

Different flavors of the Atlantic Multidecadal Variability

Davide Zanchettin · Oliver Bothe ·
Wolfgang Müller · Jürgen Bader ·
Johann H. Jungclaus

Received: 20 August 2012 / Accepted: 8 January 2013 / Published online: 23 January 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract We investigate how differently-constructed indices for North Atlantic sea-surface temperatures (NASSTs) describe the “Atlantic Multidecadal Variability” (AMV) in a suite of unperturbed as well as externally-forced millennial (pre-industrial period) climate simulations. The simulations stem from an ensemble of Earth system models differing in both resolution and complexity. Different criteria exist to construct AMV indices capturing different aspects of the phenomenon. Although all representations of the AMV maintain strong multidecadal variability, they depict different characteristics of simulated low-frequency NASST variability, evolve differently in time and relate to different hemispheric teleconnections. Due to such multifaceted signatures in the ocean-surface as well as in the atmosphere, reconstructions of past AMV may not univocally reproduce multidecadal NASST variability. AMV features under simulated externally-forced pre-industrial climate conditions are not unambiguously distinguishable, within a linear framework, from AMV features in corresponding unperturbed simulations. This prevents a robust diagnosis of the simulated pre-industrial AMV as a predominantly internal rather than externally-forced phenomenon. We conclude that a multi-perspective assessment of multidecadal NASSTs variability is

necessary for understanding the origin of the AMV, its physics and its climatic implications.

Keywords Atlantic Multidecadal Variability · Sea-surface temperatures · Earth system model · Teleconnections · Natural climate variability · Externally-forced climate variability · Climate of the last millennium

1 Introduction

North Atlantic sea-surface temperatures (NASSTs) fluctuate between alternating warm and cold phases as described by countless observational studies (e.g., Folland et al. 1986; Enfield and Mestas-Núñez 1999; Wyatt et al. 2011; Wu et al. 2011b; Liu 2012), reconstructions (e.g., Gray et al. 2004; Mann et al. 2009; Saenger et al. 2009; Chylek et al. 2011; Knudsen et al. 2011) and numerical climate simulations (e.g., Zhang and Delworth 2007; Grosfeld et al. 2008; Ting et al. 2011; Zanchettin et al. 2010, 2012a; Henriksson et al. 2012; Wei and Lohmann 2012). Variations are typically paced at a timescale of 50–90 years. Based on its oscillatory character during the twentieth century, this component of NASST variability has been commonly referred to as “Atlantic Multidecadal Oscillation” or AMO (e.g., Kerr 2000; Enfield et al. 2001). Nowadays, the term “Atlantic Multidecadal Variability” (AMV) is increasingly preferred since the observed signal may not truly represent a physical oscillatory mode of oceanic circulation (Vincze and Jánosi 2011; Liu 2012).

Traditionally, the AMV is thought to describe the low-frequency *internal* contribution to NASST variability (Schlesinger and Ramankutty 1994). The AMV signal is generally extracted by spatially averaging NASSTs after

D. Zanchettin (✉) · O. Bothe · W. Müller · J. Bader ·
J. H. Jungclaus
Max Planck Institute for Meteorology,
Bundesstr. 53, 20146 Hamburg, Germany
e-mail: davide.zanchettin@zmaw.de

O. Bothe
University of Hamburg, KlimaCampus Hamburg,
Hamburg, Germany

J. Bader
Bjerknes Centre for Climate Research,
Uni Research, Bergen, Norway

partitioning them from external signals. Pre-processing usually includes either statistical methods of different complexity (Trenberth and Shea 2006; Knight 2009; Ting et al. 2009; DelSole et al. 2011) or a priori assumptions about the externally-driven portion of NASST variability (see, e.g., Mann and Emanuel 2006). AMV-type features have also been identified, e.g., in the first principal component of detrended NASSTs (e.g., Wu et al. 2011b), in the second principal oscillatory pattern mode of Northern Hemisphere SSTs (Park and Latif 2010), or in a high-order mode of last century's global SSTs pre-processed for global trend partitioning (Dima and Lohmann 2009). Such definitions imply the capability to unambiguously discern the internally-generated and externally-forced/global components of NASST variability. The separation of both contributions to multidecadal SST signals is, however, non-trivial (Dommenget and Latif 2008; Knight 2009), and a simplistic approach may generate artifacts in the AMV index evolution (Zanchettin et al. 2012a).

An ongoing debate concerns the nature of the twentieth century evolution of the AMV, i.e., whether it was indeed due to internal variations rather than predominantly externally-forced (e.g., Ting et al. 2009; Knight 2009; Enfield and Cid-Serrano 2010; Otterå et al. 2010; Wu et al. 2011a; Zanchettin et al. 2012a; Booth et al. 2012). Neither AMV reconstructions (e.g., Gray et al. 2004; Mann et al. 2009; Saenger et al. 2009; Chylek et al. 2011; Knudsen et al. 2011) nor numerical climate simulations (e.g., Enfield and Cid-Serrano 2006; Otterå et al. 2010; Park and Latif 2010; Zanchettin et al. 2010, 2012a, b) provide conclusive arguments about the persistence and stationarity of AMV-type fluctuations in NASSTs over multi-centennial or even millennial timescales, i.e., about the tendency of the climate system to robustly generate such variability. Most recently, significant spatial variability in reconstructed AMV-type fluctuations has been detected during the last several centuries (Chylek et al. 2012). Similarly, although common features characterize the evolutions of reconstructed and ensemble-simulated AMV indices during periods dominated by external forcings (Zanchettin et al. 2012a), their projection on simulated NASSTs for the last ~500 years produces different patterns. In other words, AMV indices emphasizing NASST variability in different portions of the basin can have, at least over limited periods, an indistinguishable temporal signature. This further implies that the coherence of local NASST fluctuations can vary through time.

These findings challenge a univocal description of the AMV. They particularly question its description and interpretation as the multidecadal synchronous wavering of NASSTs as implied by the common spatially-averaged indices. These may not fully represent the basin- or sub-basin-scale decadal and multidecadal variability related to,

e.g., modifications in the wind-driven circulation or sea-ice variability. Indeed, while convergent findings address the North Atlantic thermohaline circulation as of critical importance for setting the preferred timescale(s) of AMV fluctuations, other aspects of AMV generation mechanism(s) remain unclear (Liu 2012). For instance, the North Atlantic Oscillation (NAO) strongly imprints on NASSTs, primarily via modifying the trajectory of the Gulf Stream/North Atlantic Current due to NAO-related meridional shifts in the jet (Marshall et al. 2001).

A more robust statistical interpretation of the AMV assists in a better understanding of the phenomenon's dynamics. It requires assessing whether multiple definitions of the AMV are compatible with prominent multidecadal NASST variability though capturing different aspects of it. This study is motivated by the question: can we better grasp multidecadal NASST variability when faceting it through alternative definitions of the AMV? In the following, we illustrate the distinctive traits of differently-constructed AMV indices evaluated for a set of unperturbed as well as (pre-industrial) externally-forced multicentennial and millennial climate simulations conducted with a suite of Earth system models differing in both resolution and complexity. Focusing on statistics but keeping a glance at dynamical implications, we evaluate, in a first step, how differently-constructed indices of the AMV capture different aspects of simulated multidecadal NASST variability and of its hemispheric atmospheric teleconnections. Mis-representation or under-representation of relevant NASST variability is going to influence our understanding of past climate variability. Thus, we explore, in a second step, how the different AMV expressions may affect our interpretation of reconstructed past climate variability.

2 Data and methods

Table 1 lists the climate simulations used in this study. They differ in the employed model, in the resolution and complexity of its atmospheric, oceanic and land components, in the imposed forcing/boundary conditions and in the simulated climate's background state. We consider four state-of-the-art Earth system models (CCSM4, GISS-E2-R, MPI-ESM-P/-MR) contributing to the third phase of the Paleoclimate Modelling Intercomparison Project (PMIP3). We additionally include simulations from a previous-generation model (COSMOS-Mil) for comparison. The latter includes not only modules for land and ocean biogeochemistry but also a fully-interactive carbon cycle with prognostic CO₂ calculation in the atmosphere. All other models apply prescribed atmospheric CO₂. We do not expect the implemented biogeochemistry and/or interactive

Table 1 Characteristics of the Earth system model simulations used in this study

Model (acronym)	Atm/oce components (resolution)	Unperturbed simulation (reference period) [trend]	Forced simulation(s) [trend]	Reference/sources
Max Planck Institute COSMOS Earth System Model (COSMOS-Mil)	ECHAM5 (T31L19)/MPIOM(GR30L40, horiz. res: 20–350 km)	mi10001 (800–1,799) [0]	mi100010 [1]	Jungclauss et al. (2010)
Max Planck Institute Earth System Model (MPI-ESM-P)	ECHAM6 (T63L47)/MPIOM(GR15L40, horiz. res: 12–180 km)	piControl-r1i1p1-P (1,850–2,849) [0]	past1000-r1i1p1 [1]	Giorgetta et al. (2012), Jungclauss et al. (2012)
Max Planck Institute Earth System Model (MPI-ESM-MR)	ECHAM6 (T63L95)/MPIOM(TP04L40, horiz. res: ~0.4° globally)	piControl-r1i1p1-MR (1,850–2,849) [0]	–	Giorgetta et al. (2012), Jungclauss et al. (2012)
NCAR Community Climate System Model 4 (CCSM4)	CAM4 (1.25° × 0.9°L26)/Parallel Ocean Model 2 (1°L60)	piControl-r1i1p1 (800–1,300) [0]	–	Gent et al. (2011), Landrum et al. (2011)
NASA GISS-E2-R (GISS-E2-R)	ModelE (2° × 2.5°L40)/Russell (1° × 1.25°L32)	piControl-r1i1p141 (1,112–2,011) [1]	past1000_r1i1p124 [2] past1000_r1i1p125 [1]	http://data.giss.nasa.gov/modelE/ar5/

Columns: (1) Model, with the acronym used in the text in brackets; (2) atmospheric and oceanic components, with resolution in brackets; (3) name of the unperturbed simulation with considered reference period (in brackets) and degree of the polynomial trend removed for the AMV indices' calculations (in squared brackets); (4) same as (3), but for the forced simulation; (5) references and/or web sources providing details on the models and the simulations

