


# Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble

Katinka Bellomo<sup>1</sup>  · Lisa N. Murphy<sup>2</sup> · Mark A. Cane<sup>1</sup> · Amy C. Clement<sup>2</sup> · Lorenzo M. Polvani<sup>1,3</sup>

Received: 6 March 2017 / Accepted: 24 July 2017 / Published online: 28 July 2017  
© Springer-Verlag GmbH Germany 2017

**Abstract** Previous studies suggest that internal variability, in particular the Atlantic Meridional Overturning Circulation (AMOC), drives the Atlantic Multidecadal Oscillation (AMV), while external radiative forcing only creates a steady increase in sea surface temperature (SST). This view has been recently challenged and new evidence has emerged that aerosols and greenhouse gases could play a role in driving the AMV. Here we examine the drivers of the AMV using the Community Earth System Model (CESM) Large Ensemble and Last Millennium Ensemble. By computing the ensemble mean we isolate the radiatively forced component of the AMV, while we estimate the role of internal variability using the ensemble spread. We find that phase changes of the AMV over the years 1854–2005 can be explained only in the presence of varying historical forcing. Further, we find that internal variability is large in North Atlantic SST at timescales shorter than 10–25 years, but at longer timescales the forced response dominates. Single forcing experiments show that greenhouse gases and tropospheric aerosols are the main drivers of the AMV in the latter part of the twentieth century. Finally, we show that the forced spatial pattern of SST is not distinct from the internal variability pattern, which has implications for detection and attribution.

## 1 Introduction

The dominant mode of climate variability in the North Atlantic is characterized by relatively warmer SST anomalies in the mid-latitudes (above 40°N) and in the tropics, and exhibits fluctuations on multidecadal timescales (Kushnir 1994). In this paper we refer to this mode as the Atlantic Multidecadal Variability (AMV) and here we consider this to contain both the internal and forced variability. This mode is sometimes referred to as the Atlantic Multidecadal Oscillation (AMO), but it has been suggested to use AMO to refer to the internal variability part only to avoid confusion (Booth 2015). The timescale of the AMV has been estimated to be 60–70 years, but because of the limited length of the observed record, it is still an open question as to whether there is a statistically significant spectral peak at these timescales. The AMV has been linked to several climatic impacts, including rainfall changes in Africa and North America, and tropical cyclones, among others (Enfield et al. 2001, Sutton and Hudson 2005; Knight et al. 2006; Zhang and Delworth 2006; Nigam et al. 2011; Sutton and Dong 2012; Hu and Veres 2016). Resilience to these impacts could improve, if we increase the predictive skill of the AMV.

However, the driving mechanisms of the AMV are debated (Buckley and Marshall 2016). The prevailing view is that ocean dynamics, namely the Atlantic Meridional Overturning Circulation (AMOC), drives the AMV (e.g., Bjerknes 1964; Kushnir 1994; Delworth et al. 1993; Danabasoglu et al. 2012; Gulev et al. 2013; Cheng et al. 2013; Zhang et al. 2013, 2016; McCarthy et al. 2015; O’Reilly et al. 2016). Bjerknes (1964) first hypothesized that the atmosphere drives SST anomalies at short timescales but the slow oceanic circulation drives SST anomalies at longer timescales. Kushnir (1994)’s EOF analysis of observed

---

✉ Katinka Bellomo  
kbellomo@ldeo.columbia.edu

<sup>1</sup> Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

<sup>2</sup> Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL, USA

<sup>3</sup> Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA

SST anomalies supports Bjerknes' hypothesis. They show that on short timescales SST anomalies resemble anomalies driven by the North Atlantic Oscillation (NAO) while at longer timescales SST anomalies become larger in the subpolar gyre where you would expect an influence from the AMOC. However, although Cunningham et al. (2013) argue that the AMOC contributed to the cold tropical SST anomalies in the winter 2009–2010, there is no other observational evidence for a role of the AMOC in driving the AMV in the tropics at long timescales. Moreover, ocean reanalysis products are constrained by observations of the AMOC only since 2004, hence attributing climate anomalies to the AMOC on long timescales is problematic (Baringer et al. 2014). Gulev et al. (2013) provide indirect observational evidence of the Bjerknes' hypothesis using reconstructed surface fluxes. They show that on short timescales the net surface flux is positive and into the ocean (i.e., the atmosphere warms SST from above), while on longer timescales the net surface flux is negative and out of the ocean (i.e., the ocean warms SST from below). They use this as evidence that the AMV is driven by changes in the AMOC. However, Cane et al. (2017) use a simple mathematical model and results from coupled GCMs to show that surface fluxes cannot be used to conclusively reveal the ocean's influence on SST, undermining the interpretation made by Gulev et al. (2013). Other evidence that ocean dynamics drive the AMV comes from model simulations: in some models the AMOC leads the AMV by a few years. However, a lead–lag relationship does not necessarily imply causation and this result seems to be model dependent (Tandon and Kushner 2015; Wang et al. 2017). Further, leads and lags in low-pass filtered data can be contaminated by the filter (Cane et al. 2017; Trenary and DelSole 2016; Foukal and Lozier 2016).

Clement et al. (2015) emphasize that atmospheric processes may also significantly impact the AMV. By comparing fully coupled climate models with atmospheric GCMs coupled to a slab ocean, they show that the model's AMV can be driven by atmospheric white noise integrated by the oceanic mixed layer, and can exist in the absence of ocean dynamics. Murphy et al. (2017) further document that North Atlantic SST has higher variance at decadal and longer timescales in observations and historical simulations compared to preindustrial control simulations. They suggest that atmospheric historical forcings play a substantial role in setting the phase changes and amplitude of the observed North Atlantic SST (NASST) index over the last century, although most CMIP5 models still underestimate the observed variance of North Atlantic SST variability even when historical forcings are included. Their findings are consistent with previous studies (e.g.,; Mann and Emanuel 2006, Ottera et al. 2010; Terray 2012; Booth et al. 2012). Among these, Booth et al. (2012) show in one

model that forced aerosol indirect effects can explain the observed variance in the North Atlantic, with a small role for internal variability.

Nevertheless, the role of external radiative forcings in driving the AMV is far from settled. For example, Zhang et al. (2013) rebut the study of Booth et al. (2012), claiming that in the model they used long-term trends in SST may be too sensitive to aerosol loading. Knight (2009), Ting et al. (2009, 2014) and DelSole et al. (2010) apply statistical methods and multi-ensemble averaging to extract the externally forced signal from internal variability in the AMV in CMIP models. They argue that external forcing cannot explain the variance of the observed AMV anomalies, which they suggest is driven by internal climate variability. Here we build on these previous studies and examine the role of external radiative forcing and internal variability in driving the temporal and spatial characteristics of the AMV using the CESM Large Ensemble. We take advantage of the large ensemble design by using the ensemble mean to isolate the external forcing and the ensemble spread to estimate the role of internal variability.

## 2 Data

We compare observed SST from the Extended Reconstructed SST version 4 (ERSSTv4) reanalysis (Huang et al. 2014) with historical simulations in the National Center for Atmospheric Research (NCAR) Community Earth System Model (CESM) Last Millennium Ensemble (Otto-Bliesner et al. 2016) and the Large Ensemble (Kay et al. 2015). We examine the ten historical simulations available from the CESM Last Millennium Ensemble (LME), each forced with the same observational estimates of historical forcings but initialized from different atmospheric conditions (ocean initial conditions are the same in all members). The model used is version 1.1 of CESM Community Atmosphere Model version 5 (CESM-CAM5) run at 2° resolution for atmosphere and land and 1 degree for the ocean and sea ice. The LME spans the years 850–2005, but we analyze the years 1854–2005 to be consistent with observations. In addition to the LME, we examine the 42 historical simulations available from the CESM Large Ensemble (LE), which are run with the same model but at a finer atmospheric resolution of 1°. Similarly to the LME, each ensemble member in the LE is run with the same prescribed forcings and initialized from different atmospheric conditions. The LE spans the years 1920–2005. The prescribed historical forcings include greenhouse gases, ozone, tropospheric and stratospheric aerosols, land use, and solar radiance. These forcings are the same in the LME and LE for the overlapping years of coverage except that the LME also includes

orbital changes in insolation, which are minor over the last century. More details on the sources of the forcings and the design of the LME can be found in Otto-Bliesner et al. (2016). We compare historical LE and LME ensemble members with their respective preindustrial control simulations in which climate variability is solely internally driven since all forcings are fixed throughout the simulations. The length of the control simulations is 1800 years for the LE and 1000 years for the LME. We will show that differences between the LME and LE do not arise from atmospheric resolution or the inclusion of orbital forcing in the LME, but from the different years analyzed. We note that results throughout the paper do not change if we use different observational SST datasets.

