correlation coefficients between indices of the AMOC and NASST. Unless otherwise noted, we define the NASST index to be the spatial average of SST over 0°–60°N, 7.5°–75°W, following Ting et al. (2009). This NASST index is also sometimes referred to as the Atlantic multidecadal oscillation (AMO) index (e.g., Knight et al. 2005; Sutton and Hodson 2005). As detailed below, we also test using an alternative midlatitude NASST index in which we average over just 40°–60°N.

Following Medhaug and Furevik (2011), we define the AMOC index to be the maximum of the meridional overturning streamfunction in the Atlantic basin north of 20°N. We have also tested defining the AMOC index as the maximum of the meridional overturning streamfunction at 30°N, and none of our conclusions was affected. We did not calculate the meridional streamfunction from the meridional velocity as in Zhang and Wang (2013). Instead we used the basin-averaged meridional streamfunction variable (''msftmyz'') provided directly by the modeling centers. This allowed for a much more efficient analysis, especially for the PIC simulations, for which three-dimensional fields would require much more storage space.

To determine statistical significance for our lag correlations, we first compute the effective degrees of freedom (EDOF) for each ensemble member or time chunk. [See Eq. (31) in Bretherton et al. (1999). This formulation is also used in Zhang and Wang (2013).] Several of our figures show lag correlations for multiple ensemble members or time chunks on the same axes. To maintain visual clarity in such plots, we show only the significance level associated with the lowest (most conservative) value of the EDOF among the ensemble members or time chunks. Furthermore, it will become clear below that the correlations for most models have clear structure with respect to lag. This provides an additional indication of statistical significance, since one would not obtain such structure when correlating, for example, two white noise processes.

For comparison purposes, we also present some results from the NOAA Extended Reconstructed SST, version 3b (ERSST; Smith et al. 2008). We have tested using the merged Hadley Centre and NOAA Optimal Interpolation dataset (Hurrell et al. 2008), and none of our conclusions was affected. As mentioned above, there is wide scatter among twentieth-century reconstructions and reanalyses of the AMOC (Munoz et al. 2011), so we do not attempt to incorporate these into our analysis.

We discard the first 10 years of each realization of each historical simulation prior to any computation. This helps ensure that our computations are not heavily influenced by initial conditions. Our choice of 10 years comes from Branstator et al. (2012) and Branstator and Teng (2014), who, based on a multimodel analysis, found that the AMOC and NASST take approximately 10 years to lose memory of their initial conditions.

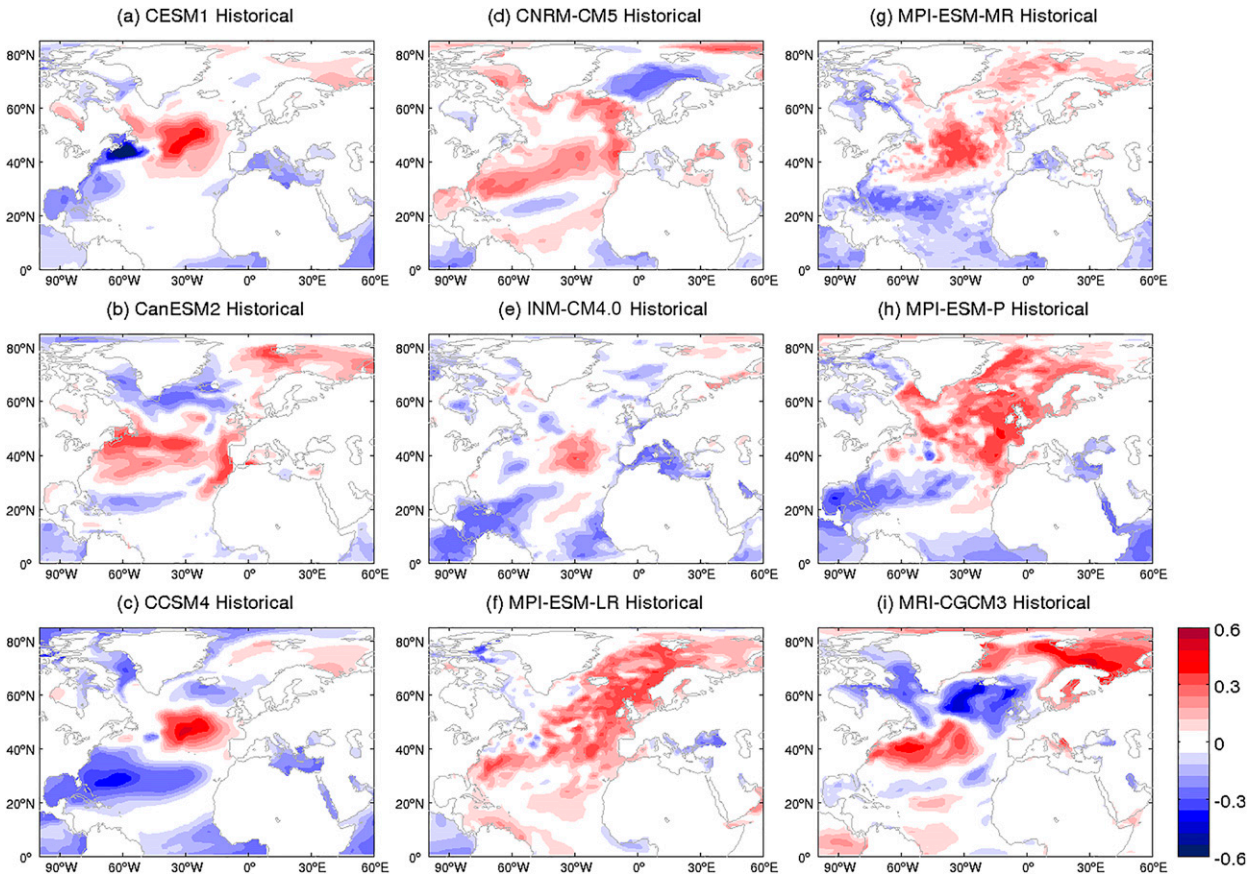

FIG. 1. Correlation between the detrended annual mean AMOC index and the detrended annual mean SST at each grid point for nine coupled model simulations with historical forcings. Each historical simulation covers the period 1850–2005, and we have discarded the first 10 years of each simulation prior to any computation. (See section 2 for details.) For models with multiple realizations, only the first realization is shown.

3. Analysis of CMIP5 historical simulations

As described in the introduction, the results of Zhang and Wang (2013) suggested that coupled models are highly inconsistent in how they represent the relationship between the AMOC and NASST. We have reproduced some of their findings in Figs. 1 and 2. Specifically, Fig. 1 shows the correlation coefficient between the detrended annual mean AMOC index and the detrended annual mean SST at each grid point in the North Atlantic. The data used for the calculation come from simulations of coupled models with historical forcings covering the period 1850–2005, as submitted to CMIP5. Figure 1 shows that, while all models have regions of positive correlation, they disagree on the location and extent of these regions. For example, CESM1 (Fig. 1a) shows peak positive correlation in the Labrador Sea, but MRI-CGCM3 (Fig. 1i) shows peak positive correlation farther south. Meanwhile, MPI-ESM-LR (Fig. 1f) shows positive correlation spread over most of the North Atlantic.

Figure 2 shows the lag correlations between the detrended AMOC and NASST indices for the same nine models. Here, lag is measured in years, and positive lag indicates that changes in the AMOC lead changes in NASST. Comparing the results for the first ensemble members of each model (thick black lines), the models disagree on the overall magnitude, sign, and phasing of the AMOC–NASST correlation. For example, CanESM2 (Fig. 2b), CNRM-CM5 (Fig. 2d), MPI-ESM-LR (Fig. 2f), MPI-ESM-MR (Fig. 2g), and MRI-CGCM3 (Fig. 2i) show significant positive correlation at nonnegative lag. But CCSM4 (Fig. 2c) shows negative correlation over most of the lag range. And if CESM1 (Fig. 2a), INM-CM4.0 (Fig. 2e), and MPI-ESM-P (Fig. 2h) show any significant correlation, it is at negative lag, when the AMOC lags NASST.