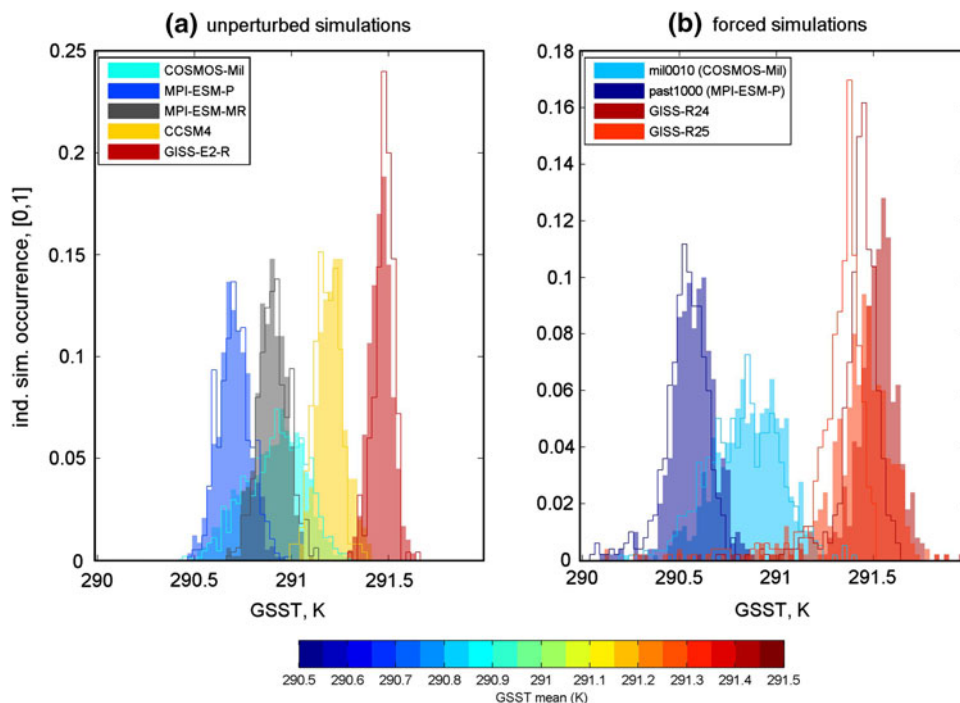
carbon cycle to substantially affect oceanic and atmospheric physics and variability at the investigated timescales, though tight climate-biogeochemical cycle interactions are diagnosed in Earth-system-model simulations following episodes of strong forcing (e.g., Brovkin et al. 2010).

For a detailed description of individual simulations, the reader is referred to the references listed in Table 1. Here, we outline the most relevant distinguishing characters. The top-of-atmosphere is set at different geopotential heights, ranging from 0.01 hPa in MPI-ESM-P/-MR to 10 hPa in COSMOS-Mil. High-top models allow for a more realistic representation of stratospheric processes and of stratospheric/tropospheric interactions, compared to low-top models that only partly resolve the stratosphere and stratospheric/tropospheric interactions (e.g., Manzini et al. 2012). COSMOS-Mil and MPI-ESM-P/-MR share the same ocean model (MPIOM), but in different resolutions. Magnitude and patterns of model biases against observations of surface temperature and salinity in MPIOM-based coupled models highlights the ocean's grid configuration rather than the coupled atmosphere component (Jungclauss et al. 2012), which adds interest to our comparative assessment of the -P and -MR versions of MPI-ESM. Our ensemble therefore includes models from different families as well as models from the same family but in different configurations. We believe that our selection, though subjective, is effective in spanning the range of AMV features simulated by available climate models.

Our evaluation includes unperturbed simulations as well as a selection of forced simulations for the last millennium (pre-industrial era only). Unperturbed simulations differ, among other aspects, in the imposed constant boundary conditions throughout the simulation (e.g., the consideration of a background volcanic-aerosols forcing or the reference year for the orbital forcing and greenhouse gases concentrations) and in the implementation of a module for dynamical vegetation. Forced simulations differ in the imposed forcings with respect to their numerical implementation and the input data. The two forced GISS-E2-R simulations (hereafter: GISS-R24 and GISS-R25), for instance, use the volcanic forcing reconstructions by Crowley et al. (2008) (see also: Crowley and Unterman 2012) and Gao et al. (2008), respectively. The solar forcing in the forced simulations is based on reconstructions of total solar irradiance exhibiting a total increase of 0.1 % (~1.3 Wm⁻²) from the Maunder Minimum (1647–1715 AD) to present-day (Vieira et al. 2011, see also: Jungclauss et al. 2010). Schmidt et al. (2011) and Fernández-Donado et al. (2012) provide an overview on forcing estimates used, respectively, in PMIP3 and pre-PMIP3 simulations.

Figure 1 illustrates the empirical probability distributions of the global-average SST (GSST) for half-periods of

Fig. 1 Empirical probability distributions of global-average SST (GSST) for the investigated simulations (**a** unperturbed, **b** forced). Filled contours (*lines*) refer to the first (*latter*) half of each simulation. Colors reflect the full-period GSST mean, according to the *color bar* on the bottom, except for MPI-ESM-MR in **a** (set to *grey* for distinguishability). SST data are not preprocessed for trend removal before the GSST calculation



the individual investigated simulations. Half-periods are taken from all considered integration years in the unperturbed simulations (up to 1,000 years, see Table 1), and are the AD 850–1349 and AD 1350–1849 periods in the forced millennium simulations. Stationary GSST is generally detected in unperturbed simulations (compare half-periods in Fig. 1a), with the larger range for COSMOS-Mil likely due to the overly strong interannual variability originated in the tropical Pacific SSTs (Jungclaus et al. 2010; Zanchettin et al. 2012a). We can expect the forced climate variability to differ only slightly from internally-generated variability, due to the characteristics of the imposed forcings. This is, however, not necessarily the case, as pre-industrial forced climates within a simulation-ensemble can individually be statistically distinguishable from the corresponding unperturbed climate (concerning COSMOS-Mil, see: Bothe et al. 2012). In our ensemble, not only the following features differentiate GSST in forced (pre-industrial) and unperturbed simulations: (i) a millennial cooling trend (compare half-periods in Fig. 1b) due to orbital forcing and slow oceanic adjustment to background cooling from volcanic aerosols; (ii) cold temperature excursions due to strong volcanic eruptions, which can leave a detectable trace on multidecadal timescales (e.g., Jungclaus et al. 2010; Otterå et al. 2010); (iii) empirical distributions spanning a wider range of values, especially for GISS-R24/-R25; (iv) appreciably colder average conditions in the past1000 simulation of MPI-ESM-P compared to its unperturbed counterpart. We may therefore expect the imposed forcing, though generally weak, to affect the simulated AMV by directly

modulating the NASST decadal evolution and by setting the background climate conditions, which can crucially influence simulated processes in the North Atlantic Ocean (e.g., Yoshimori et al. 2010).

The following standardized AMV indices are evaluated based on annual-average NASSTs covering the domain [90 W–10E; 0–80 N] but excluding the interior of the Labrador Sea as in Zanchettin et al. (2012a): (i) AMV1 as spatially-averaged NASSTs; (ii) AMV2 as spatially-averaged Δ NASSTs, where Δ NASSTs indicate local NASST anomalies from the GSST; (iii) AMV3 as the corresponding principal component of NASSTs' first EOF evaluated using an area-weighted covariance matrix. The sign of the AMV3 index is chosen as to be positively correlated with the other indices; the sign of AMV3 regression patterns for ensemble mapping is chosen as to produce positive spatial correlations between the different models (this particularly concerns GISS-E2-R). The ocean data from COSMOS-Mil and MPI-ESM-P/-MR are regridded to a regular $1^\circ \times 1^\circ$ grid before the analysis. The long-term trend component in SSTs differs among the considered simulations due to millennial-scale forcing and/or due to slow convergence towards stationary states (compare half-period distributions of GSST in Fig. 1). In order to align them, a local full-period trend component is therefore removed from the NASST and Δ NASST fields prior to the AMV indices' calculations (see Table 1 for details). We acknowledge that in doing so we neglect the effect of possible interdependencies between millennial- and multidecadal-to-centennial-scale NASST variability and alter the characteristics of AMV indices (e.g., their

autocorrelation profiles). Anyway, the trend components generally account for a small fraction of NASST local variability (not shown), and trend removal allows for a more reliable correlation-based assessment of O(10–100 years) NASST features. For each index the explained fraction of total NASST variability is calculated as the grid-area-weighted sum of locally explained variances based on the local AMV-SST squared correlation coefficient.

The selected AMV indices and models/simulations are only a subset of those available and/or described in the literature. A more comprehensive investigation is not required since we aim to demonstrate that multidecadal NASST variability is captured by differently-defined AMV indices and to discuss possible implications for their statistical as well as dynamical interpretation. We do not intend to describe all the possible manifestations of the AMV, nor do we claim to individuate the best AMV descriptor or, similarly, the model with the best skills in simulating its observed characteristics. We further acknowledge that unperturbed simulations alone are known non-optimal tools for clarifying details of internally-generated AMV dynamics (e.g., Liu 2012). Therefore, we avoid attempting to distinguish specific feedbacks and mechanisms (e.g., does the AMV predominantly result from a coupled mode or from a damped oceanic mode?). Rather, we rely on statistical analyses to infer congruency between AMV representations and (aspects of its) dynamics in unperturbed and forced simulations.

The scaling between AMV indices evaluated for the unperturbed and for the corresponding forced simulations is assessed by (i) projecting the NASST-AMV regression pattern for the unperturbed simulation (scaled for the relative local standard deviations and grid-area-weighted) on the forced NASSTs and (ii) linearly regressing the obtained standardized index on the corresponding forced AMV index. The scaling and the standardization make local amplification/dampening of forced versus unperturbed signals marginal for this exercise, whereas changes in the sign of local NASST-AMV correlations leading to a disruption of the regression pattern are highlighted. Local amplification/dampening due to external forcing is checked against internally-generated amplification/dampening by comparing the scaling patterns obtained over half-periods of the forced and the corresponding unperturbed simulations. More specifically, local NASST-AMV regression coefficients r from two simulations/half-periods are used to construct a scaling coefficient $\alpha = \text{sign}(r_1/r_2) \cdot \max(|r_1/r_2|, |r_2/r_1|)$. Local amplification/dampening is attributed to external forcing if all the α values for both forced half-periods versus both unperturbed half-periods (i.e., four values) differ in the sign and/or exceed the α value calculated between the two unperturbed half-periods. Half-periods are defined as above. For consistency, AMV

indices are re-calculated over the half-periods and may therefore differ from the full-period counterparts.