### 3 Methods

Observations of SST are influenced by both internal variability and external forcing, and so does the SST simulated by each historical member in the large ensembles. To quantify the contribution of radiative forcing alone in driving SST, we compute the ensemble mean of the LME (10 members) and LE (42 members). The ensemble mean filters out internal variability originating from different atmospheric initial conditions in each ensemble member while retaining the radiative forcing. We also examine single forcing experiments that are included as part of the LME project: well-mixed greenhouse gases (3 members), tropospheric aerosols plus ozone (2), solar (4), stratospheric aerosols (5), land use (3), and orbital (3). We present only the ensemble mean for each of these single forcing experiments and note that the small ensemble size of the single forcing experiments may preclude the full removal of internal variability.

For both observations and each ensemble member we compute monthly SST anomalies by removing the long-term monthly mean from each month. We linearly detrend SST to roughly remove the monotonic increase in SST due to the net effect of CO<sub>2</sub>, and focus on the role of the other forcing agents in driving SST variability. Therefore, the detrended ensemble mean in the LME or LE should be interpreted as being radiatively forced by all forcing agents except the linear warming trend. Murphy et al. (2017) examined other methods to remove the global warming trend in the North Atlantic in CMIP5 models, including the LE. They found similar results if, instead of linearly detrending, they removed the warming rate associated with CO<sub>2</sub> through linear regression. Instead, they found that using the Trenberth and Shea (2006) method, which consists of subtracting the global mean SST from the NASST index, removes most of the forced component and not just the net warming trend.

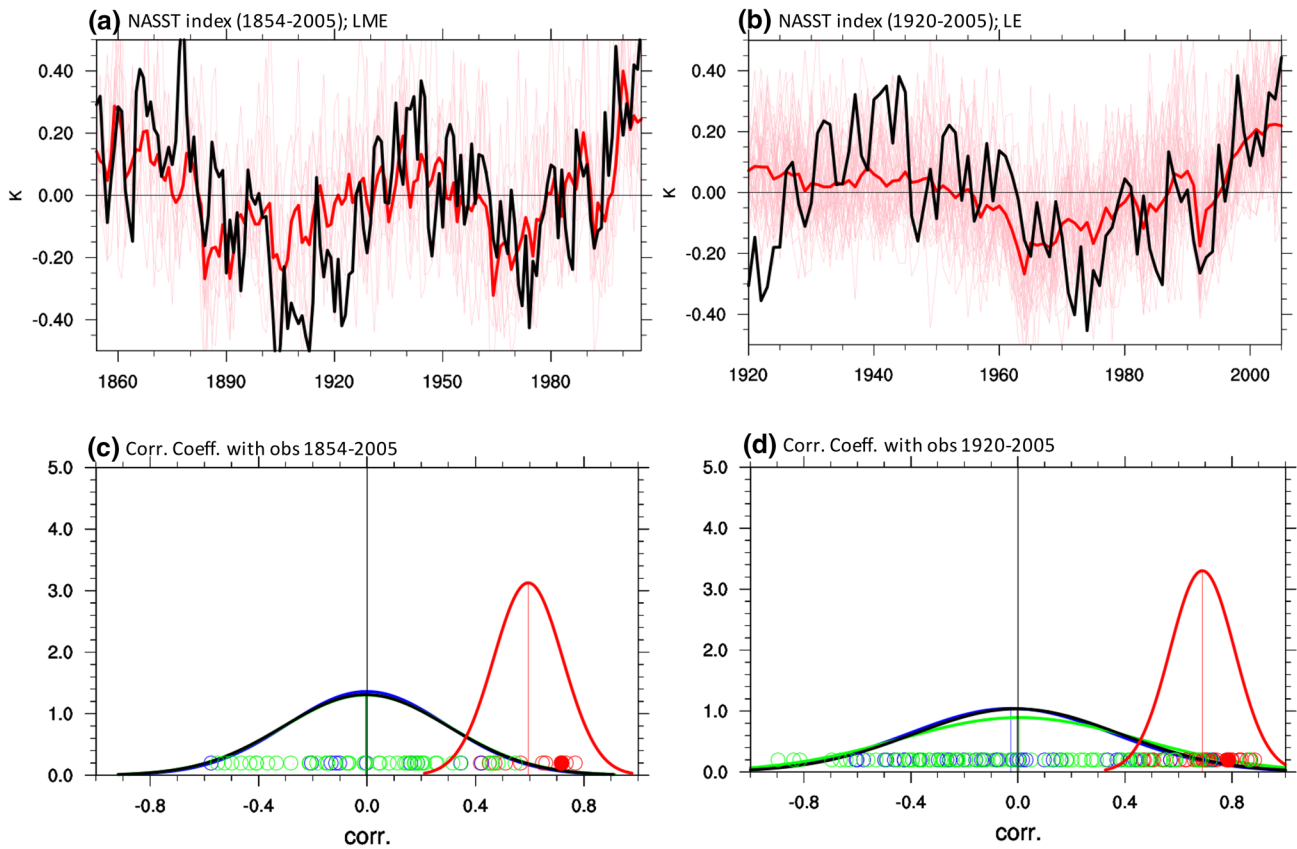
## 4 Results

### 4.1 Forced and internal variability

The black curves in Fig. 1a, b represent the observed NASST index over the years 1854–2005 (Fig. 1a) and 1920–2005 (Fig. 1b). Each thin red line represents an ensemble member from the LME (Fig. 1a) and LE (Fig. 1b), while the thick red lines are the ensemble mean from each ensemble. The ensemble mean alone seems to capture the phase changes of the AMV, except in Fig. 1b where the ensemble mean seems to change phase from positive to negative a few years before the mid 1960s, unlike observations. This indicates that remaining forcings (i.e., aerosols, ozone, solar radiance, etc.) drive variability in the NASST, in addition to their impact on the longer term SST trend.

In Fig. 1c, d, we apply a 20-year Lanczos low-pass filter to isolate the multi-decadal variability in the NASST index (Fig. 1a, b). Using a 10-year filter rather than 20-year does not significantly change the results. Hereafter, we will refer to the low-pass filtered linearly detrended NASST index as the AMV index. Figure 1c shows the correlation coefficient between the AMV index in the model and observations. The higher the correlation coefficient the better the model simulates the observed phase changes of the AMV. Each open red dot represents the correlation between a historical ensemble member and observations, while the filled red dot is the correlation coefficient between the ensemble mean and observations. As an aid to visualization, a red Gaussian probability distribution function (PDF) is fitted from the average correlation and the standard deviation computed from the open red dots. The blue curve and dots represent the PDF of correlations between observations and each ensemble member minus the ensemble mean (hereafter, termed “de-meaned”), while the green curve and dots represent the PDF between observations and time series of the same length as observations drawn from the long preindustrial control simulations. Both the blue and green dots and respective distributions represent the correlation between the model AMV driven solely by internal variability and the observed AMV. The black PDF encompasses correlations between 20-year low-pass filtered and linearly detrended white noise time series generated with random numbers (where the mean of the time series is 0 and standard deviation is 1) and observations. An equivalent analysis and corresponding colors are used for the LE and respective years in Fig. 1d.