Based on these findings, one might suspect that coupled models are exhibiting fundamentally different physical behavior, as suggested by Medhaug and Furevik (2011) and Zhang and Wang (2013). However, it is also possible that internal or forced variability are confounding

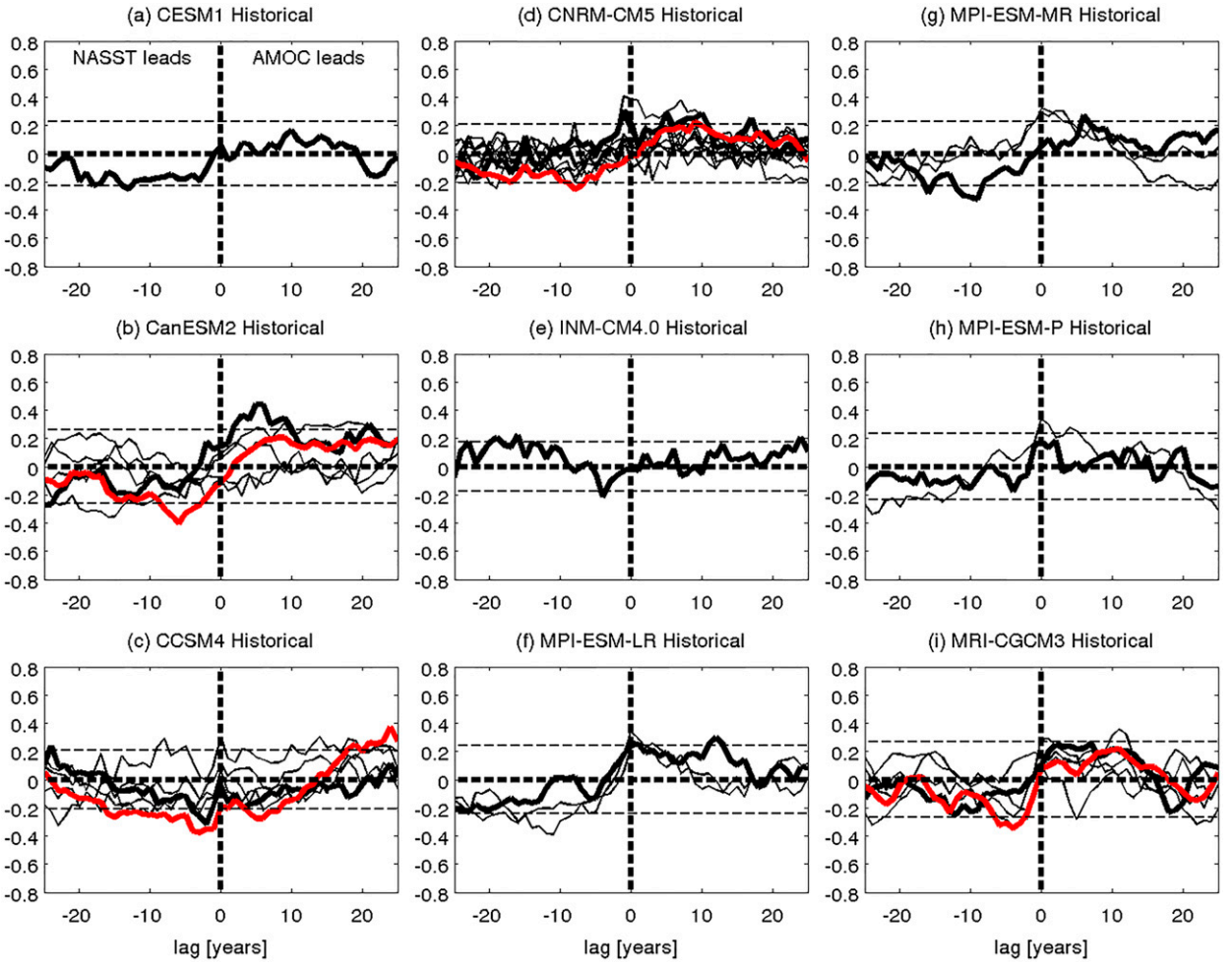

FIG. 2. Lag correlation between the detrended annual mean AMOC index and the detrended annual mean NASST index for nine coupled model simulations with historical forcings. The abscissa shows the lag (yr), with positive values indicating that AMOC changes lead NASST changes. For models with multiple realizations, the thick black curve indicates the lag correlation for the first realization. (This facilitates comparison with Fig. 1.) For models with five or more realizations, the red curve indicates the lag correlation between the ensemble means of the detrended AMOC and NASST indices. The thin dashed lines indicate statistical significance at the 95% level for the black curves. (This significance level is not applicable to the red curves. See section 2 for additional details.)

factors. One way to test for this is by considering all available realizations from the CMIP5 historical simulations. These additional realizations (thin black lines in Fig. 2) are subjected to the same external forcings as the first ensemble member, but their initial conditions differ. These show that different realizations produce very different AMOC–NASST correlations. While the first realization of CanESM2 produces significant positive correlation for positive lag, other realizations produce negligible—or even negative—correlation (Fig. 2b). While the first realization of CCSM4 shows negative correlation at lag 0, other realizations produce positive correlation (Fig. 2c). There is large intraensemble scatter in CNRM-CM5 (Fig. 2d) and MRI-CGCM3 (Fig. 2i) as well. The MPI models (Figs. 2f–h) show less scatter, but this may be because they include fewer (two or three) realizations. So the apparent discrepancies shown in Medhaug and Furevik (2011) and Zhang and Wang (2013) cannot be explained simply in terms of intermodel differences; even for a single model, choosing a different realization produces a very different AMOC–NASST correlation.

For the alternate realizations, the spatial structure of the simultaneous AMOC–SST correlations is very different from that shown in Fig. 1. For example, realization 3 of CCSM4 produces positive correlation over most of the North Atlantic with some negative correlation off the coast of Florida (not shown). Realization 6 of CNRM-CM5 shows negative correlation over most of the North Atlantic, with some positive correlation off

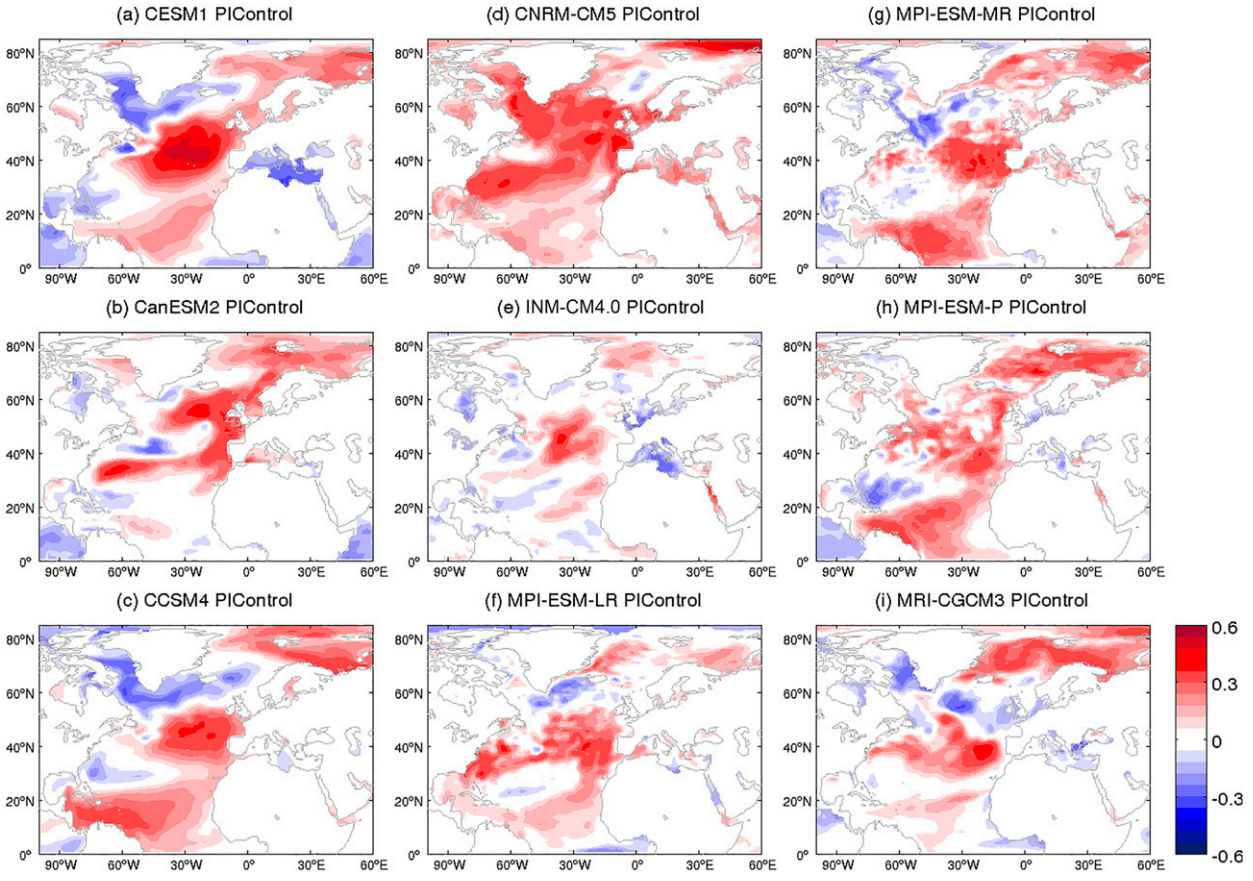

FIG. 3. As in Fig. 1, but for the first 146 years of nine PIC simulations.