Reconstructions of different AMV indices are tested following Gray et al. (2004). In a 1,000-sample Monte-Carlo iteration, ten locations are randomly sampled over the land-only domain within [120 W–60E; 0–80 N], and the leading four principal components of associated 11 years smoothed annual surface (2 m) air temperatures (SATs) are used as predictors (plus an intercept coefficients) in a multiple linear regression model for each of the considered AMV indices. Coefficient of determination (R^2 as r-square statistics for the full regression model) and coefficient of error (CE) (Cook et al. 1994) are used as skill metrics within a cross-calibration-verification procedure. CE is defined as $[1.00 - \sum_{t=1:N} (x_t - y_t)^2 / \sum_{t=1:N} (x_t - x_{mv})^2]$, where x_t and y_t are, respectively, the observed and the predicted index in year t , x_{mv} is the observed mean index over the validation period and N is the number of years in the validation period. Two subsequent half-periods (defined as above) are considered as calibration/validation and validation/calibration periods. We only show statistics for the validation periods. We also assess the robustness of the reconstruction against noise. To this purpose, for each iteration, 50 % of the selected SAT series are substituted with simulated autocorrelated processes whose parameters are estimated from the substituted original series (Neumaier and Schneider 2001). In particular, the order is subjectively set to be equal to the lag for which the autocorrelation of the substituted original series falls below the threshold of $1/e$.

3 Results

3.1 Unperturbed AMV variability and patterns

Figure 2 illustrates the temporal evolutions and spectral densities of the differently defined AMV indices for the selected unperturbed simulations. There is general agreement between the evolutions of AMV1 and AMV2 (see also the correlations in Fig. 2a). In COSMOS-Mil, MPI-ESM-MR and CCSM4, AMV3 shares less than one-third of the total variability with AMV1 and AMV2. A test of the strength of the correlations in COSMOS-Mil based on bootstrapping the decadal-smoothed AMV3 spectral phases (e.g., Zanchettin et al. 2008) reveals that a random process described by the AMV3 spectrum would correlate with AMV1 and AMV2 with an upper 99 %-confidence R^2 of 0.09 and 0.1, respectively. Hence, the variability expressed by AMV3 in COSMOS-Mil is very likely (linearly) independent from that expressed by AMV1 and AMV2. Such a separation between AMV3 and AMV1/AMV2 is less evident in the other models and lacks in

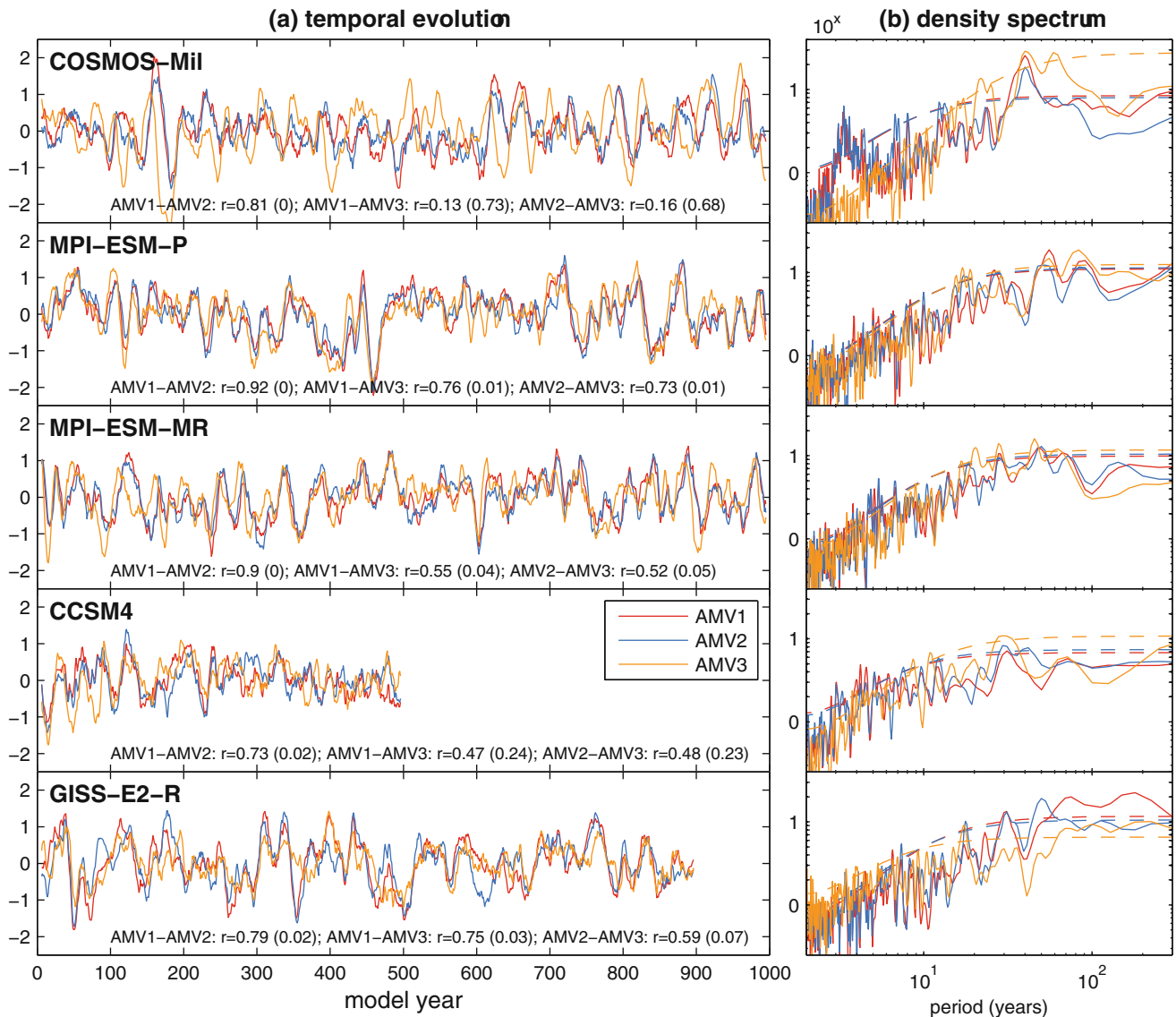


Fig. 2 Time series (a, 11 years moving average smoothing) and density spectra (b, linearly detrended data) of AMV indices for the investigated unperturbed simulations. Pearson's correlation coefficients (r) between the indices are reported in a with associated

p -levels in brackets (accounting for autocorrelation in the series). Hatched lines in b are 95 % confidence level for a random red-noise process with lag-1 autocorrelation equal to the corresponding AMV series

GISS-E2-R, where correlations among AMV indices are comparable.

All indices in all models display prominent multidecadal variability (Fig. 2b), but the spectra differ in their exact peak frequencies. These range from prominently interdecadal as, e.g., in CCSM4 (but note the shorter integration time) to near-centennial as, e.g., AMV1 in GISS-E2-R. Within the O(50–90 years) band, spectral power is most comparable among indices in MPI-ESM-P/-MR. Conversely, within this band, AMV indices peak at most notably different frequencies in GISS-E2-R, highlighting the ambiguity inherent in the assessment of AMV's "typical" fluctuation frequencies.

Figure 3 illustrates the ensemble regression patterns of the global SST field on the different AMV indices, by means of mapped local ensemble averages (top) and standard deviations (bottom). There is strong inter-model agreement in the general characteristics of the NASST-AMV patterns for AMV1 and AMV2, but not as much for AMV3 and for the AMV1 and AMV2 signatures on SSTs outside the North Atlantic. Warm AMV1 phases robustly correspond to warm SST anomalies spreading over almost the entire North Atlantic, with maxima in the tropical band spreading northward along the eastern boundary. AMV1 imprints robustly and positively on equatorial Pacific and tropical Indian Ocean SSTs, with individual simulations

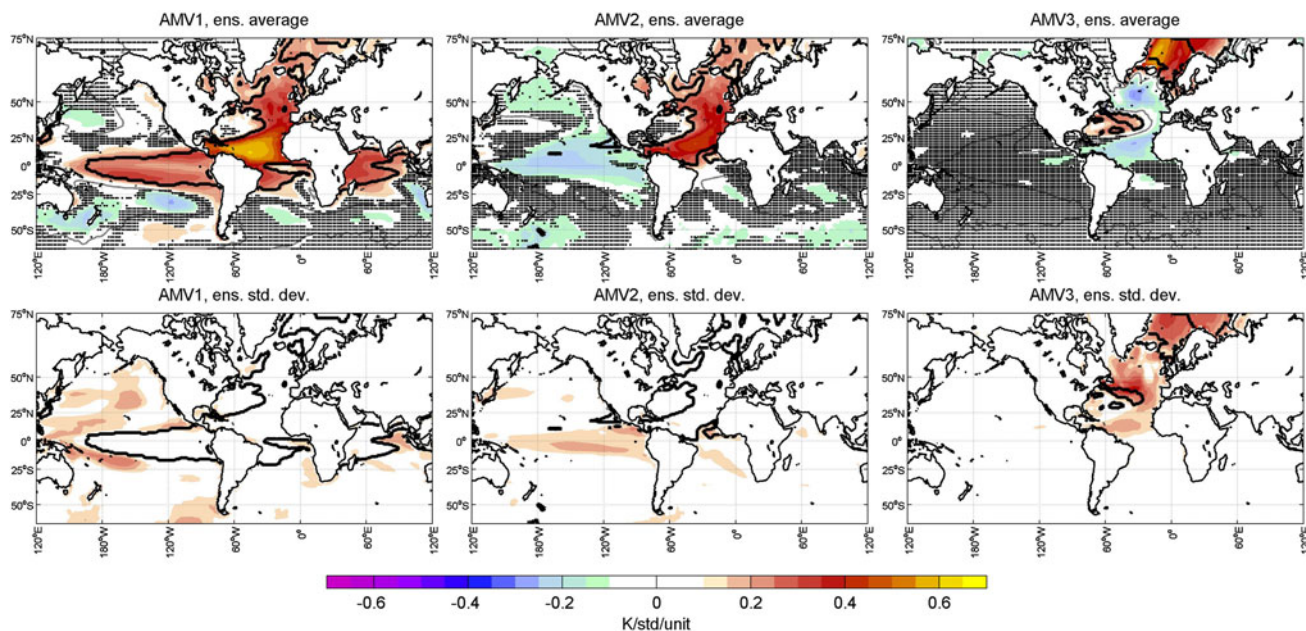


Fig. 3 Ensemble regression patterns of standardized global SSTs on the AMV indices for the investigated unperturbed simulations. *Top* ensemble average; *bottom* ensemble standard deviation. *Thick contours (black dots)* in all (*top*) panels individuate regions where