Removing the ensemble mean from each ensemble members leads to the same PDF (blue) as the preindustrial control simulation (green), which indicates that removing the ensemble mean is an effective way of removing the radiative forcing while leaving internal variability. It further suggests



**Fig. 1** **a, b** Time series of the unfiltered detrended NASST index (SST averaged over the North Atlantic: 0–60N, 80W–0) in observations (*black*), ensemble members (*thin red lines*), ensemble mean (*thick red line*) for the years **a** last millennium ensemble (1854–2005) and **b** large ensemble (1920–2005). **c, d** Correlation coefficients and

PDF between observed AMV index (computed as the detrended 20 year low-pass filtered NASST index) and ensemble members. *Red* historical, *red filled* ens mean, *blue* de-means (historical minus ensemble mean), *green* preindustrial control, *black* random numbers

that internal and forced variability tend to add linearly, rather than overwrite each other. The blue and green PDFs resemble Gaussian distributions generated by a random process with red noise characteristics. In fact, the black distribution obtained from random numbers strikingly resembles the blue and green distributions. Although the black distribution is obtained from white noise time series, because we low-pass filter them before computing the PDF, we introduce low-frequency variability and the ultimate time series are red noise. Therefore, the correlation coefficients due to internal variability (blue and green curves) can be interpreted as indistinguishable from randomly generated red noise time series. A Kolmogorov–Smirnov test confirms that the red (historical) PDF belongs to a different distribution than the other PDFs. The filled red dots representing external forcing fall within the upper range of the red PDFs. The filled red dot value is 0.72 for the LME for the years 1854–2005, 0.79 for the LE for the years 1920–2005, and 0.85 for the LME computed for the years 1920–2005, suggesting that forcings may play a larger role in the latter part of the twentieth century (although

it might also be that estimates of historical forcings are more accurate post 1920). All these results underscore a fundamental role for radiative forcing in setting the timing of the phase changes of the AMV index. Without external forcing (green and blue PDFs) it is extremely unlikely that the model captures the timing of phase changes in the AMV. A similar result is shown for other CMIP5 models in Murphy et al. (2017).

We introduce a framework to interpret the magnitude of correlations in Fig. 1 and quantify the forced to total variability ratio for the correlation coefficients. The correlation between the ensemble mean and observations (filled red dots) can be written as:

$$\rho_{ensm} = \frac{\text{cov}(AMV_{obs}, AMV_{ensm})}{\sqrt{\sigma_{obs}^2} \times \sqrt{F_{ensm}^2 + \frac{I_{ensm}^2}{n}}} \quad (1)$$

where “*ensm*” stands for ensemble mean, “*cov*” for covariance, “*obs*” for observed, “*F*<sup>2</sup>” for forced variance and “*I*<sup>2</sup>”

for internal variance, while “ $n$ ” is the number of ensemble members for the model. Here it is assumed that the total variance “ $\sigma^2$ ” is the linear combination of forced and internal variance, and that the internal variability in the observations does not co-vary with the forcing in the model.  $I_{ensm}^2$  is computed as the average of the correlation of de-measured ensemble members (internal variability) and observations. Instead, the average correlation between observations and all the ensemble members (red lines) can be written as:

$$\rho_{avg} = \frac{1}{n} \sum_{j=1}^n \frac{cov(AMV_{obs}, AMV_j)}{\sqrt{\sigma_{obs}^2} \times \sqrt{\sigma_j^2}} \tag{2}$$

As a check, the correlation ratio  $\frac{\rho_{ensm}}{\rho_{avg}}$  is 1.18 for the LME (1854–2005) and 1.17 for the LE (1920–2005), while the ratios of the RHS of Eqs. (1) and (2) are 1.21 and 1.14, respectively (we ignore the fact that the very small term  $\frac{I_{ensm}^2}{n}$  is, strictly speaking, not part of the forced variability).

We do not know the forced to total variability ratio in the observations, but we can obtain some informative bounds for it. We can write the “perfect” correlation between observations and ensemble mean in the hypothetical case that the observations have zero internal variability:

$$\rho_{per} = \frac{cov(AMV_{obs}, AMV_{ensm})}{\sqrt{F_{obs}^2} \times \sqrt{F_{ensm}^2 + \frac{I_{ensm}^2}{n}}} \tag{3}$$

Using Eq. (1) we obtain:

$$\rho_{per} = \rho_{ensm} \times \sqrt{\frac{\sigma_{obs}^2}{F_{obs}^2}} = \frac{\rho_{ensm}}{\mu} \tag{4}$$

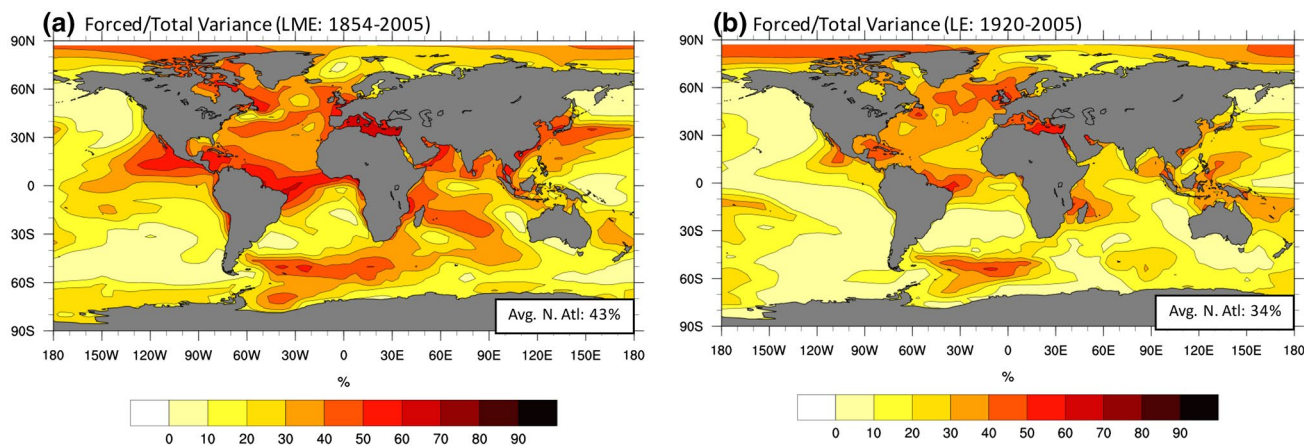
where “ $\mu^2$ ” is the ratio of observed forced variance to observed total variance  $\frac{F_{obs}^2}{\sigma_{obs}^2}$ . Because the largest possible

value for  $\rho_{per}$  is by definition (i.e. in the hypothetical case that the model perfectly captures the observed signal), Eq. (4) implies that the smallest possible value for  $\mu^2 = \frac{F_{obs}^2}{\sigma_{obs}^2}$  is

$\rho_{ensm}^2$ . Since  $\rho_{per} \geq \rho_{ensm}$  the largest possible value for  $\mu^2$  is 1, which would hold if there were no observed internal variability at all. In any case the range of allowed values for  $\rho_{per}$  indicates that the forced variance in the model is in this sense accounting for a great deal of the observed forced variability. If we plug the  $\rho_{ensm}$  values from Fig. 1 (filled red dots) of 0.72 for the LME and 0.79 for the LE into Eq. (4) we find the following bounds for the observed forced to total variance: for the LME period 1854–2005 is  $0.52 \leq \frac{F_{obs}^2}{\sigma_{obs}^2} \leq 1$  while for the LE period 1920–2005 is  $0.62 \leq \frac{F_{obs}^2}{\sigma_{obs}^2} \leq 1$ . That is, in the LME period 1854–2005

forced part is at least 52% of the total, while in the LE period 1920–2005 it is at least 62%. While this analysis allows for the observed forced total variance to be as high as 100%, it is most likely that the true value is somewhere around the lower limits.