the coast of Portugal (not shown). So the large intraensemble scatter in the lag correlations of Fig. 2 is reflective of the large intraensemble scatter in the overall spatial pattern of AMOC–SST correlation.

4. Analysis of CMIP5 preindustrial control simulations

This raises a pivotal question: Is there any robust AMOC–NASST relationship in coupled models, or is it all washed out by internal variability? To consider this, we performed the same analysis on PIC simulations for the same nine models. In these PIC simulations, there is no time-varying external forcing beyond the seasonal cycle. Each PIC time series is 499 years long, whereas the historical simulations analyzed above were each 146 years long. (As detailed in section 2 we discard the first 10 years of each historical simulation.) Thus, to more directly compare with the historical simulations, we select 146-yr chunks from the PIC simulations and calculate correlations after detrending each chunk. We have tested an alternative approach in which we detrend the entire PIC time series before selecting the time chunks, and perform no additional detrending thereafter. This alternative approach produces very similar results, and the most noticeable quantitative difference is in CNRM-CM5, which exhibits large oscillations on centennial time scales.

Figure 3 shows the correlation between the AMOC index and SST at each grid point for the first 146 years of each PIC simulation. Compared to the historical simulations (Fig. 1), the PIC simulations show much more agreement regarding the spatial structure of the AMOC–SST correlation. Specifically, all models show a tripole pattern resembling that found in earlier studies (Delworth et al. 1993; Msadek and Frankignoul 2009; Teng et al. 2011; Marini and Frankignoul 2014), although the precise location and shape of the tripole pattern varies from model to model. Similar results were obtained by computing the correlation for the full 499-yr time series (not shown).

Figure 4 shows the AMOC–NASST lag correlations computed from the PIC simulations. Each black curve corresponds to a correlation computed from a 146-yr chunk of the PIC simulation. We have selected 10 such chunks, which means that the chunks are not

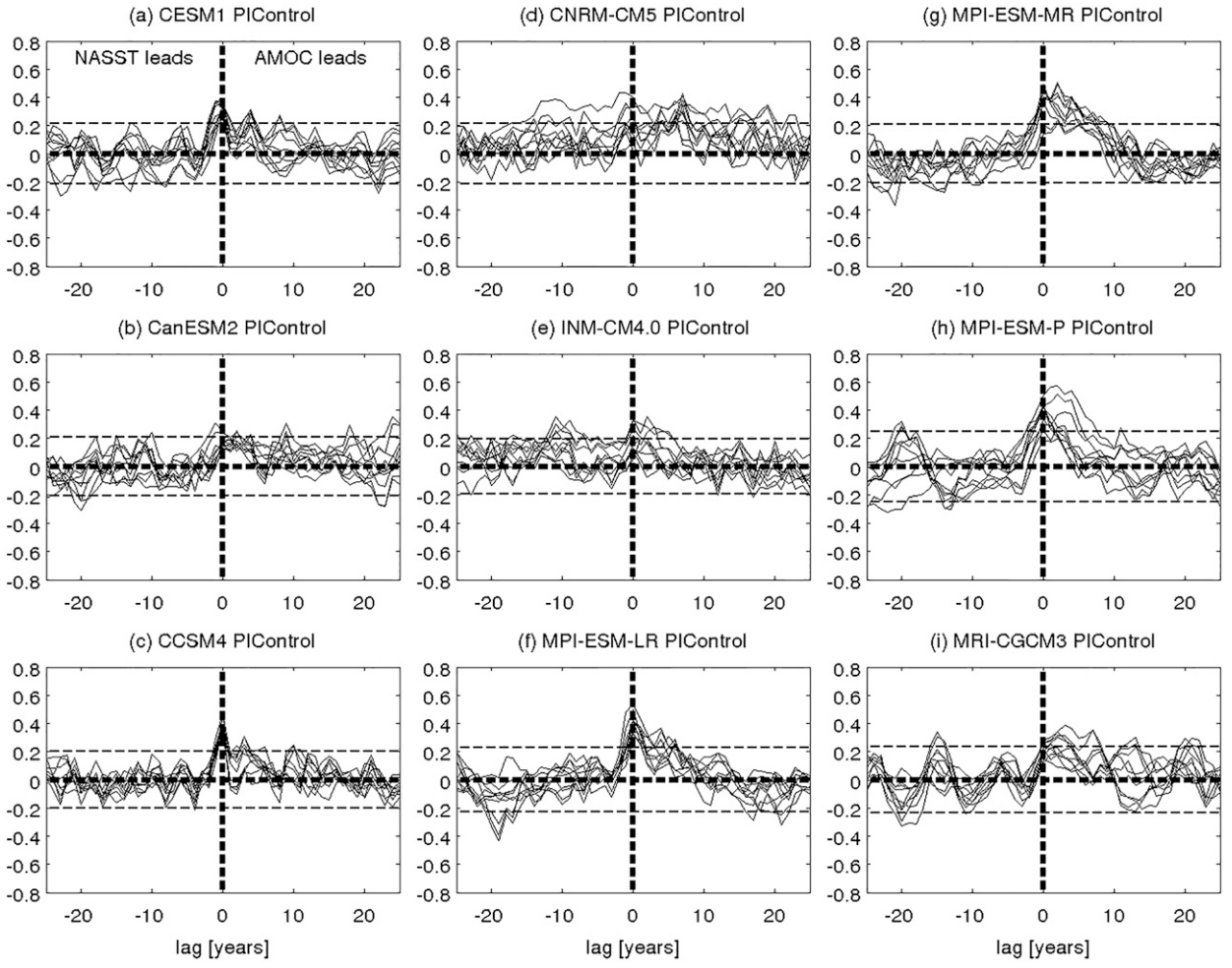

FIG. 4. Lag correlation between the detrended annual mean AMOC index and the detrended annual mean NASST index for nine PIC simulations. The abscissa shows the lag (yr), with positive values indicating that AMOC changes lead NASST changes. Each black curve indicates the lag correlation computed from a continuous 146-yr chunk selected from a total of 499 years of data. For each model, 10 such chunks were selected (i.e., the chunks overlap each other), and each chunk was linearly detrended before computing the lag correlation. The thin dashed lines indicate statistical significance at the 95% level. (See section 2 for additional details.)

independent of each other, but this still provides some preliminary comparison with the historical simulations. Figure 4 shows that in the absence of forced variability, all chunks of all the PIC simulations produce positive simultaneous correlations between the AMOC and NASST indices. Thus, the correlations computed from the PIC simulations show an AMOC–NASST relationship in much closer agreement with earlier findings (Delworth et al. 1993; Vellinga and Wu 2004; Knight et al. 2005; Teng et al. 2011; Ba et al. 2014) and in accord with the idea that a strengthening of the AMOC is associated with stronger poleward heat transport in the upper ocean. Marini and Frankignoul (2014) have also shown that some models produce a higher AMOC–NASST correlation in control simulations compared to historical simulations.

Recall that in the historical simulations, several models produced at least one realization in which the peak AMOC–NASST correlation occurs at positive lags of more than 5 years (Figs. 2b,c,d,f,g,i). This would suggest that the AMOC is providing strong predictive power for NASST anomalies. However, such strong predictive power is mostly absent in the PIC simulations, except perhaps in the MPI and MRI models (Figs. 4f–i). Thus, forced variations affect not just the simultaneous AMOC–NASST correlation but the lagged correlation as well.