agreeing less on the AMV1 signature at the margins of the Pacific warm pool region. In some models at least the extensive signature of the AMV1 over the tropical Pacific reflects the dominant global character of the El Niño–Southern Oscillation (ENSO) (see, e.g., Zanchettin et al. 2012a, concerning COSMOS-Mil). A similar, inter-model coherent pattern characterizes the AMV2 signature on NASSTs, with a generally weaker tropical signature than AMV1. Outside the North Atlantic, the AMV2 signature is generally weak and/or less robust compared to AMV1, see, e.g., the negative imprint some simulations produce over the equatorial Pacific. As seen in individual simulations (not shown), the NASST-AMV3 pattern generally consists of a tri-polar structure with two same-sign centers located in the subtropical western North Atlantic and in the Nordic Seas, and one opposite-sign center located in the subpolar North Atlantic. This structure is largely smeared out in the ensemble average pattern due to the different location of the centers in the individual simulations. Large spread in the ensemble signature of the index is detected within the North Atlantic, especially along the Gulf Stream trajectory. AMV3 thus possibly reflects the different representations of ocean boundary currents and gyre circulation, and the decadal SST variability they generate locally (e.g. Liu 2012). AMV3 seems to describe especially wind-driven and sea ice-related features of the upper North Atlantic circulation, preferably capturing mid- and high-latitude variability. Therefore, it generally explains a smaller fraction of NASST variability (generally well below 10 %)

the regression is significant (non-significant) at $p < 0.05$ accounting for autocorrelation in the series in all (at least three-fourth of) the simulations. *Maps* are created by homogenizing the original regression patterns on a regular $1^\circ \times 1^\circ$ grid

than the other indices, in our definition and in these models. This is however not true for GISS-E2-R, where the AMV3 pattern highlights variability in the subpolar gyre and tropical regions (individual simulation not shown). In this model, the strong signature of AMV3 in the subpolar gyre region and its weak signature in the Nordic Seas suggest weaker sea-ice variability and/or a weaker influence of sea-ice variability on this mode compared to the other models, where the correlation is particularly strong along the trajectory of the Eastern Greenland current.

Figure 4 illustrates the ensemble regression patterns of Northern Hemisphere winter (DJF) sea-level pressure (SLP) on the different AMV indices. AMV1 imprints hemispherically on DJF SLPs, with the ensemble consistently pointing to a seesaw between SLPs over the mid-latitude North Atlantic and the eastern tropical and subtropical Pacific oceans, and SLPs over the tropical Indian and the western tropical Pacific oceans. The hemispheric signature is particularly strong in CCSM4 and COSMOS-Mil (individual simulations not shown), possibly reflecting interferences by ENSO (e.g., Deser et al. 2012, concerning CCSM4). Regions of significant AMV1-SLP correlation often correspond to regions of largest ensemble spread, as in the case of the tropical western Atlantic Ocean. The consistent AMV2 imprint on SLP is similar to the AMV1 signature over the North Atlantic, but consistently significant AMV2-SLP correlations lack over remote regions. The ensemble spread over the North Atlantic is appreciably reduced compared to AMV1, particularly in the equatorial

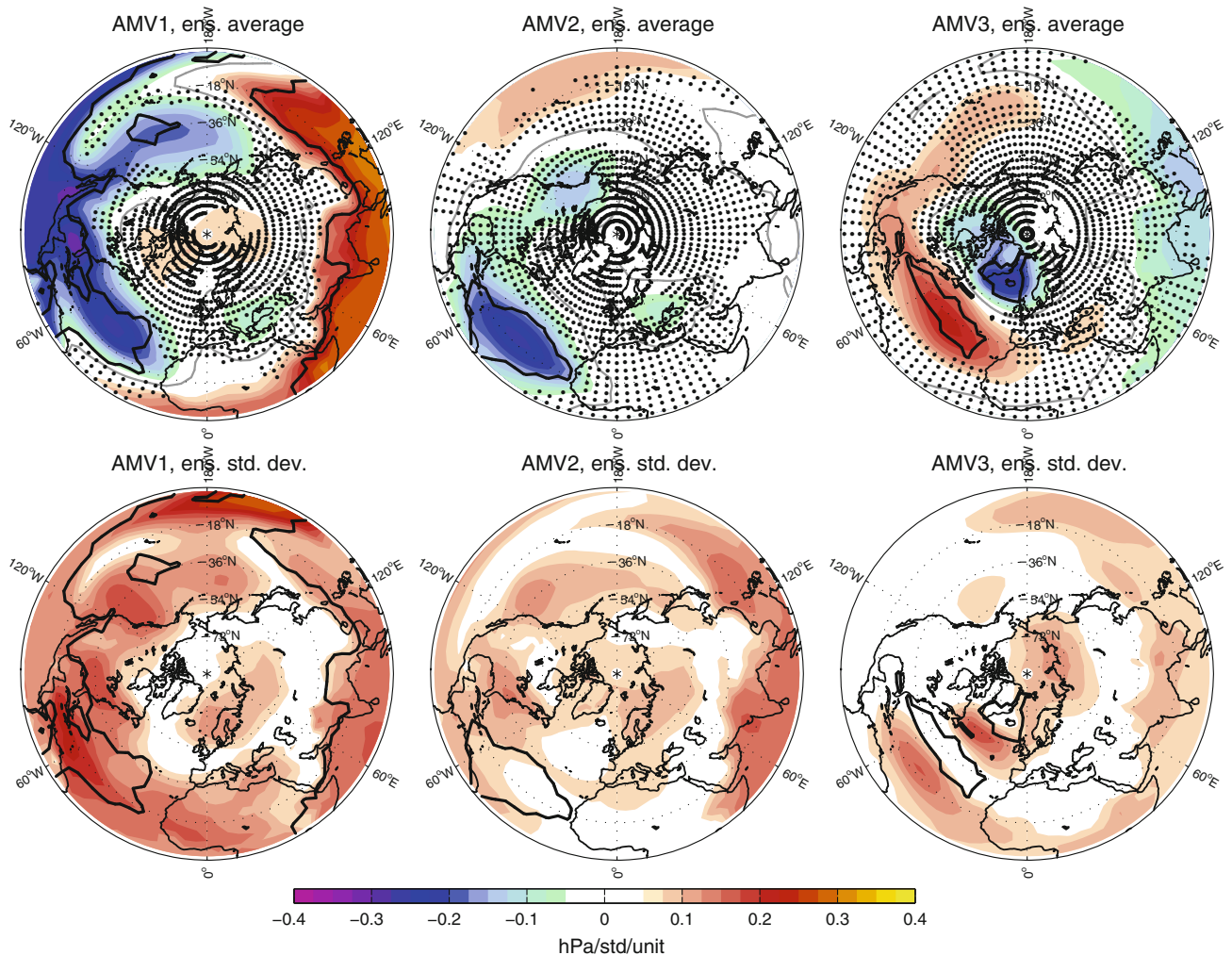


Fig. 4 Ensemble regression patterns of standardized Northern Hemispheric winter (DJF) sea level pressure and AMV indices for the investigated unperturbed simulations. *Top* ensemble average; *bottom* ensemble standard deviation. *Thick contours (black dots)* in all (*top*)

panels individuate regions where the regression is significant (non-significant) at $p < 0.05$ accounting for autocorrelation in the series in three-fourth of the simulations. *Maps* are created by homogenizing the original regression patterns on the COSMOS-Mil T31 grid

to subtropical band. The AMV3 signature on DJF SLPs is local and corresponds to a positive NAO pattern. The robustness of the signature is highlighted by the small ensemble spread over the NAO's two centers of action.

Figure 5 similarly illustrates the ensemble regression patterns of Northern Hemisphere winter upper tropospheric zonal circulation on the different AMV indices. A hemispheric signature characterizes the correlation pattern between AMV1 and the winter 250 hPa zonal wind (U_{250}) corresponding to a circumglobal southward shift of the subtropical jet under positive AMV1 phases. The pattern is especially strong in COSMOS-Mil and CCSM4 while correlations are only locally significant in GISS-E2-R (individual simulations not shown), which results in a generally large ensemble spread over remote regions. The AMV2 ensemble signature on DJF U_{250} is weak, meaning that atmospheric variability related to AMV2 is confined to

the low tropospheric levels. The AMV3- U_{250} regression pattern emphasizes modifications of the jet straddling the North Atlantic. The strong, inter-model consistent signature of AMV3 over the North Atlantic sector contrasts with the weaker and much less consistent signature at the hemispheric scale, highlighting again the regional character of this mode. Lag-correlations between AMV indices and modes of atmospheric variability indicate a strong positive lag (0) AMV3-NAO correlation in COSMOS-Mil and MPI-ESM-P/-MR (not shown). The distinctive response to the atmospheric forcing may explain why large AMV3 fluctuations often appear to anticipate the corresponding ones in AMV1 and AMV2 (see Fig. 2a). In GISS-E2-R, contrasting with the weak U_{250} signature of AMV1 and AMV2, AMV3- U_{250} correlations are particularly strong and suggest a clearer separation between the polar and the subtropical jet over the North Atlantic Ocean