In Fig. 1 we estimated the contribution of radiative forcing in driving *phase changes* of the AMV through correlations between the observed and modeled AMV. To quantify the role of radiative forcing in driving the *amplitude* of AMV anomaly patterns, we now compare the variance of the AMV in the model with and without radiative forcing. In Fig. 2 we show the percentage of forced to total variance in each grid point, obtained dividing the ensemble mean (forced) variance of SST by the average of all



**Fig. 2** Spatial pattern of forced variance divided by total variance of SST. Forced variance is the variance of the ensemble mean SST. Total variance is the average of the SST variances across all ensemble members. All data are detrended and filtered with a 20 year low-

pass Lanczos filter. **a** Last millennium ensemble (1854–2005), **b** large ensemble (1920–2005). The insets in the lower right in both panels indicate the forced to total variance averaged in the North Atlantic basin

(forced + internal) variances of SST computed as the average across all ensemble members. It clearly emerges that external forcing affects low-frequency SST variability in the North Atlantic, unlike regions such as the tropical Pacific where we expect internal variability to be larger than external forcing. If we take an average of the North Atlantic basin from Fig. 2, we find that external forcing explains 43% of the total variance in the LME for the years 1854–2005 and 34% in the LE for the years 1920–2005.

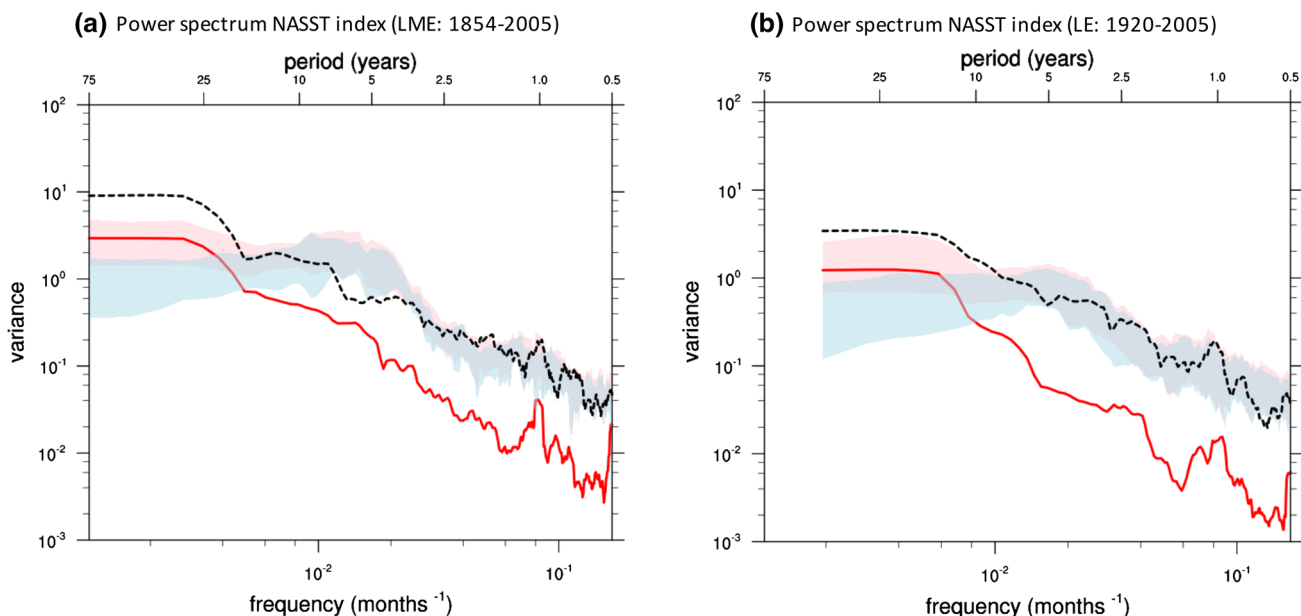
To further assess the robustness of these results and the choice of the ensemble size, we subsampled the LE into 33 groups of ten ensemble members each (starting from the first to tenth member, second to eleventh, third to twelfth, etc.), which is the same ensemble size as the LME. We find that the forced to total variance in the North Atlantic across the groups ranges 34–47% and the average is 39% for the years 1920–2005. The forced to total variance in the LME for the years 1920–2005 is 39%, which is within the

range of each group of ten ensemble members from the LE. The forced to total variance for the ensemble mean AMV index (20-year low-pass filtered NASST index) is 68% for the LME (1854–2005) and 72% for the LE (1920–2005). If we use the LME for the years 1920–2005 we obtain 70%. The range across the 33 groups of ten ensemble members in the LE is 70–81% and the average is 73%. These values are summarized in Table 1. We note that this sub-sampling technique might give less weight to the simulations that are at the beginning or end of the ensemble, which are included less times than those in the middle. Therefore, we repeated this calculation drawing random groups of ten ensemble members but found similar results (not shown). As a comparison the observational estimates of forced to total variance of the AMV index derived above were at least 52% for the period 1854–2005 (vs. 68% in the model) and 62% for the period 1920–2005 (vs. 72% in the model). Finally, if we apply 20-year high-pass filter to the ensemble mean NASST index in Fig. 1 we obtain that the forced to total variance is 12% for the LME and 8% for the LE, indicating that variability at timescales <20 years is largely internally driven.

**Table 1** Forced to total variance averaged over the North Atlantic (first row) and for the AMV index (second row)

	LME (1854– 2005)	LE (1920– 2005)	LME (1920– 2005)	LE (33 groups of ten members each)
Average North Atl	43%	34%	39%	Range 34–47% Average 39%
AMV index	68%	72%	70%	Range 70–81% Average 73%

To better understand the forced component of the AMV variance as a function of timescale and compare with observations, we compute the power spectra of the unfiltered NASST indices (Fig. 3). Black lines represent the power spectrum of the observed NASST for the years 1854–2005 (Fig. 3a) and 1920–2005 (Fig. 3b). Superimposed thick red lines are the power spectra of the ensemble



**Fig. 3** Power spectra of unfiltered and detrended NASST index. Black line is observations. Red line is the ensemble mean (forced component). Red envelope spans the historical ensemble members

(forced + internal variability). Light blue envelopes span de-meaned ensemble members (internal variability). **a** Last millennium ensemble (1854–2005), **b** large ensemble (1920–2005)

mean NASST for the LME 1854–2005 (Fig. 3a) and LE 1920–2005 (Fig. 3b). The red and blue envelopes represent the ensemble spread of the NASST in the (red) historical and (blue) de-meaned ensembles. We note that both models and observations have red power spectra with variance increasing as a function of period and no spectral peak. At multi-decadal timescales the observations lie above the ensemble variance, indicating that the CESM-CAM5, in common with most models in the CMIP5 archive, underestimates multi-decadal NASST variability (Peings et al. 2016; Murphy et al. 2017).

The key finding here is that the ensemble mean (thick red line) has lower variance than the ensemble spread (red envelope) up to 10–25 year timescales, but it falls within the ensemble spread at longer timescales. This confirms that variability in the AMV index at timescales <10–25 years is predominantly internally driven, while at timescales >10–25 years radiative forcing drives a substantial part of the variance (as computed above, 68–72% of the 20-year low-pass filtered NASST index is estimated to be forced). Comparing the red and blue envelopes further corroborates this point: internal variability only (blue) overlays internal plus forced variability (red) up to 10–25 years, then internal variability decreases, which means that the higher variance observed at longer timescales must be mostly driven by radiative forcing. This analysis agrees with the earlier estimates of forced to total variance in the low-pass/high-pass AMV index and reported in Table 1. The results found in Fig. 3 do not change if instead of using de-meaned ensembles for the blue envelopes we use chunks of the same length as observations drawn from the preindustrial control simulation. We note, however, that the forced variance emerges at >10 years in the LE for the years 1920–2005 and at >20 years in the LME for the years 1854–2005. If we compute the power spectra for the years 1920–2005 in the LME we find that the forced variance also emerges at >10 years as in the LE, indicating that the forced to total variance ratio may be dependent on the time frame chosen in the analysis.