Although the models all agree on the sign of the correlation near lag 0, they disagree on the strength and characteristic time scale of the lag correlation. For example, MPI-ESM-P (Fig. 4h) shows significant positive correlation at lag 0, with evidence of a coherent 20-yr

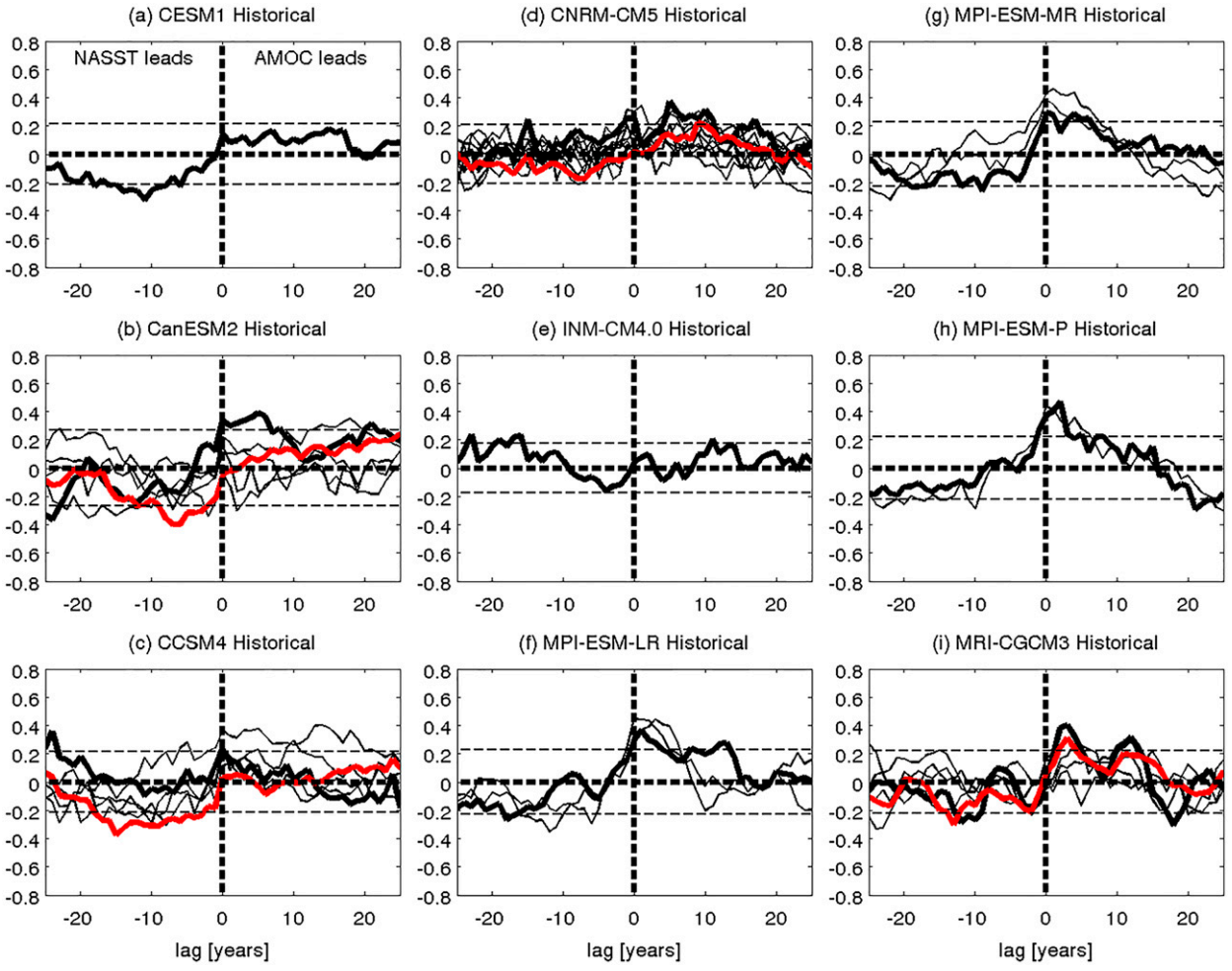

FIG. 5. As in Fig. 2, but for the modified midlatitude NASST index, which is defined as the spatial average of NASST between 40° and 60°N.

oscillation. Meanwhile, CanESM2 (Fig. 4b) shows some positive correlation at lag 0 but no coherent oscillation. That said, the scatter among the chunked lag correlations of the PIC simulations is much less than the intraensemble scatter in the historical simulations.

As mentioned above, the overall spatial structure of the AMOC–SST correlation in the PIC simulations is very different from that of the historical simulations (cf. Figs. 1 and 3). However, in the midlatitude region between 40° and 60°N, all of the historical and PIC simulations produce mostly positive AMOC–NASST correlation. Figure 5 shows that, if we confine our NASST index to the 40°–60°N band, then the simultaneous AMOC–NASST correlation is more robust and positive for all the CMIP5 historical simulations. This suggests that forced variations in the tropics might account for some of the scatter in the lag correlation computation shown in Fig. 2. However, even when this midlatitude NASST index is adopted for computing lag correlations from the PIC simulations (not shown, but similar to Fig. 4), important differences remain between the PIC and historical simulations in terms of the sign and phasing of the AMOC–NASST correlation. This suggests that, even with a heuristically applied midlatitude NASST index, external forcing still interferes with the AMOC–NASST relationship.

5. Physical interpretation

Precisely how is external forcing interfering with the AMOC–NASST relationship? One possibility is that externally forced changes in NASST simply mask any NASST changes due to the AMOC. This would suggest that forced NASST changes are decoupled from the AMOC. However, numerous earlier studies have shown that this is not the case. A positive thermal forcing from

increased CO_2, for example, is expected to weaken the AMOC through changes in stratification and salinity (e.g., Mikolajewicz and Voss 2000; Gregory et al. 2005; Weaver et al. 2007). So AMOC strength is coupled to external forcing, and the effect of external forcing on the AMOC–NASST relationship cannot be viewed as a simple masking effect.

Instead, when analyzing the AMOC–NASST relationship in the presence of forced variations, there are actually two different outcomes we might expect. First, we might expect an unforced bottom-up effect in which AMOC strengthening (weakening) produces NASST warming (cooling). Second, we might expect a forced top-down effect in which NASST warming (cooling) produces AMOC weakening (strengthening). In other words, we would expect that unforced variability produces a positive AMOC–NASST correlation with AMOC variations leading NASST variations, whereas we would expect forced variability to produce a negative AMOC–NASST correlation with NASST variations leading AMOC variations.

The correlations computed from the CMIP5 historical simulations show characteristics of both the top-down and bottom-up effects. Most of the models produce at least one realization in which there is significant negative correlation at negative lag, reflecting the forced top-down effect (Fig. 2). Seven of the models produce at least one realization with significant positive correlation at nonnegative lag, reflecting an unforced bottom-up effect. This interpretation is further supported by the PIC simulations, in which all models show significant positive correlation at nonnegative lag, indicating that only the unforced bottom-up effect is at work (Fig. 4).

Over time, we would expect negative feedbacks associated with both the top-down and bottom-up effects: a weaker (stronger) AMOC would produce cooler (warmer) NASST, weaker (stronger) ocean stratification, and eventually a stronger (weaker) AMOC. Numerous studies going back to Delworth et al. (1993) have shown evidence for such a feedback, though there are additional details of the feedback mechanism (i.e., involving changes in salinity and air–sea fluxes; Ba et al. 2014) that are beyond our focus here. Thus, we would expect forced and unforced variations to produce different AMOC–NASST correlations not only near lag 0, but also for the longer lead–lag times associated with these feedbacks. This might help to explain the apparently inflated lag correlations shown above in some of the CMIP5 historical simulations.

The strength of our conclusions here is limited somewhat by the sample size. Only one model (CNRM-CM5) includes more than five historical realizations. So it is desirable to make a sharper distinction between the effects of forced and unforced variations, and we do this next by analyzing the CESM1 LE.