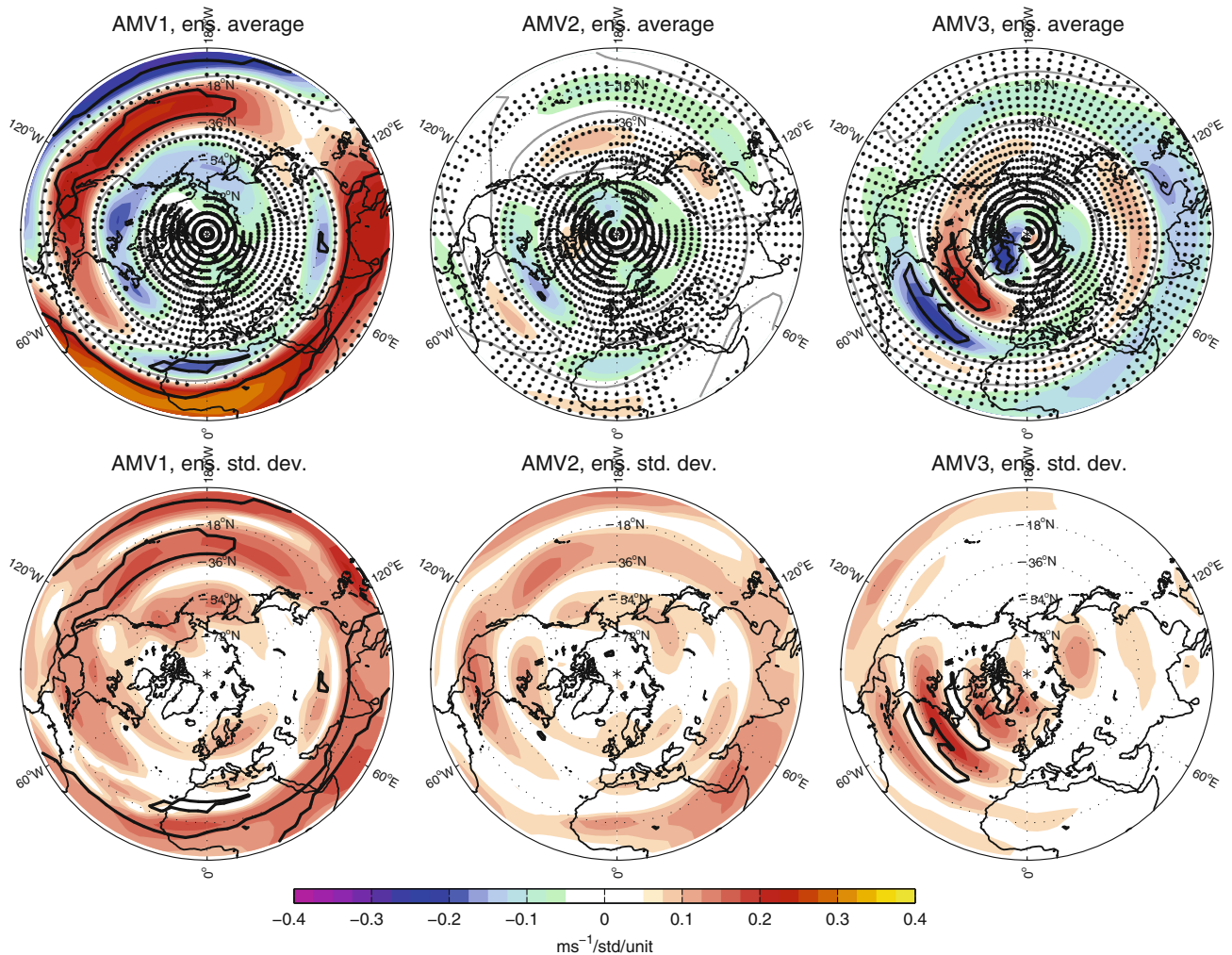


Fig. 5 Ensemble regression patterns of standardized Northern Hemispheric winter (DJF) 250 hPa zonal wind and AMV indices for the investigated unperturbed simulations. *Top* ensemble average; *bottom* ensemble standard deviation. *Thick contours (black dots)* in all (*top*)

under positive AMV3, and vice versa (individual simulation not shown). The AMV3–NAO correlation in CCSM4 and GISS-E2-R is not as strong as in the other models (not shown). This highlights the model-dependent features of the separation into modes of large-scale atmospheric circulation over the North Atlantic/Europe and their coupling to the oceanic circulation.

We shortly recapture, that AMV1 and AMV2 both display a strong basin wide signal though it is weaker for AMV2 in the tropical domain; AMV1 extends its imprint on tropical SSTs outside the Atlantic basin, reflecting the global character of the index. AMV3, on the other hand, generally explains a smaller fraction of NASST variability and appears to relate primarily to the wind-driven component of NASST variability. Teleconnections usually extend throughout the hemisphere for AMV1 but not for

panels individuate regions where the regression is significant (non-significant) at $p < 0.05$ accounting for autocorrelation in the series in three/fourth of the simulations. *Maps* are created by homogenizing the original regression patterns on the COSMOS-Mil T31 grid

AMV2. The AMV3 atmospheric signature extends from the surface to the upper troposphere and prominently features the zonal wind structure over the North Atlantic.

3.2 Relation of AMV variability under forced and unperturbed conditions

Forced (pre-industrial period) simulations confirm the general agreement and the distinguishing details between the temporal evolutions of AMV indices (Fig. 6). Among the considered simulations, AMV1 and AMV2 share (linearly) at most 65 % of the variability. The fraction of shared variability drops to about 11 % between AMV2 and AMV3 in COSMOS-Mil. On the other hand, the temporal evolutions of AMV1 and AMV3 are practically indistinguishable in the GISS-E2-R simulations. The imprint of

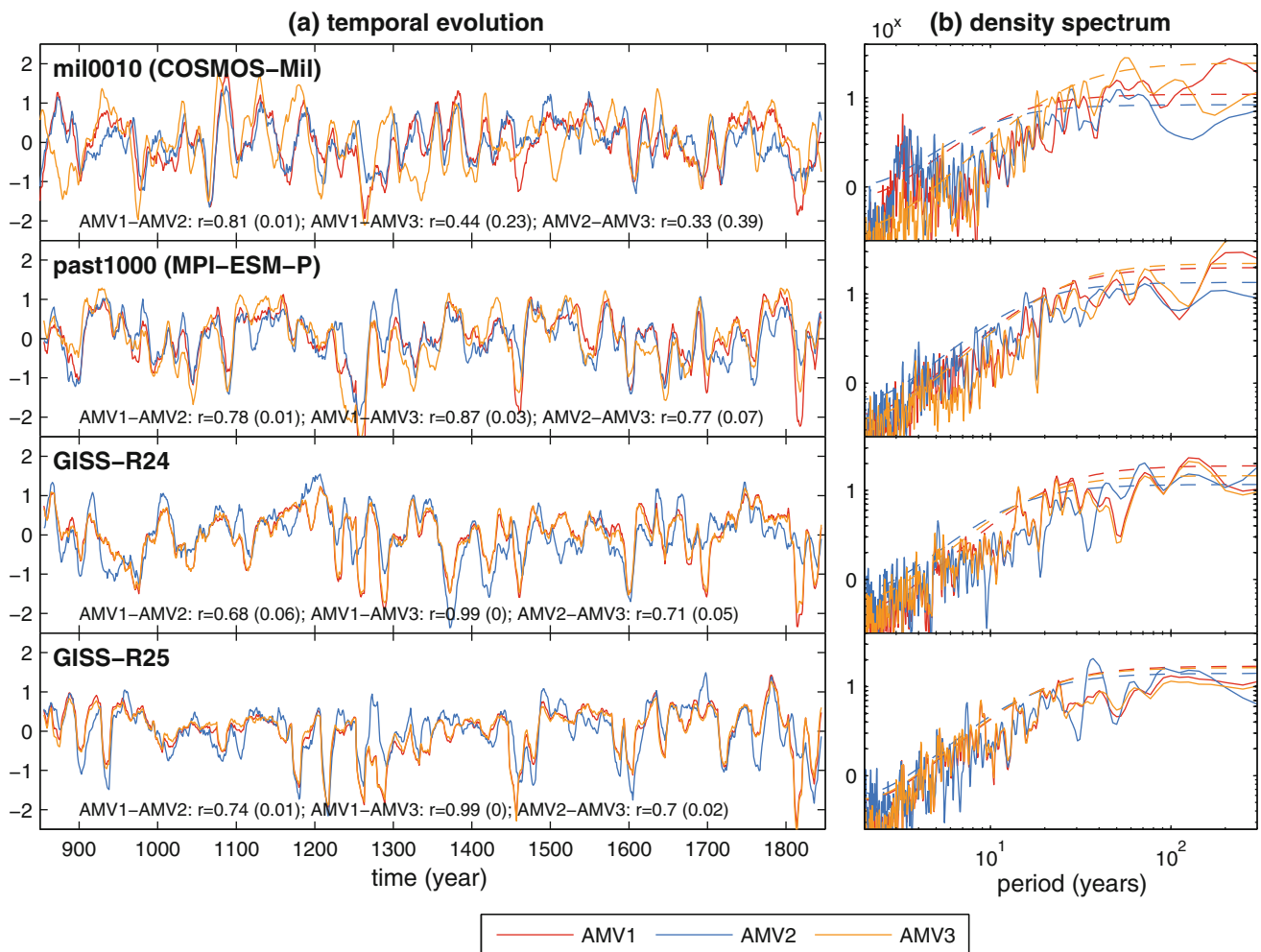


Fig. 6 Time series (**a** 11 years moving average smoothing) and density spectra (**b**) of AMV indices for the investigated forced simulations. Pearson's correlation coefficients (r) between the indices are reported in *panel a*) with associated p-levels in *brackets*

tropical strong volcanic eruptions, especially those clustered in the XIII, late XVI and early XIX centuries (see: Crowley et al. 2008; Gao et al. 2008), is discernible in all models and indices. Volcanically-driven fluctuations often have the weakest amplitude in AMV2. In COSMOS-Mil and MPI-ESM-P, multidecadal variability remains prominent together with emerging strong bicentennial variability, especially for AMV1 (Fig. 6b). In GISS-E2-R, fluctuations in the low-frequency portion of the multidecadal band contribute significantly to AMV2 (but with different peak frequencies in the two simulations) but less so to AMV1 and AMV3.