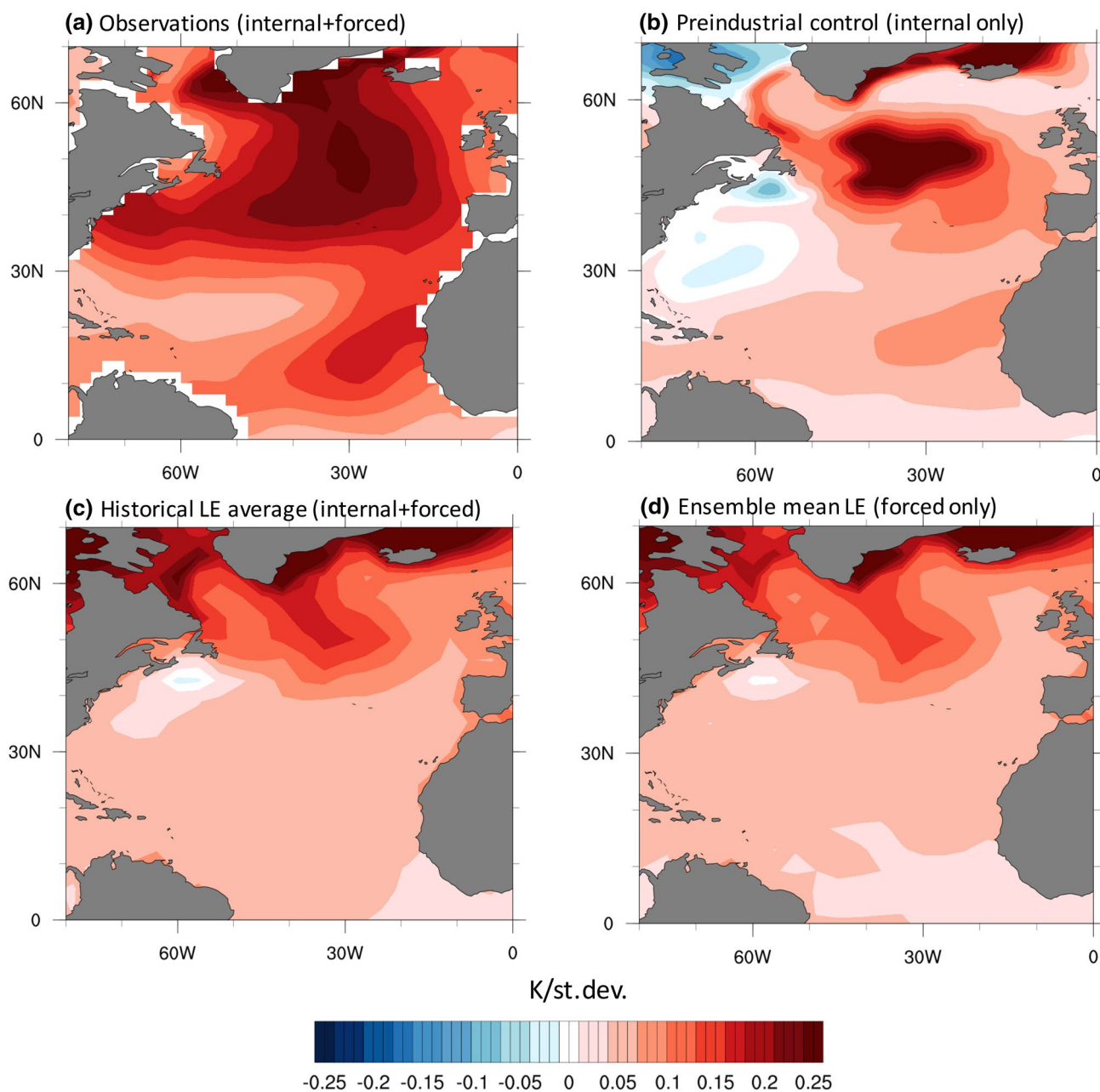
## 4.2 Spatial patterns

By computing regressions of local SST on the AMV index (Fig. 4), we characterize the role of forcing in driving the spatial pattern associated with the AMV. In this figure we show the years 1920–2005 for observations and the LE. The observed regression of SST on the AMV index is shown in Fig. 4a. This pattern is somewhat reproduced in the preindustrial control simulation (Fig. 4b), however observations show a more widespread warming, while the preindustrial control shows an enhanced warming over the subpolar gyre. The average regression across the historical simulations (Fig. 4c) is less similar to observations,

and smaller in amplitude in the tropics, which is a problem shared with several other global climate models (e.g., Kavvada et al. 2013; Ba et al. 2014; Martin et al. 2014; Yuan et al. 2016; Bellomo et al. 2016). Figure 4d shows the spatial pattern associated with radiative forcing alone (from the LE), and is obtained by computing the ensemble mean SST at each grid point and regressing it onto the ensemble mean standardized AMV index. It emerges from these figures that radiative forcing (Fig. 4d) projects onto the unforced AMV especially in the subpolar gyre region. We note that these results do not change much if we compute the first EOF instead of regression or use de-meaned ensembles instead of the preindustrial control simulation. Results do not change much if we use LE or LME for the years 1920–2005 (although there are some slight differences between the 1854–2005 and 1920–2005 periods), or subsample the LE into groups of ten members.

## 4.3 Single forcing experiments

To evaluate the influence of each forcing agents on the AMV, we examine single forcing experiments from the LME. In Fig. 5 we show which of the forcing agents contribute to the total forced variability in the LME. We compute the correlation coefficient between the AMV index from the all forcing experiment (LME) and the AMV index computed from single forcing experiments, linearly adding one AMV index from a single forcing experiment at the time (shown in Fig. 5b). If adding a single forcing increases the correlation, that forcing contributes to the total forced variability, otherwise it only adds uncorrelated variability (or variability correlated with a different forcing ahead of it in the list in Fig. 5b, therefore not contributing further in increasing the correlation). Figure 5a shows the (black) ensemble mean AMV index from the all forcings experiment (LME) and the (red) sum of the AMV indices from the following single forcing experiments, which maximize the correlation coefficient shown in Fig. 5b: greenhouse gases, tropospheric aerosols plus ozone, stratospheric aerosols (i.e., volcanic eruptions), and solar. By separating the time series into different time periods we find that stratospheric aerosols mostly add an increase in the correlation in the earlier part of the record (before 1950, not shown). Therefore, in the latter part of the twentieth century greenhouse gases and tropospheric aerosols plus ozone explain most of the phase changes of the forced AMV. Also note that the AMV time series in each of the forcing experiments were added starting from the time series with largest variance to the lowest, but changing the order in the sum of the AMV time series still leads to the same result (i.e., that the four forcings that maximize the correlation are the ones shown in Fig. 5a).