6. Analysis of the CESM1 large ensemble

To further investigate the effects of internal and forced variability, we turn to the CESM1 LE, for which 29 integrations were performed from 1920 to 2005 using the exact same external forcings but slightly different initial conditions (Kay et al. 2015). When we take an average of all these ensemble members, most of the internal variability averages out and we obtain an estimate of the forced variability. We can then consider the effect of including or removing this forced variability. Similar approaches have been taken in earlier studies (e.g., Kravtsov and Spannagle 2008; Knight 2009; DelSole et al. 2013; Mann et al. 2014; Schmith et al. 2014).

Time series for the NASST and AMOC indices are shown in Fig. 6. The ensemble mean NASST (Fig. 6a, red curve) is mostly flat between 1920 and 1960, dips abruptly around 1965, and trends positive thereafter. The ensemble mean is qualitatively similar to the multimodel mean computed from the CMIP5 historical simulations (pink curve), but with a difference of 0.3°–0.4°C. Mann et al. (2014) obtained qualitatively very similar estimates of forced changes in Northern Hemisphere (NH) mean surface temperature based on a large ensemble of integrations with GISS-E2-R, as well as an average of 40 coupled models from CMIP5. The NASST index computed from ERSST data (blue curve) does not depart far beyond the ensemble spread of CESM1. [Mann et al. (2014) show very similar results with different observational products.] The ensemble mean of the AMOC index (Fig. 6b, red curve) decreases from 1920 to 1950, followed by a positive trend until 1980. The CMIP5 multimodel mean exhibits qualitatively similar behavior, but is about 10 Sv ($1\,\text{Sv} \equiv 10^6\,\text{m}^3\,\text{s}^{-1}$) weaker than the CESM1 LE average.

Surprisingly, almost nothing changes qualitatively if we subtract the linear fit from each time series (Figs. 6c,d). This reflects the fact that there is very little long-term trend in the original time series, and that nearly all of the forced variability occurs on time scales of 10–30 yr. This means that linearly detrending is not suitable for removing forced variability, and thus lag correlations computed from detrended time series may still reflect an influence of external forcing. The detrended time series also show more clearly that forced variations from the CESM1 LE (red curves) are of comparable amplitude to those obtained from CMIP5 (pink curves).

Since linear detrending does not properly remove forced variability, we take the alternative approach of subtracting from each realization the ensemble mean of the CESM1 LE (Figs. 6e,f). That is, to obtain each black

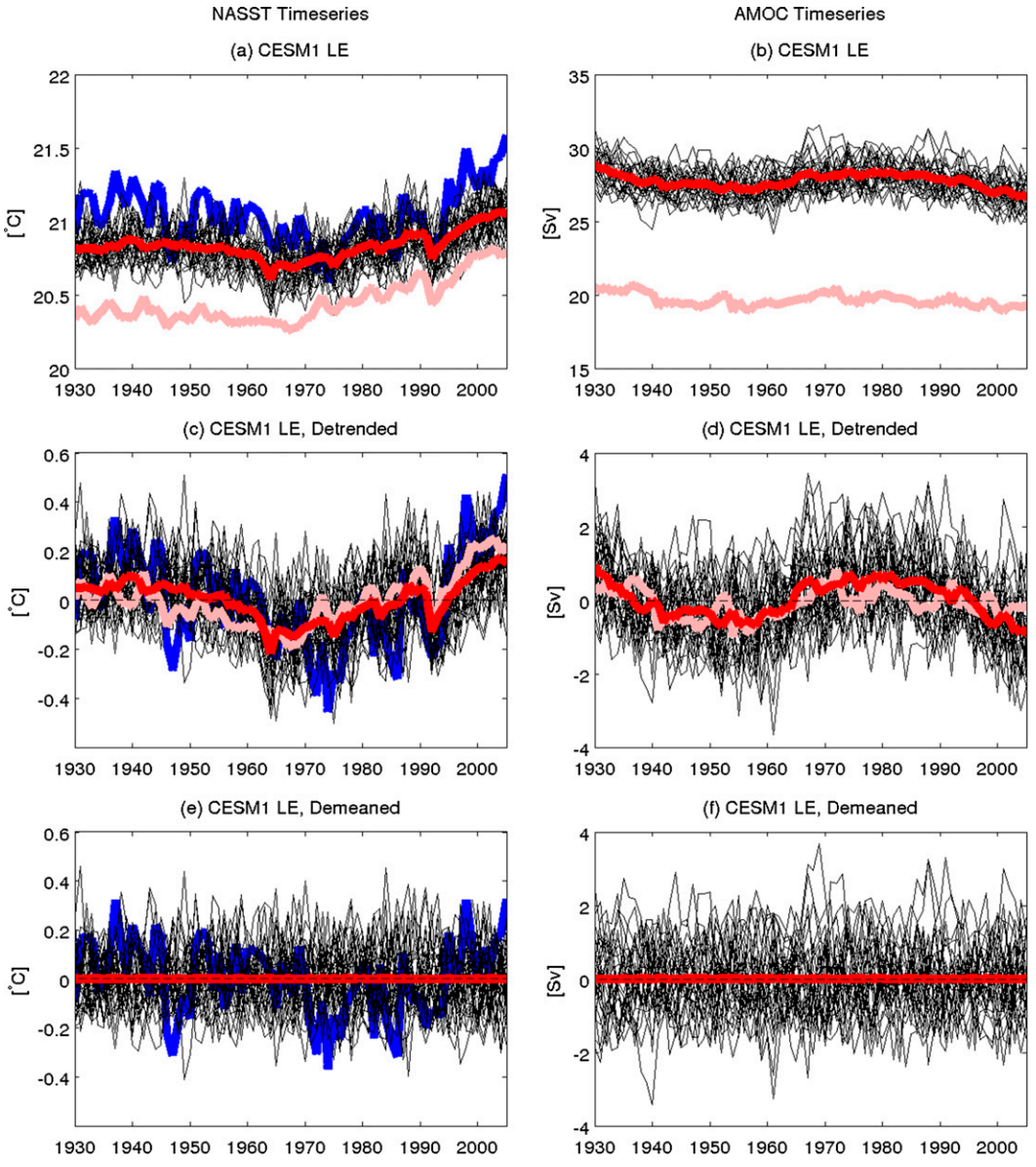

FIG. 6. Time series of the (a) NASST and (b) AMOC indices for the CESM1 29-member large ensemble. Each black curve corresponds to a single realization. The red curves show the ensemble means (i.e., the average of the black curves). For comparison, the pink curves show the multimodel means, computed by taking the average of the first ensemble members of the nine CMIP5 models analyzed in section 3. The blue curve shows the NASST time series from ERSST data. (c),(d) The same time series after subtracting the linear fit of each time series. (e),(f) The same time series after subtracting the CESM1 ensemble mean from each ensemble member as well as the ERSST data. In (e), the ERSST data have been recentered so that their time mean equals zero.

curve in Fig. 6e, we take the difference between each black curve and the red curve in Fig. 6a, and to obtain each black curve in Fig. 6f, we take the difference between each black curve and the red curve in Fig. 6b. We refer to this process as "demeaning" the data. After doing this, the average of all the black curves is, by construction, completely flat, reflecting the fact that all the forced variations have been removed. Furthermore, if we assume that the CESM1 forced response is the same as nature's forced response, we can consider the effect of removing this forced response from observations. After doing this (Fig. 6e, blue curve), we find that there is still multidecadal variability in observed NASST, but its amplitude is lower than in the detrended time series.