In our assessment of how much external-forcing contributes to shaping the pre-industrial AMV we compare, in a first step and for the different AMV definitions, the spatial characteristics of the AMV pattern in unperturbed and corresponding forced simulations. Fig. 7 illustrates how the forced NASST-AMV regression maps scale to the

(accounting for autocorrelation in the series). *Hatched lines* in **b** are 95 % confidence level for a random red-noise process with lag-1 autocorrelation equal to the corresponding AMV series

corresponding unforced ones, highlighting regions where strong amplification and strong dampening of the AMV signature occur. The dominance of red and blue tones, which individuate respectively amplification and inverse correlation together with amplification, reflects the generally larger SST variance per AMV unit in the forced simulations. Dots in Fig. 7 individuate regions where we cannot unambiguously attribute the local amplification/dampening to external forcing, i.e., where it is compatible with internal variability and reflects sampling. Most common forced features for AMV1 are the stronger imprints over the equatorial/tropical North Atlantic and, to a lesser extent, over the western part of the North Atlantic storm track region (which is locally ambiguous in MPI-ESM-P and lacks in COSMOS-Mil). The individual patterns hint towards a weaker imprint in the subpolar North Atlantic, but this feature lacks in MPI-ESM-P and is ambiguous in GISS-R24. The AMV2 scaling map is

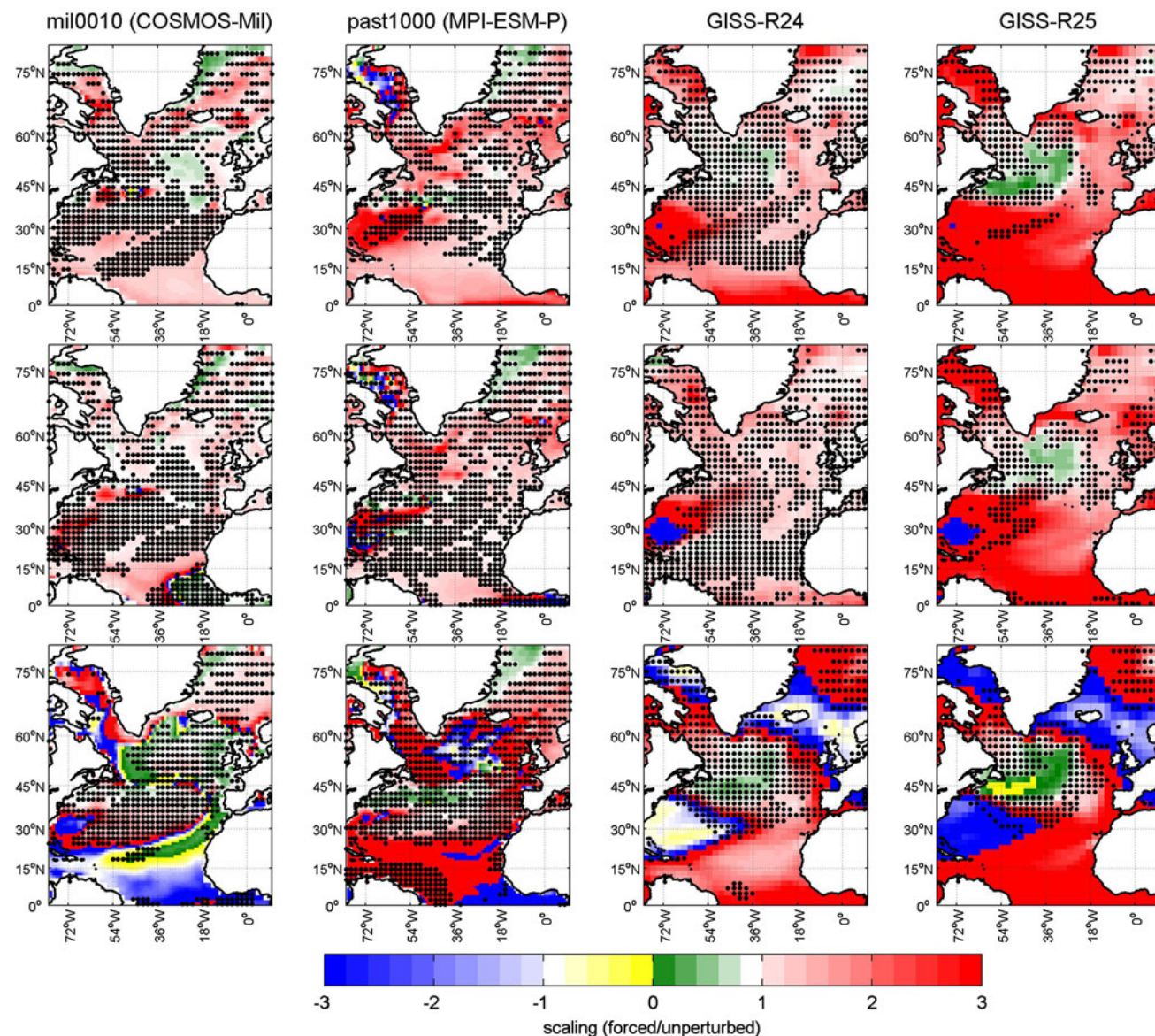


Fig. 7 Ratio between the forced and the corresponding unperturbed local regressions of NASSTs on AMV indices (*top* AMV1; *middle* AMV2; *bottom* AMV3) for the different simulations (full-period analysis). *Black dots* individuate regions where the local amplification/

dampening due to external forcing is compatible with amplification/dampening due to internal variability (based on half-periods, see “Methods” section)

overall patchier than for AMV1, reflecting that this mode captures internal variations. NASST sensitivity to AMV2 scales differently over most regions in the two GISS-E2-R simulations. This feature can be interpreted as a consequence of the applied different forcing, but the interplay between externally-forced and internally-generated AMV variability may also contribute to the differences between the two simulations. In all models, profound differences are revealed in the way AMV3 captures NASST and gyre variability in unperturbed and forced simulations (Fig. 7, bottom panels). The presence of regions featuring

amplified negative scaling (blue) and strong dampening (yellow and green) indicates a disruption of the unperturbed AMV3 pattern (compare with Fig. 3) in the corresponding forced simulations. Such structural modifications may reflect changes in the covariance structure of NASSTs linked to different processes/phenomena dominating NASST variability and modifications to the same dominant processes/phenomena (e.g., changes in the position and intensity of the mid-latitude westerlies). The large extent of dotted regions prevents, at least in some simulations, to attribute the disruption of the AMV pattern

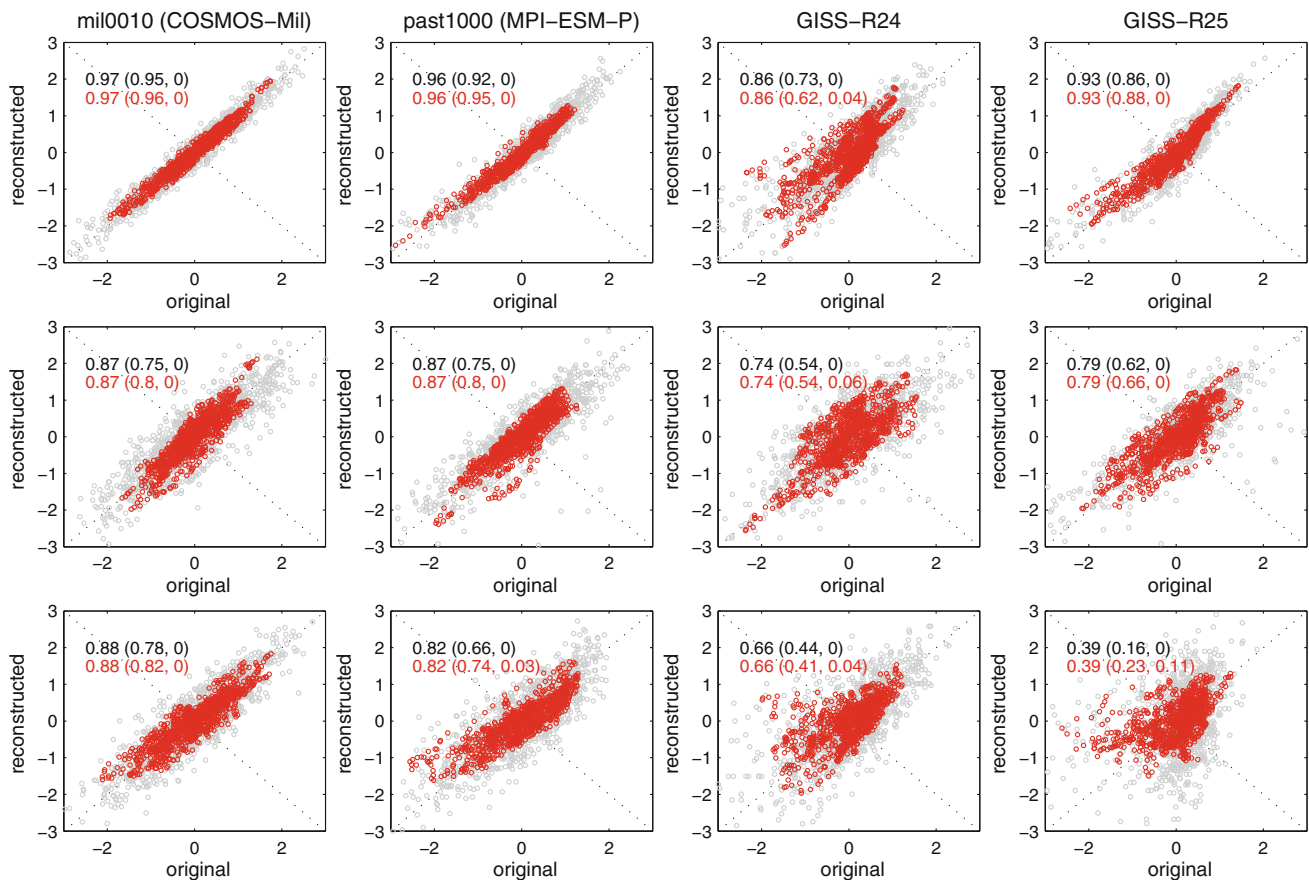


Fig. 8 Scatterplots of AMV indices for the forced simulations (x axis) versus the corresponding standardized AMV indices reconstructed by projecting the unperturbed NASSTs-AMV regression map on the NASST data for the forced simulations (y axis). *Top* AMV1;

middle AMV2; *bottom* AMV3. Grey annual data; red 11 years moving average smoothed data. Reported numbers are linear regression coefficients and, in brackets, R^2 with associated p value

to the presence of external forcing. The fact that the scaling pattern is simulation-specific (compare, for instance, the scaling over the western tropical North Atlantic in GISS-R24 and GISS-R25) rather confirms that, as noted for AMV2, each combination of forcing and ongoing internal variability imprints uniquely on AMV3 characteristics. Not surprisingly, and contrasting AMV1 and AMV2 features, AMV3 features evaluated for half-periods differ from the corresponding full-period ones (e.g., time series share only about 2/3 of the total variance, not shown). This reflects the continuous modifications in the covariance structure of NASSTs that can help disentangling the competition between the multiple processes/mechanisms implicated in the AMV.