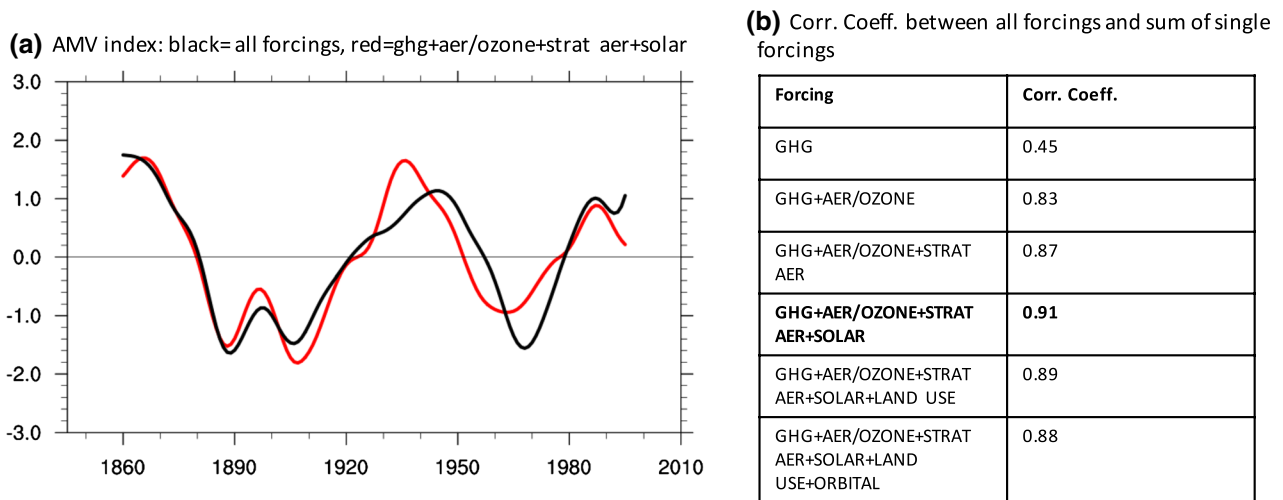
**Fig. 4** Regression of SST on the standardized AMV index: **a** observations (years 1920–2005), **b** preindustrial control (200 years), **c** historical simulation obtained by averaging the regression of SST on the AMV index across all the LE members, **d** LE mean (forced com-

ponent only) obtained as the regression of the ensemble mean SST at each grid point on the ensemble mean AMV index. All data are detrended and low pass filtered with a 20 year Lanczos filter. Units are K per standard deviation of the AMV index

In Fig. 6 we show the spatial patterns associated with the four single forcing experiments highlighted in Fig. 5. These patterns are obtained as in Fig. 4d by regressing the ensemble mean SST at each grid point onto the ensemble mean AMV index, except that now the ensembles only contain single forcing experiments and fewer members. These plots interestingly show that all forcings

contribute to anomalies in the mid-latitudes over the sub-polar gyre, but that tropospheric aerosols plus ozone also enhances SST in the tropics (as argued in Booth et al. 2012). It remains to be tested with more models and ensemble members the extent to which tropospheric aerosols affect the tropics and what is the role for cloud–aerosol interactions.





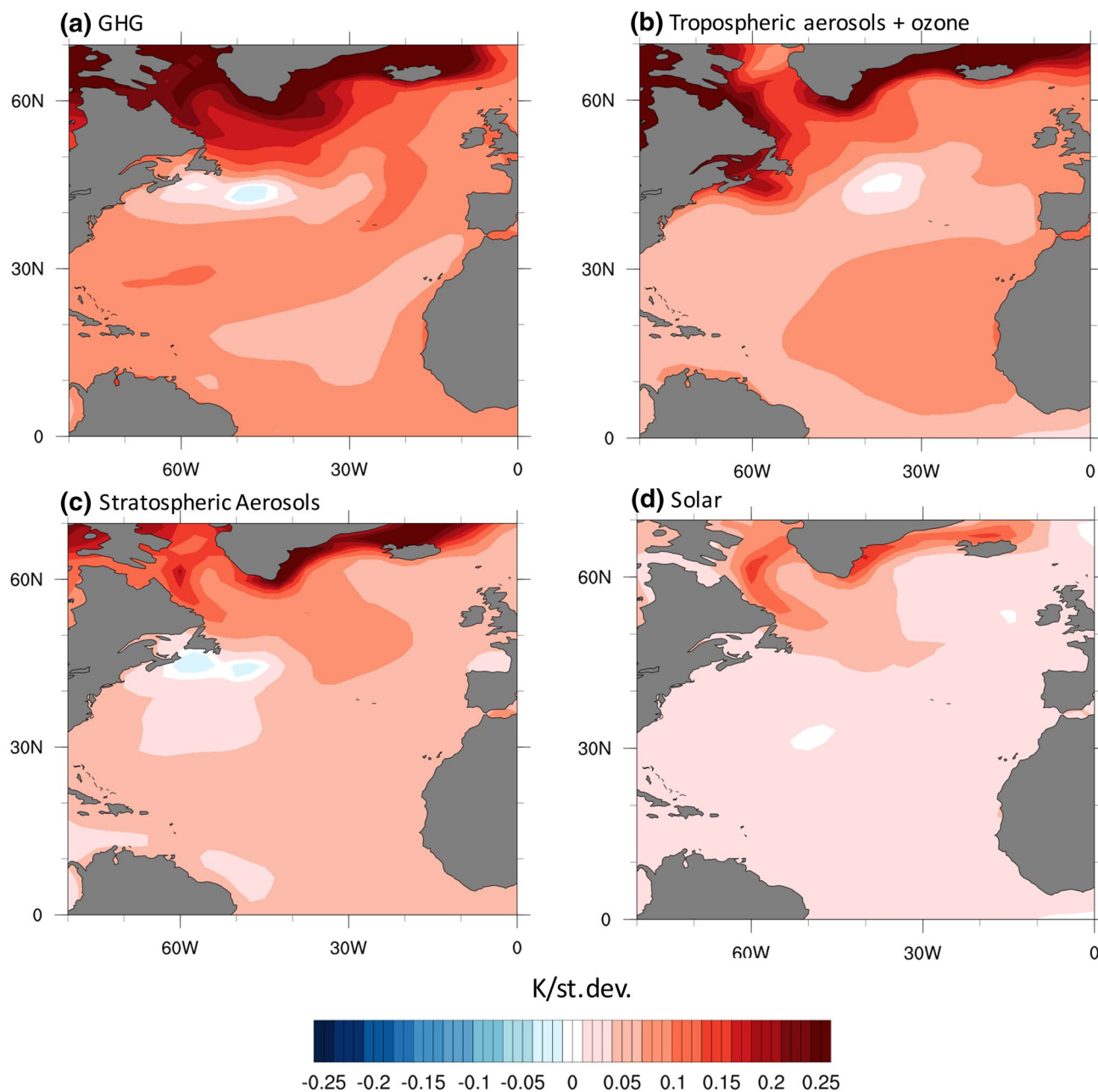
**Fig. 5** **a** AMV index: (black) ensemble mean of all historical forcings (LME) and (red) sum of the four forcings that give the highest correlation coefficient (see **b**): ghg + aer/ozone + strat aer + solar; **b** correlation coefficient between ensemble mean AMV index and AMV index in single forcing experiments adding one forcing each

### 5 Discussion and conclusions

Using one climate model (CESM-CAM5) and a large ensemble, we find that historical forcings play a substantial role in driving the observed AMV. We show that the phase changes of the AMV (as measured by the correlation coefficient) are very unlikely to be driven by internal variability. We further estimate the forced to total variance ratio in the model, which measures the influence of radiative forcing on the amplitude of AMV anomalies: we find that at timescales longer than 20 years, 34–43% of the average of North Atlantic SST variance, and 68–72% of the AMV index variance, are radiatively forced. In contrast, at timescales shorter than 20 years only 8–12% of the AMV index variance is forced. Showing power spectra of the ensemble mean and spread, we provide additional evidence that most part of the AMV variance is largely radiatively forced at timescales >10–25 years but internally driven at shorter timescales. In addition, we use regression analysis to show that radiative forcing spatially projects onto the unforced AMV pattern with largest loadings in the subpolar gyre region. Lastly, we examine single forcing experiments from the LME and show that greenhouse gases, tropospheric aerosols plus ozone, stratospheric aerosols (volcanic eruptions) and solar forcings are the largest contributors to AMV variability over the twentieth century. All of these four forcings project onto mid-latitude anomalies. Only the aerosols plus ozone single forcing experiment shows enhanced SST anomalies in the tropics, which could be related to aerosol–cloud interactions or low-level cloud feedbacks (Martin et al. 2014; Evan et al. 2013; Bellomo

et al. 2016). However, the small ensemble size of the single forcing experiments limits the ability to filter out internal variability.