We now compare the AMOC–NASST correlations computed from detrended data with those computed

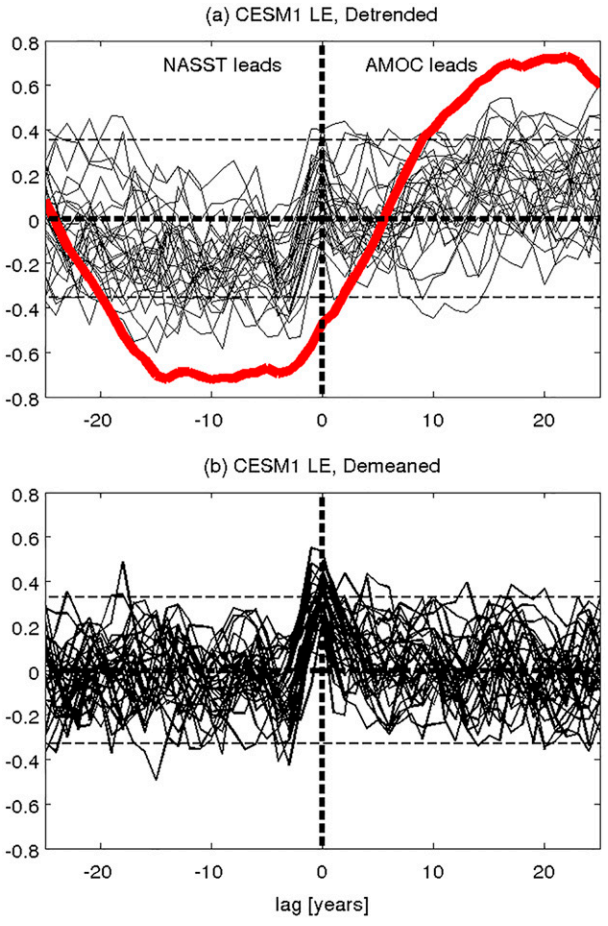

FIG. 7. (a) Lag correlation between the detrended annual mean AMOC index and detrended annual mean NASST index for 29 realizations of CESM1. (b) The same computation after subtracting the ensemble mean from each realization. The red curve in (a) is the lag correlation between the ensemble means of the detrended AMOC and NASST indices. The thin dashed lines indicate the 95% significance level for the CESM1 realizations. (Note that this significance level is not applicable to the red curve. See section 2 for additional details.)

from demeaned data (Fig. 7). The lag correlations on detrended data (Fig. 7a), have a large spread; at lag 0, some realizations show a negative correlation, while others show a positive correlation. However, if the ensemble mean is first removed (Fig. 7b), then all realizations produce a positive simultaneous correlation. Demeaning the data also affects the correlation at lead–lag times exceeding 10 yr: the detrended correlations show that for most realizations, the AMOC and NASST are positively (negatively) correlated at positive (negative) lag, whereas demeaned data show no preferred sign at positive (negative) lag. This may indicate, as suggested in section 5, that forced variations and their associated feedbacks are acting over a long enough time scale that they affect the AMOC–NASST relationship on multidecadal time scales.

Notice that the correlation with the demeaned data shows a strong resemblance to the correlation computed from PIC data (Fig. 4a). This provides additional assurance that our approach for removing forced variations is appropriate for this model. This also demonstrates that the effects of external forcing and internal variability are linearly separable, provided the separation is performed on all fields prior to computing correlations. Furthermore, this suggests that varying the ocean initial conditions for the CESM1 LE realizations would likely not affect any of our conclusions, because the PIC time chunks sample different ocean states. (Recall from earlier that only the atmospheric initial conditions are varied in the LE realizations.)

Figure 8a shows the lag 0 correlation between the AMOC and NASST for each CESM1 realization. This illustrates more clearly a key point of Fig. 7: detrended data (red bars) produce some correlations with negative sign, whereas demeaned data (blue bars) produce correlations that are all positive. Furthermore, the intraensemble spread of the correlation coefficients (error bars) is larger for the detrended data than for the demeaned data. Thus, properly removing forced variability reduces the intraensemble spread in the correlation calculation. Note that there is still substantial spread in the correlation coefficients in the demeaned output, so internal variability still plays a role. But forced variability can increase the scatter enough that the AMOC–NASST correlation actually changes sign, and linearly detrending the data does not resolve this problem.

Figure 8b shows the correlation coefficients for each realization when AMOC variations lead NASST variations by 20 yr. This reiterates another key point of Fig. 7: for most realizations, removing the forced component reduces the lag 20 correlation. In other words, forced variations appear to inflate the predictive power of the AMOC in CESM1, and when forced variations are removed, the AMOC provides little predictive power for NASST changes.

We can also consider, as we did earlier, the effect of removing the tropical contribution to NASST variations by adopting the same midlatitude NASST index that we used in Fig. 5. In agreement with the results of Fig. 5, most realizations show positive simultaneous correlation (only three realizations do not), but external forcing still exerts substantial influence on the lagged correlations (not shown).

7. Changes in persistence

One of the reasons we are so interested in the AMOC–NASST relationship is because of the possibility of predicting NASST changes years in advance based on changes

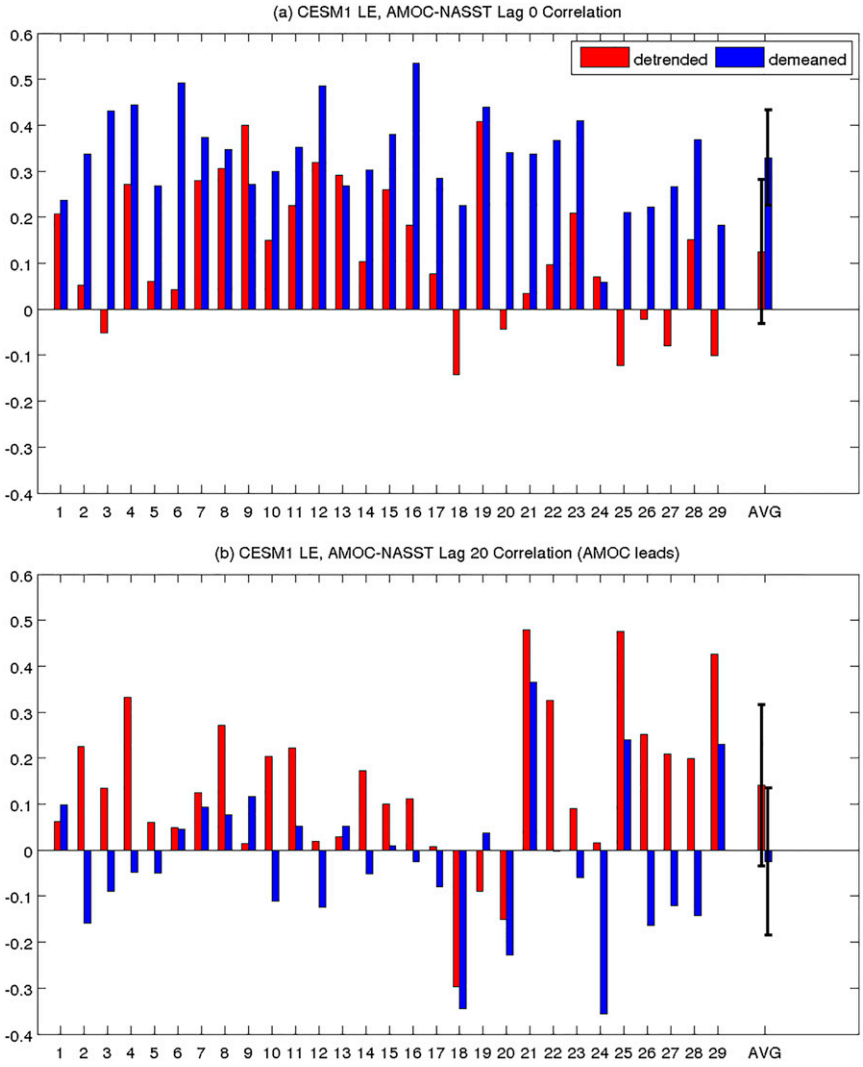

FIG. 8. (a) Lag 0 correlation between the annual mean NASST index and annual mean AMOC index after subtracting the linear trend (red) and after subtracting the ensemble mean (blue). Values are shown for each ensemble member of the CESM1 LE, as well as the average ("AVG") and standard deviation (error bars) of the correlation coefficients. (b) As in (a), but when the AMOC index leads the NASST index by 20 yr.

in AMOC strength. One way to characterize the predictability of NASST is by measuring its persistence, or the degree to which the field is autocorrelated. DelSole et al. (2013) have shown that NASST persistence is about the same as its predictability based on a more sophisticated linear regression model derived from a multimodel dataset. Indeed, NASST persistence is often thought to be due to the ocean circulation, so our earlier results raise questions about a possible influence of external forcing.