3.3 Implications for AMV reproducibility and dynamical interpretation

The scaling maps in Fig. 7 indicate that, at least for AMV1 and AMV2, the difference between unperturbed and corresponding forced patterns concerns mostly the local

amplitude of the signal. This means, the unperturbed pattern seems to contain enough information to accurately reconstruct the evolution of the corresponding forced AMV index by projecting the unperturbed AMV pattern on the corresponding forced NASSTs. Fig. 8 summarizes the results of such reconstructions following the approach described in the methods section. Results are at least encouraging for AMV1 and reach a peak of 96% of recovered decadal AMV1 variability in COSMOS-Mil. Reconstructions remain skillful in all cases for AMV2 recovering at least half of the decadal variability. Conversely, the reconstruction of AMV3 can be deceptive (as for GISS-E2-R simulations) and is more simulation-dependent (compare especially GISS-R24 and GISS-R25). We remark that, as noted above for half-period index correlations, AMV3 reproducibility via linear scaling is inherently limited by continuous modifications in the covariance structure of NASSTs. Hence, we would expect analogous results if the same exercise was performed on an unperturbed simulation by sampling data over different sub-periods.

Nonetheless, since the tails of the scatterplots in Fig. 8 do not consistently deviate from the 45° line, the linear scaling hypothesis seems to be viable even under strong forcing events such as strong volcanic eruptions. The latter are, however, sources of low-frequency ocean variability entailing modifications to the natural interactions between NASSTs, the strength of the thermohaline circulation and the associated northward ocean heat transport (Zanchettin et al. 2012b). Therefore, a simple linear scaling between unperturbed and forced patterns may be insufficient for capturing complex forced interactions, despite allowing for an overall satisfactory reproduction of forced AMV evolutions. This is exemplified by the time-varying correlations between the AMV and the ocean heat transport through the Iceland Scotland Ridge (Q_{ISR}) simulated under unperturbed and forced conditions by MPI-ESM-P (Fig. 9). Q_{ISR} reflects the complexity of the simulated interactions between gyre and overturning components of ocean heat transport in the North Atlantic, and play a key role in the simulated dynamical decadal response to strong volcanic eruptions (Zanchettin et al. 2012b). Correlations between Q_{ISR} and AMV indices are only sporadically <0.5 over multidecadal periods in both the forced and the unperturbed simulation (Fig. 9, left panels). The strength and the lag of full-period peak correlations with Q_{ISR} differ for the different indices and in the two simulations (Fig. 9, insets in the right panels), particularly highlighting a different Q_{ISR} -AMV1 relation under forced and unperturbed conditions. Under forced conditions, multidecadal AMV- Q_{ISR} correlations (Fig. 9, right panels) evolve similarly across the AMV indices but also show distinctive details. For example, around AD 1400 the correlation becomes temporarily largely negative for AMV1 and describes a near-centennial shift to (weak) anti-correlation for AMV2. Similarly, at the end of the XVII century indices either correlate strongly positive (AMV2 and AMV3) or negative (AMV1) with Q_{ISR} . Fig. 9, right panels, further displays possible drawbacks of the linear scaling from unforced to forced patterns since correlations with Q_{ISR} evolve nearly identical for all reconstructed indices (grey lines) but remarkably different from the original ones (black lines). Therefore, decisions made in defining an AMV index may affect our interpretation of simulated phenomena, and the simple scaling limits our ability to interpret the forced dynamics.

Figure 10 illustrates the lag-correlation profiles between AMV indices and the Atlantic meridional overturning circulation (AMOC) at 26.5°N and 1,000 m depth for the unperturbed COSMOS-Mil and MPI-ESM-P/-MR simulations (note: the three models share the same ocean model, see Table 1). It provides an additional example of how differently-defined AMV indices can assist in interpreting simulated dynamics in different models. Independent on

the AMV definition, full-period AMV-AMOC correlations are weakest in COSMOS-Mil, but tight AMV-AMOC correlations are still found in this simulation over multi-centennial sub-periods, indicating that competing processes temporarily dominate multidecadal NASST variability. Correlations peaking at lag zero characterize the AMV1 and AMV2 profiles for MPI-ESM-P/-MR, indicating that in these models fluctuations in the AMV robustly corresponds to fluctuations in the AMOC. In both models, peak correlations are found for AMV3 leading the AMOC of a few years, highlighting the implications of SSTs over the oceanic deep convection regions on the AMOC variability.

3.4 Implications for reconstructed AMV

The results illustrated above stress the influence of the different AMV expressions on our assessment and understanding of simulated climate variability. A pertinent question concerns whether such differences are smeared out when the AMV is reconstructed by using surrogate data rather than being diagnosed from full-field NASSTs. This means, whether we can be confident about which expressions of past NASSTs variability is most likely described in proxy-based AMV reconstructions. The focus here is on AMV reconstructions based on surface (2 m) air temperatures (SATs), following the Gray et al. (2004) approach as described in the methods section. Briefly recalling it: it consists of a multiple linear regression using the first 4 PCs from 10 SAT series sampled from land-only locations and separating the temporal domain in equally populated calibration and validation sets.

Figure 11 summarizes the skill statistics for the SAT-based AMV reconstructions for the considered unperturbed (top panels) and forced (bottom) simulations. We use perfect-knowledge and noise-degraded approaches as described in the methods section. Generally, the 5th–95th percentile bars for the different AMV indices overlap for both skill statistics. No more than about 40 % of the original total AMV variability is explained by the reconstruction, on average, over the validation set. The generally modest skills in reconstructing AMV from land SATs reflect that only few locations over land contain a significant portion of AMV-type signals. Over the ocean the AMV signatures on SATs trace those on SSTs (not shown, compare with Fig. 3), but they generally weaken considerably inland just off the coast and are strongly simulation- and index-dependent. AMV-SAT correlations over landmasses are most extensively significant ($p < 0.1$) in MPI-ESM-P/-MR, where they are mainly located over Western Europe and the UK (for AMV1 and AMV2) and over Scandinavia (for AMV3) (not shown). Correlations are least extensively significant in CCSM4, which consequently shows the overall worse predictive skills (CE < 0 , meaning that the reconstruction is worse than the

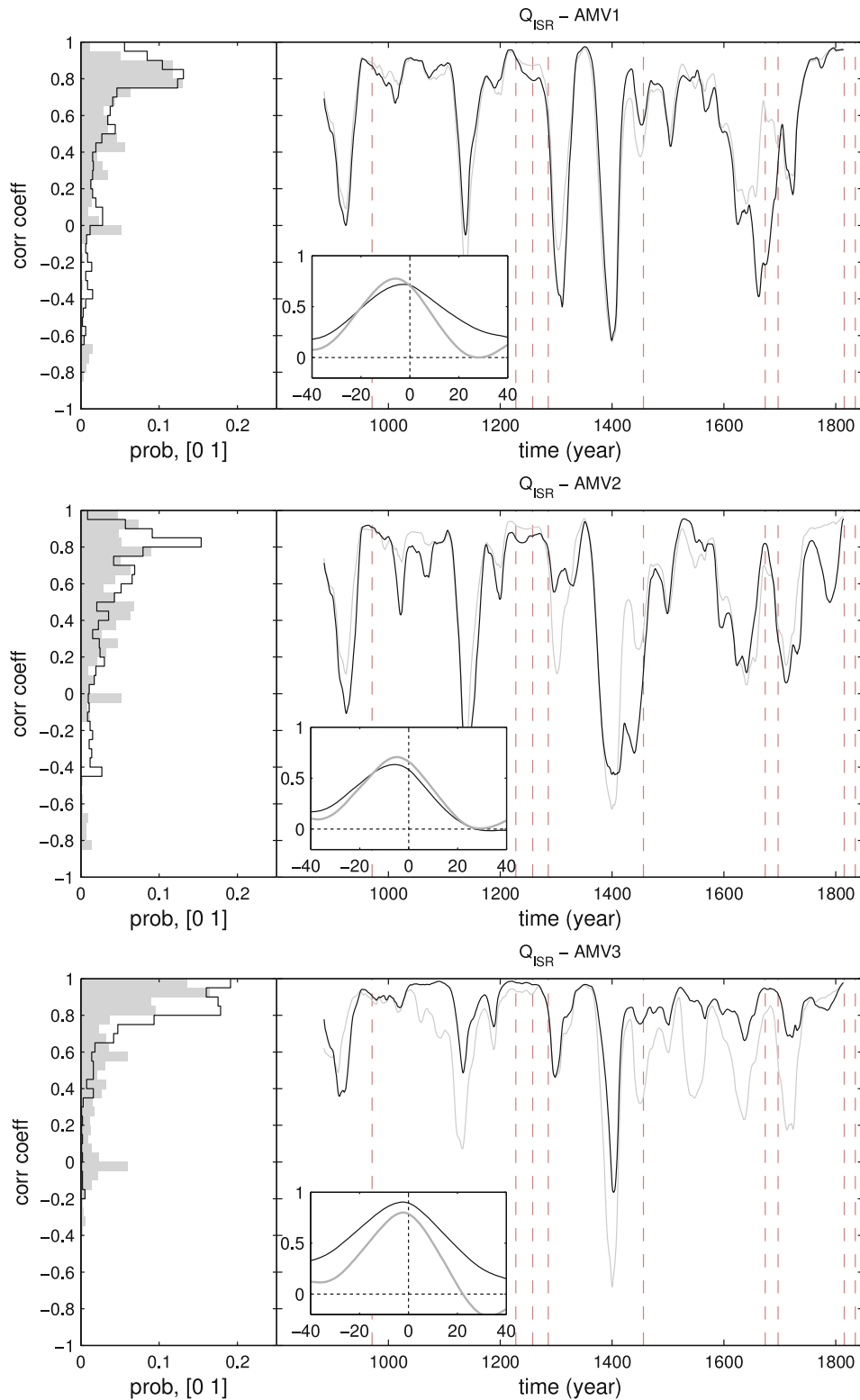


Fig. 9 Moving window correlations (71 years) between smoothed (31 year moving average) ocean heat transport through the Iceland-Scotland-Ridge (Q_{ISR}) and AMV indices in MPI-ESM-P unforced and forced simulations. *Left panels* empirical probability distribution of the correlations for the forced (empty black bars) and for the unforced (filled gray bars) simulation. *Right panels*: evolution of the correlations in the

forced simulations for the original AMV signal (black) and for the AMV reconstructed via linear scaling (grey). *Vertical dashed lines* individuate the reconstructed occurrence of major tropical volcanic eruptions. *Insets* full-length [i.e., O(1,000 years)] lagged correlations for past 1000 (black) and for MPI-ESM-P piControl (gray). AMV indices lead (lag) for positive (negative) lags. Positive Q_{ISR} is defined as for the Nordic Sea gaining heat