Here we do not provide a direct measurement of the observed variance explained by external forcing in the model, although we estimate that the observed forced to total variance of the AMV index is at least 52% for the period 1854–2005 and 62% for the period 1920–2005. Murphy et al. (2017) directly compared observed and simulated variability of the AMV using CMIP5 models, including the LE (see their Fig. 1b). They found that while models generally underestimate observed variance (including CESM-CAM5), certain ensemble members of some models are able to simulate variance of the AMV of the same strength if not bigger than observed, only when historical forcings are included. Similar results were found by Peings et al. (2016) for other CMIP5 models. In this study we also do not evaluate the possible influence of the North Atlantic Oscillation (NAO) on the AMV, through its influence on the AMOC. Recent studies by Delworth and Zeng (2016) and Delworth et al. (2017) demonstrate that artificially increasing the persistence of the NAO in the GFDL model through surface fluxes creates long-term fluctuations in the AMOC and associated heat transport, which in turn affect the amplitude of the AMV at multi-decadal timescales, making it more similar to observations. Low-frequency variability in the NAO is generally underestimated in CMIP5 models, including CESM-CAM5 (Wang et al. 2017) and could potentially bring the variance of the AMV closer to observations even in the absence of forcings, but this requires an explanation for why the models



**Fig. 6** Regression of ensemble mean SST on the standardized ensemble mean AMV index in single forcing experiments (representing forced component only as in Fig. 4d): **a** GHG, **b** Tropospheric

aerosols plus ozone, **c** Stratospheric aerosols, **d** Solar forcing. All data are detrended and low pass filtered with a 20 year Lanczos filter. Units are K per standard deviation of the AMV index

underestimate the NAO, which could be related to models underestimating the radiative forcing effect on the NAO. It also remains to be clarified how the AMOC influences SST in the North Atlantic in the ocean models, given the recent study of Foukal and Lozier (2016) which shows no surface inter-gyre exchange between the SST in the subtropical and subpolar gyres. Another possible caveat in this paper is that the ocean initial conditions are the same in all ensemble members, therefore affecting the persistence of the AMV

and undermining the role of the external radiative forcing. However, we note that the LME starts from year 850, therefore we believe that any ocean memory will be long lost after 1000 years.

In conclusion, the results presented herein provide evidence that a sizeable part of the observed AMV variability may be externally forced. The implication is that the predictive skill of the AMV requires a knowledge of radiative forcing and a better understanding of how it projects onto

the spatial pattern and internal variability. These results are based on one model only and should be corroborated with more models that can provide large ensembles.

**Acknowledgements** KB was funded by a cooperative agreement between NASA and Columbia University (NNX15AJ05A). AC and LM were funded by NSF P2C2 program. Large ensemble data are made available by the CESM Large Ensemble Community Project and supercomputing resources provided by NSF/CISL/Yellowstone. NOAA ERSSTv4 data is provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at <http://www.esrl.noaa.gov/psd/>. We thank two anonymous reviewers for helping improving the analysis presented here.

## References

- Ba J et al (2014) A multi-model comparison of Atlantic multidecadal variability. *Clim Dyn* 43(9–10):2333–2348
- Baringer MO, McCarthy GD, Willis J, Lankhorst M, Smeed DA, Send U, Rayner D, Johns WE, Meinen CS, Cunningham SA, Kanzow TO, Frajka-Williams E, Marotzke J (2013) Meridional overturning circulation observations in the North Atlantic Ocean. *Bull Am Meteor Soc* 95(7):S67–S69. doi:[10.1175/2014BAMSStateoftheClimate.1](https://doi.org/10.1175/2014BAMSStateoftheClimate.1)
- Bellomo K, Clement AC, Murphy LN, Polvani L, Cane MA (2016) New observational evidence for a positive cloud feedback that amplifies the Atlantic multidecadal oscillation. *Geophys Res Lett*. doi:[10.1002/2016GL069961](https://doi.org/10.1002/2016GL069961)
- Bjerknes J (1964) Atlantic air sea interaction. *Adv Geophys* 10:1–82
- Booth BB (2015) Why the Pacific is cool. *Science* 347(6225):952. doi:[10.1126/science.aaa4840](https://doi.org/10.1126/science.aaa4840)
- Booth BB, Dunstone NJ, Halloran PR, Andrews T, Bellouin N (2012) Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature* 484(7393):228–232. doi:[10.1038/nature10946](https://doi.org/10.1038/nature10946)
- Buckley MW, Marshall J (2016) Observations, inferences, and mechanisms of Atlantic meridional overturning circulation variability: a review. *Rev Geophys* 54:5–63. doi:[10.1002/2015RG000493](https://doi.org/10.1002/2015RG000493)
- Cane M, Clement A, Murphy L, Bellomo K (2017) Low pass filtering, heat flux and Atlantic multidecadal variability. *J Clim*. doi:[10.1175/JCLI-D-16-0810.1](https://doi.org/10.1175/JCLI-D-16-0810.1)
- Cheng W, Chiang JCH, Zhang D (2013) Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. *J Clim* 26:7187–7197. doi:[10.1175/JCLI-D-12-00496.1](https://doi.org/10.1175/JCLI-D-12-00496.1)
- Clement A, Bellomo K, Murphy LN, Cane MA, Mauritsen T, Rädel G, Stevens B (2015) The Atlantic multidecadal oscillation without a role for ocean circulation. *Science* 350:320–324
- Cunningham SA, Roberts CD, Frajka-Williams E, Johns WE, Hobbs W, Palmer MD, Rayner D, Smeed DA, McCarthy G (2013) Atlantic meridional overturning circulation slowdown cooled the subtropical ocean. *Geophys Res Lett* 40:6202–6207. doi:[10.1002/2013GL058464](https://doi.org/10.1002/2013GL058464)
- Danabasoglu G, Yeager SG, Kwon Y-O, Tribbia JJ, Phillips AS, Hurrell JW (2012) Variability of the Atlantic meridional overturning circulation in CCSM4. *J Clim* 25:5153–5172. doi:[10.1175/JCLI-D-11-00463.1](https://doi.org/10.1175/JCLI-D-11-00463.1)
- DelSole T, Tippett MK, Shukla J (2010) A significant component of unforced multidecadal variability in the recent acceleration of global warming. *J Clim*. doi:[10.1175/2010JCLI3659.1](https://doi.org/10.1175/2010JCLI3659.1)
- Delworth TL, Zeng F (2016) The impact of the North Atlantic oscillation on climate through its impact on the Atlantic meridional overturning circulation. *J Clim*. doi:[10.1175/JCLI-D-15-0396.1](https://doi.org/10.1175/JCLI-D-15-0396.1)
- Delworth T, Manabe S, Stouffer RJ (1993) Interdecadal variations of the thermohaline circulation in a coupled ocean–atmosphere model. *J Clim* 6:1993–2011
- Delworth TL, Zeng F, Zhang L, Zhang R, Vecchi G, Yang X (2017) The central role of ocean dynamics in connecting the North Atlantic oscillation to the Atlantic multidecadal oscillation. *J Clim*. doi:[10.1175/JCLI-D-16-0358.1](https://doi.org/10.1175/JCLI-D-16-0358.1)
- Enfield DB, Mestas-Núñez AM, Trimble PJ (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophys Res Lett* 28:2077–2080. doi:[10.1029/2000GL012745](https://doi.org/10.1029/2000GL012745)
- Evan AT, Allen RJ, Vimont DJ, Bennartz R (2013) The modification of sea surface temperature anomaly linear damping time scales by stratocumulus clouds. *J Clim* 26:3619–3630
- Foukal NP, Lozier MS (2016) No inter-gyre pathway for seasurface temperature anomalies in the North Atlantic. *Nat Commun*. doi:[10.1038/ncomms11333](https://doi.org/10.1038/ncomms11333)
- Gulev SK, Latif M, Keenlyside N, Park W, Koltermann KP (2013) North Atlantic Ocean control on surface heat flux on multidecadal timescales. *Nature* 499(7459):464–467. doi:[10.1038/nature12268](https://doi.org/10.1038/nature12268)
- Hu Q, Veres MC (2016) Atmospheric responses to North Atlantic SST anomalies in idealized experiments. Part II: North American precipitation. *J Clim* 29:659–671. doi:[10.1175/JCLI-D-14-00751.1](https://doi.org/10.1175/JCLI-D-14-00751.1)
- Huang B, Banzon VF, Freeman E, Lawrimore J, Liu W et al (2014) Extended reconstructed sea surface temperature version 4 (ERSST.v4): Part I. Upgrades and intercomparisons. *J Clim*. doi:[10.1175/JCLI-D-14-00006.1](https://doi.org/10.1175/JCLI-D-14-00006.1)
- Kavvada A, Ruiz-Barradas A, Nigam S (2013) AMO's structure and climate footprint in observations and IPCC AR5 climate simulations. *Clim Dyn* 41(5–6):1345–1364
- Kay JE, Deser C, Phillips A, Mai A, Hannay C, Strand G et al (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 Am Meteorol Soc* 96:1333–1349. doi:[10.1175/BAMS-D-13-00255.1](https://doi.org/10.1175/BAMS-D-13-00255.1)
- Knight JR (2009) The Atlantic multidecadal oscillation inferred from the forced climate response in coupled general circulation models. *J Clim* 22:1610–1625. doi:[10.1175/2008JCLI2628.1](https://doi.org/10.1175/2008JCLI2628.1)
- Knight JR, Allan RJ, Folland CK, Vellinga M, Mann ME (2005) A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys Res Lett* 32:L20708
- Knight JR, Folland CK, Scaife AA (2006) Climate impacts of the Atlantic multidecadal oscillation. *Geophys Res Lett* 33:L17706. doi:[10.1029/2006GL026242](https://doi.org/10.1029/2006GL026242)
- Kushnir Y (1994) Interdecadal variations in North Atlantic surface temperature and associated atmospheric conditions. *J Clim* 7:141–157
- Mann ME, Emanuel KA (2006) Atlantic hurricane trends linked to climate change. *Eos* 87:233–244
- Martin ER, Thorncroft C, Booth BBB (2014) The multidecadal Atlantic SST–Sahel rainfall teleconnection in CMIP5 simulations. *J Clim* 27:784–806
- McCarthy GD, Haigh ID, Hirschi JJM, Grist JP, Smeed DA (2015) Ocean impact on decadal Atlantic climate variability revealed by sea-level observations. *Nature* 521:508–510. doi:[10.1038/nature14491](https://doi.org/10.1038/nature14491)
- Murphy LN, Bellomo K, Cane M, Clement A (2017) The role of historical forcings in simulating the observed Atlantic multidecadal oscillation. *Geophys Res Lett* 44:2472–2480. doi:[10.1002/2016GL071337](https://doi.org/10.1002/2016GL071337)
- Nigam S, Guan B, Ruiz-Barradas A (2011) Key role of the Atlantic multidecadal oscillation in 20th century drought and wet periods over the Great Plains. *Geophys Res Lett* 38:L16713. doi:[10.1029/2011GL048650](https://doi.org/10.1029/2011GL048650)