Figure 9 shows the lag 1 autocorrelation of the annual mean NASST and AMOC indices for each of the CESM1 realizations. The NASST autocorrelation is about 0.5 when computed from detrended data (red bars). For all of the realizations, removing the ensemble mean leads to a reduction in NASST persistence (blue bars). Averaged across all ensemble members, this translates to a 0.15 reduction in the autocorrelation coefficient. The autocorrelation computed from observed data falls within the range of autocorrelations computed from CESM1, although it is in the tail of the distribution: only one CESM1 realization produces more persistence than observed. Subtracting the CESM1 ensemble mean from observations reduces the autocorrelation to 0.43, which is within one standard deviation of the CESM1 LE average. Taken together, these results suggest that, if our estimate of forced variability is accurate, then much of the apparent persistence of NASST is due to forced

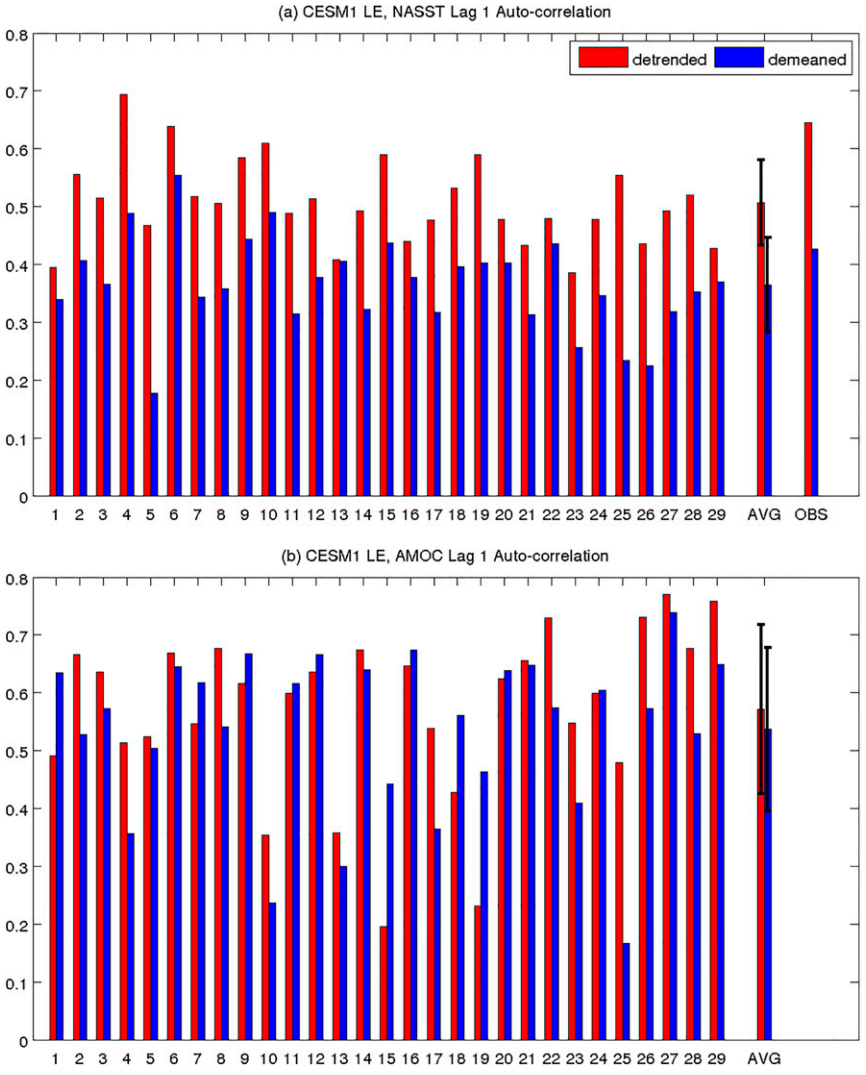

FIG. 9. Lag 1 autocorrelation of the annual mean (a) NASST index and (b) AMOC index after subtracting the linear trend (red) and after subtracting the ensemble mean (blue). Values are shown for each ensemble member of the CESM1 LE, as well as the average (''AVG'') and standard deviation (error bars) of the autocorrelation coefficients. The NASST autocorrelation coefficients are also computed from ERSST data (''OBS''), for which the demeaned value is obtained by first subtracting the ensemble mean of the CESM1 LE.

variations, and removing forced variations greatly reduces NASST persistence.

Based on this, one might expect that removing forced variations reduces AMOC persistence as well, but this is not so apparent (Fig. 9b). Specifically, demeaning the AMOC index leads to a reduction in persistence for most ensemble members, but for 11 realizations this actually increases persistence. Averaged across all ensemble members, there is a slight reduction in AMOC persistence, but the effect is much smaller than that for NASST. This may reflect the fact that external forcing can directly affect NASST through changes in surface fluxes, whereas the effect on AMOC strength is more indirect, requiring more persistent forcing over longer periods in order to detect a signal (cf. Gregory et al. 2005; Weaver et al. 2007).

8. Discussion and conclusions

Overall, the results of our CESM1 LE analysis support our interpretation of the CMIP5 simulation results. Specifically, external forcing interferes with the typical unforced relationship between the AMOC and NASST, and this can lead to ambiguity in the sign and phasing of the AMOC–NASST correlation. Forced variations can produce a negative simultaneous correlation between

the AMOC and NASST, and forced variations can also inflate the correlation when NASST changes lead and lag AMOC changes.

In section 5, we provided some ideas on the physical mechanisms that might explain this effect of external forcing. We argued that there is an unforced bottom-up effect whereby AMOC changes lead to NASST changes of the same sign, and there is a forced top-down effect whereby NASST changes lead to AMOC changes of the opposite sign. The results of CESM1 LE show that correlations performed on detrended data exhibit characteristics of both the bottom-up and top-down effects: some realizations produce positive correlation at lag 0 and positive lag, while other realizations produce negative correlation at lag 0 and negative lag (Fig. 7). Furthermore, in CESM1 the forced top-down effect appears to act over a longer time scale than the unforced bottom-up effect, leading to the false impression in most realizations that AMOC changes are providing predictive power for NASST changes more than 10 years in advance.

To better visualize the effect of external forcing on the AMOC–NASST correlation, we also show the correlation between the ensemble averages of the AMOC and NASST indices (Fig. 7, red curve). This shows what the AMOC–NASST correlation might be in the absence of internal variability. In agreement with earlier interpretations, this forced correlation is strongly negative with NASST variations leading AMOC variations, clearly reflecting the forced top-down effect. Although the CMIP5 historical simulations do not include as many realizations, four of the models include five or more realizations, which may be enough for some preliminary insight. For these four models, we have also plotted the correlation between the ensemble means, and they also suggest that forced variations on their own produce a clear top-down effect, with NASST variations leading AMOC variations of the opposite sign (Fig. 2). This is apparent even after the tropical contribution to NASST is removed (Fig. 5), except perhaps in the case of MRI-CGCM3 (Fig. 5i).

Deeper understanding of these forced correlations would also require explaining the time evolution of the NASST and AMOC forced responses. Recall that in the CESM1 LE, NASST shows very little forced long-term trend over the period 1920–2005 (Fig. 6a, red curve). This makes it unlikely that well-mixed greenhouse gases are entirely responsible for the NASST forced response, since these have nonnegative trends during the entire period. In contrast, the trend of NH aerosol emissions does change sign during this period. Mann et al. (2014) argued that there is an influence of anthropogenic SO_2 emissions on decadal NH temperature trends. Based on a multimodel analysis of single forcing simulations Terray (2012) also showed evidence for a role of aerosol forcing during this period, but his dataset was not large enough to cleanly separate the effects of internal and forced contributions. Based on our physical reasoning in section 5, it is also possible that some of the forced changes in NASST are due to a feedback from forced changes in the AMOC. So more work is needed to properly attribute the causes of past NASST variations.

Our results connect with the recent debate about the extent to which historical NASST trends are externally forced or internally generated. Several studies argue that past variations are strongly influenced by internal variability (Kravtsov and Spannagle 2008; Ting et al. 2009; Knight 2009; Zhang and Wang 2013; Schmith et al. 2014), although they disagree on the timing of internal and forced contributions (cf. Terray 2012; Mann et al. 2014). Booth et al. (2012) have argued that past variations may have been primarily driven by aerosols, although Zhang et al. (2013) and Zhang and Wang (2013) have shown strong evidence that the behavior of the model used by Booth et al. (2012) is unrealistic. Overall, our conclusions most closely echo those of Otterå et al. (2010) and Mann et al. (2014) that external forcing heavily influences the phasing of—but is not the main driver of—NASST multidecadal variability and that any attempt to attribute past changes in NASST must take multiple sources into account (Terray 2012). Our study takes this a step further and shows that external forcing can also strongly influence the apparent relationship between the AMOC and NASST, so much so that the sign and phasing of their correlation become ambiguous. Thus any attempt to understand the AMOC's influence on multidecadal climate variability must take this forcing effect into account. This is especially crucial for studies inferring AMOC variations through proxies that are possibly sensitive to external forcing (e.g., Knight et al. 2005; Lippold et al. 2009; Ritz et al. 2013).

Our results also suggest that decadal predictability of NASST may be severely limited in the absence of forced variability. However, this result may be model dependent, since studies with other models have shown evidence of unforced NASST variations that are decadally predictable (Griffies and Bryan 1997; Wu and Liu 2005; Zhang 2008; Yang et al. 2013). In addition, the lag correlations from control simulations of the MPI models (Figs. 4f–h) showed evidence of multidecadal covariation between the AMOC and NASST. More work is needed with other models—including more idealized models (e.g., Farneti and Vallis 2011)—to develop a clearer picture of what determines the existence of predictable modes of multidecadal variability and whether climate truly exhibits such modes.

Acknowledgments. This work was funded by the BNP-Paribas Foundation under the PRECLIDE project. We thank Laurent Terray for overseeing PRECLIDE and supporting our work. We also thank Frédéric Laliberté for developing the cdb_query tool, which greatly facilitated the acquisition of CMIP5 model output. We acknowledge the modeling centers for contributing to CMIP5, and we acknowledge the NCAR Climate Data Guide, which helped us select suitable observational datasets. Finally, we thank Bill Merryfield and Gavin Schmidt for helpful discussions and three anonymous reviewers for constructive feedback.

REFERENCES

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

Bjerknes, J., 1964: Atlantic air–sea interaction. Advances in Geophysics, Vol. 10, Academic Press, 1–82.

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

Branstator, G., and H. Teng, 2014: Is AMOC more predictable than North Atlantic heat content? *J. Climate,* **27,** 3537–3550, doi:10.1175/JCLI-D-13-00274.1.

——, ——, G. A. Meehl, M. Kimoto, J. R. Knight, M. Latif, and A. Rosati, 2012: Systematic estimates of initial-value decadal predictability for six AOGCMs. *J. Climate,* **25,** 1827–1846, doi:10.1175/JCLI-D-11-00227.1.

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, doi:10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2.

Collins, M., and B. Sinha, 2003: Predictability of decadal variations in the thermohaline circulation and climate. *Geophys. Res. Lett.,* **30,** 1306, doi:10.1029/2002GL016504.

DelSole, T., L. Jia, and M. K. Tippett, 2013: Decadal prediction of observed and simulated sea surface temperatures. *Geophys. Res. Lett.,* **40,** 2773–2778, doi:10.1002/grl.50185.

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

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

Gregory, J. M., and Coauthors, 2005: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO_2 concentration. *Geophys. Res. Lett.,* **32,** L12703, doi:10.1029/2005GL023209.

Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate variability. *Science,* **275,** 181–184, doi:10.1126/science.275.5297.181.

Hurrell, J. W., 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, doi:10.1175/2008JCLI2292.1.

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.,* doi:10.1175/BAMS-D-13-00255.1, in press.

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

——, 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, doi:10.1029/2005GL024233.

Kravtsov, S., and C. Spannagle, 2008: Multidecadal climate variability in observed and modeled surface temperatures. *J. Climate,* **21,** 1104–1121, doi:10.1175/2007JCLI1874.1.

Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate,* **7,** 141–157, doi:10.1175/1520-0442(1994)007<0141:IVINAS>2.0.CO;2.

Lippold, J., J. Grützner, D. Winter, Y. Lahaye, A. Mangini, and M. Christl, 2009: Does sedimentary $^{231}Pa/^{230}Th$ from the Bermuda rise monitor past Atlantic meridional overturning circulation? *Geophys. Res. Lett.,* **36,** L12601, doi:10.1029/2009GL038068.

Mann, M. E., B. A. Steinman, and S. K. Miller, 2014: On forced temperature changes, internal variability, and the AMO. *Geophys. Res. Lett.,* **41,** 3211–3219, doi:10.1002/2014GL059233.

Marini, C., and C. Frankignoul, 2014: An attempt to deconstruct the Atlantic multidecadal oscillation. *Climate Dyn.,* **43,** 607–625, doi:10.1007/s00382-013-1852-3.

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, doi:10.5194/os-7-389-2011.

Meehl, G. A., C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. B. Mitchell, 2007: THE WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394, doi:10.1175/BAMS-88-9-1383.

——, and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc.,* **90,** 1467–1485, doi:10.1175/2009BAMS2778.1.

——, and Coauthors, 2014: Decadal climate prediction: An update from the trenches. *Bull. Amer. Meteor. Soc.,* **95,** 243–267, doi:10.1175/BAMS-D-12-00241.1.

Mikolajewicz, U., and R. Voss, 2000: The role of the individual air–sea flux components in CO_2-induced changes of the ocean's circulation and climate. *Climate Dyn.,* **16,** 627–642, doi:10.1007/s003820000066.

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, doi:10.1007/s00382-008-0452-0.

——, K. W. Dixon, T. L. Delworth, and W. Hurlin, 2010: Assessing the predictability of the Atlantic meridional overturning circulation and associated fingerprints. *Geophys. Res. Lett.,* **37,** L19608, doi:10.1029/2010GL044517.

Munoz, E., B. Kirtman, and W. Weijer, 2011: Varied representation of the Atlantic meridional overturning across multidecadal ocean reanalyses. *Deep-Sea Res.,* **58,** 1848–1857.

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

Pohlmann, H., M. Botzet, M. Latif, A. Roesch, M. Wild, and P. Tschuck, 2004: Estimating the decadal predictability of a

coupled AOGCM. *J. Climate,* **17,** 4463–4472, doi:10.1175/3209.1.

Rayner, D., and Coauthors, 2011: Monitoring the Atlantic meridional overturning circulation. *Deep-Sea Res.,* **58,** 1744–1753, doi:10.1016/j.dsr2.2010.10.056.

Ritz, S. P., T. F. Stocker, J. O. Grimalt, L. Menviel, and A. Timmermann, 2013: Estimated strength of the Atlantic overturning circulation during the last deglaciation. *Nat. Geosci.,* **6,** 208–212, doi:10.1038/ngeo1723.

Schmith, T., S. Yang, E. Gleeson, and T. Semmler, 2014: How much have variations in the meridional overturning circulation contributed to sea surface temperature trends since 1850? A study with the EC-Earth global climate model. *J. Climate,* **27,** 6343–6357, doi:10.1175/JCLI-D-13-00651.1.

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, doi: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, doi:10.1126/science.1109496.

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, doi:10.1175/BAMS-D-11-00094.1.

Teng, H., G. Branstator, and G. A. Meehl, 2011: Predictability of the Atlantic overturning circulation and associated surface patterns in two CCSM3 climate change ensemble experiments. *J. Climate,* **24,** 6054–6076, doi:10.1175/2011JCLI4207.1.

Terray, L., 2012: Evidence for multiple drivers of North Atlantic multi-decadal climate variability. *Geophys. Res. Lett.,* **39,** L19712, doi:10.1029/2012GL053046.

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

Vellinga, M., and P. Wu, 2004: Low-latitude freshwater influence on centennial variability of the Atlantic thermohaline circulation. *J. Climate,* **17,** 4498–4511, doi:10.1175/3219.1.

Weaver, A. J., M. Eby, M. Kienast, and O. A. Saenko, 2007: Response of the Atlantic meridional overturning circulation to increasing atmospheric CO_2: Sensitivity to mean climate state. *Geophys. Res. Lett.,* **34,** L05708, doi:10.1029/2006GL028756.

Wu, L., and Z. Liu, 2005: North Atlantic decadal variability: Air–sea coupling, oceanic memory, and potential Northern Hemisphere resonance. *J. Climate,* **18,** 331–349, doi:10.1175/JCLI-3264.1.

Yang, X., and Coauthors, 2013: A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J. Climate,* **26,** 650–661, doi:10.1175/JCLI-D-12-00231.1.

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, doi:10.1002/jgrc.20390.

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

——, and Coauthors, 2013: Have aerosols caused the observed Atlantic multidecadal variability? *J. Atmos. Sci.,* **70,** 1135–1144, doi:10.1175/JAS-D-12-0331.1.