- O'Reilly CH, Huber M, Woollings T, Zanna L (2016) The signature of low-frequency oceanic forcing in the Atlantic multidecadal oscillation. *Geophys Res Lett*. doi:[10.1002/2016GL067925](https://doi.org/10.1002/2016GL067925).
- Otterå OH, Bentsen M, Drange H, Suo L (2010) External forcing as a metronome for Atlantic multidecadal variability. *Nat Geosci* 3:688–694
- Otto-Bliesner BL et al (2016) Climate variability and change since 850 CE: an ensemble approach with the Community Earth System Model. *Bull Am Meteor Soc* 97:735–754. doi:[10.1175/BAMS-D-14-00233.1](https://doi.org/10.1175/BAMS-D-14-00233.1)
- Peings Y, Simpkins G, Magnusdottir G (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. doi:[10.1002/2015JD024107](https://doi.org/10.1002/2015JD024107)
- Sutton RT, Dong B (2012) Atlantic Ocean influence on a shift in European climate in the 1990s. *Nat Geosci* 5:788–792. doi:[10.1038/ngeo1595](https://doi.org/10.1038/ngeo1595).
- Sutton RT, Hodson DLR (2005) Atlantic Ocean forcing of North American and European summer climate. *Science* 309:115–118. doi:[10.1126/science.1109496](https://doi.org/10.1126/science.1109496)
- Tandon NF, Kushner PJ (2015) Does external forcing interfere with the AMOC's influence on North Atlantic sea surface temperature? *J Clim* 28:6309–6323. doi:[10.1175/JCLI-D-14-00664.1](https://doi.org/10.1175/JCLI-D-14-00664.1)
- Terray L (2012) Evidence for multiple drivers of North Atlantic multi-decadal climate variability. *Geophys Res Lett* 39:L19712. doi:[10.1029/2012GL053046](https://doi.org/10.1029/2012GL053046)
- Ting M, Kushnir Y, Seager R, Li C (2009) Forced and internal twentieth-century SST in the North Atlantic. *J Clim* 22:1469–1481
- Ting M, Kushnir Y, Li C (2014) North Atlantic multidecadal SST oscillation: external forcing versus internal variability. *J Mar Syst* 133:27–38
- Trenary L, DelSole T (2016) Does the Atlantic multidecadal oscillation get its predictability from the Atlantic meridional overturning circulation? *J Clim* 29:5267–5280
- Trenberth KE, Shea DJ (2006) Atlantic hurricanes and natural variability in 2005. *Geophys Res Lett* 33:L12704. doi:[10.1029/2006GL026894](https://doi.org/10.1029/2006GL026894)
- 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*. doi:[10.1002/2016JD025979](https://doi.org/10.1002/2016JD025979)
- Wunsch C (1999) The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull Am Meteor Soc* 80:245–255
- Yuan T, Oreopoulos L, Zelinka M, Yu H, Norris JR, Chin M, Plattnick S, Meyer K (2016) Positive low cloud and dust feedbacks amplify tropical North Atlantic multidecadal oscillation. *Geophys Res Lett* 43:1349–1356. doi:[10.1002/2016GL067679](https://doi.org/10.1002/2016GL067679)
- Zhang R, Delworth TL (2006) Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys Res Lett* 33:L17712. doi:[10.1029/2006GL026267](https://doi.org/10.1029/2006GL026267)
- Zhang R, Delworth TL, Sutton R, Hodson D, Dixon KW, Held IM, Kushnir Y, Marshall D, Ming Y, Msadek R, Robson J, Rosati A, Ting M, Vecchi GA, 2013: Have aerosols caused the observed Atlantic multidecadal variability?. *J Atmos Sci*. doi:[10.1175/JAS-D-12-0331.1](https://doi.org/10.1175/JAS-D-12-0331.1)
- Zhang R, Sutton R, Danabasoglu G, Delworth TL, Kim WM, Robson J, Yeager SG (2016) Comment on “The Atlantic multidecadal oscillation without a role for ocean circulation”. *Science*. doi:[10.1126/science.aaf1660](https://doi.org/10.1126/science.aaf1660)