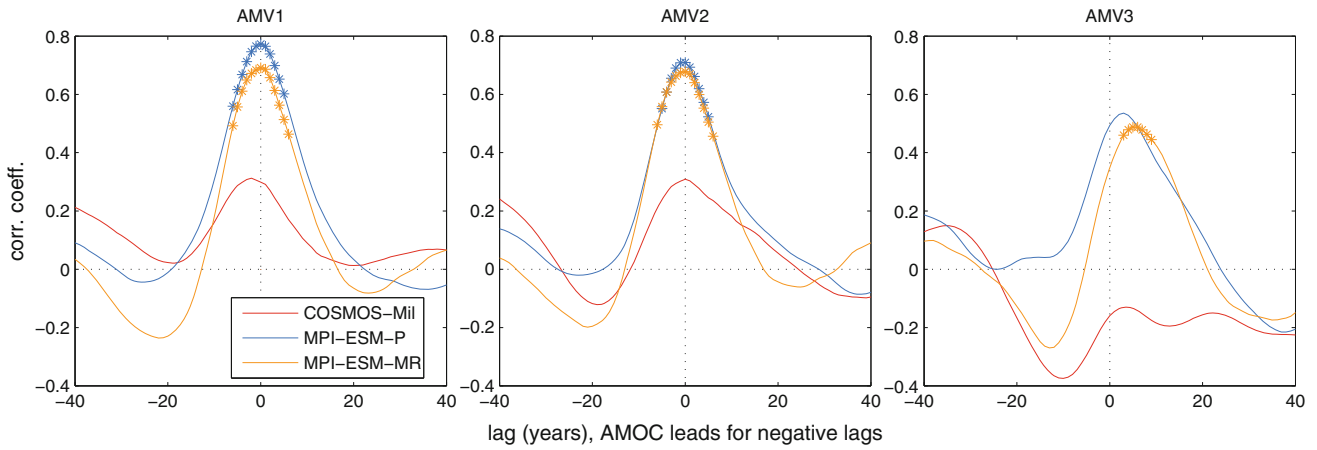


Fig. 10 Full-period lag-correlation profiles between AMV indices (left AMV1; center AMV2; right AMV3) and Atlantic meridional overturning circulation (AMOC) at 26.5°N and 1,000 m depth in the unperturbed COSMOS-Mil and MPI-ESM-P/-MR simulations. Stars

mark correlations significant at $p = 0.1$ accounting for autocorrelation in the time series (see, e.g., Zanchettin et al. 2012a). Data were low-pass filtered by applying an 11 years running moving average prior to the evaluation of lag-correlations

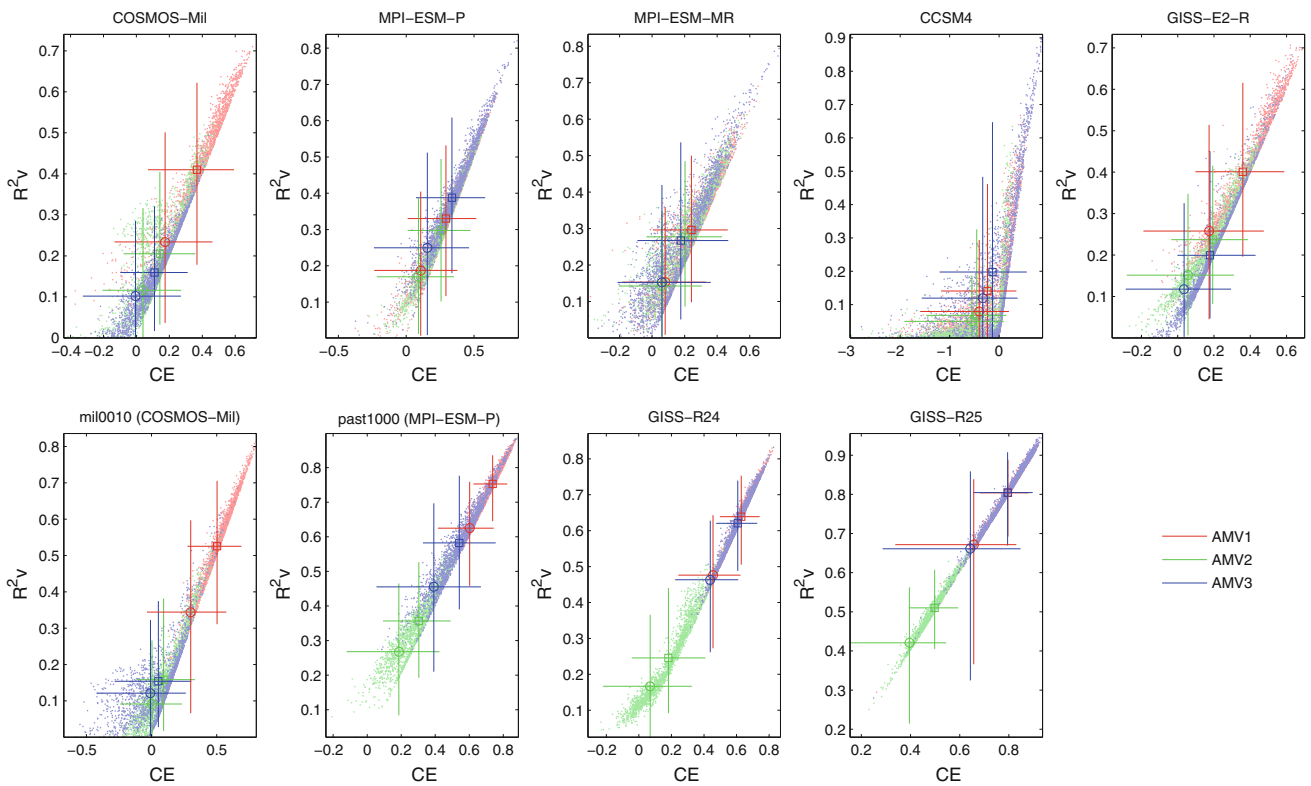


Fig. 11 Squared Pearson's correlation (R^2) and coefficient of error (CE) for within-model AMV indices reconstruction within unperturbed (top) and forced (bottom) simulations based on multiple linear regression using near-surface air temperature predictors (detrended using the same-order polynomial fit used for the AMV, see Table 1). Dots illustrate the results for an ensemble of 1,000 models for the

perfect-knowledge approach with 11 years running average smoothed data. Markers (error bars) individuate the ensemble-average (5th–95th percentile interval). Squared markers perfect-knowledge approach; Circle markers regressions degraded by noise (see “Methods” section). Skill statistics refer to the validation sets of the cross-calibration-verification procedure

average over the validation period). In COSMOS-Mil and GISS-E2-R, AMV1 stands out as the most-likely recovered signal, but there is a tendency towards recovering AMV3 as well in MPI-ESM-P/-MR. The inclusion of artificial noise

generally strongly degrades the regression performances although skillful reconstructions are still possible. It does not change the general picture gained from the perfect-knowledge approach. Thus, following the chosen approach, there is

either a tendency towards more likely recovering AMV1 or AMV3 but not AMV2 signals.

The amount of recovered AMV variability increases under forced pre-industrial conditions compared to unperturbed conditions (Fig. 11, bottom panels). The increased reconstruction skills are due to a substantial strengthening of the forced AMV signatures on SATs over landmasses (not shown). Possible explanations for such strengthening are the stronger AMV-related SST variability, which we expect to be especially important on coastal regions, and the near-simultaneous response of SSTs and land SATs to temporary cooling events due to strong volcanic eruptions. In COSMOS-Mil, the AMV1 reconstruction still outperforms that of AMV2 and AMV3, but in MPI-ESM-P AMV1 now performs at least as good as AMV3. As expected, in GISS-E2-R the skill performances for AMV1 and AMV3 are nearly indistinguishable but still outperform those for AMV2. Therefore, we conclude that the reconstructions more likely recover AMV1/AMV3 than AMV2 forced signals. Additional analyses (not shown) indicate that conclusions similar to those drawn for land SATs stand also for SATs sampled over the North Atlantic Ocean.

4 Summarizing discussion

Investigating multidecadal variability of NASSTs discloses different flavors of the AMV in a selection of unperturbed and forced multicentennial and millennial pre-industrial climate simulations performed with a suite of Earth system models. Whereas the AMV maintains strong O(50–90 years) variability as a main character under unperturbed and externally-forced conditions, differently-defined AMV indices exhibit different signatures on NASSTs, different temporal evolutions and different teleconnections. As seen in the NASST-AMV regression patterns, under unperturbed conditions spatial-average-based indices have a pan-North Atlantic signature but are more closely related to tropical regions, while EOF-based indices generally appear more closely related to extra-tropical variability. Differently-defined AMV indices can be practically (linearly) unrelated in some of the simulations, highlighting the separation between sub-basin-scale NASST variability and particularly between the tropical and extra-tropical ones. Since the spectra of differently-defined AMV indices differ as well in their exact peak frequencies, different AMV manifestations may describe different physical processes, signal propagation pathways and/or resonance (e.g., Liu 2012). The less clear separation between AMV features in other, especially forced, simulations raises however the question whether such AMV-based (statistical) separation is robust and meaningful under all circumstances.

The robust detection of O(50–90 years) variability in the unperturbed simulations supports the hypothesis of the AMV as an internally-generated climate phenomenon (e.g., Ting et al. 2009; Enfield and Cid-Serrano 2010; Park and Latif 2010; Wu et al. 2011a; Zanchettin et al. 2012a). The performed simple scaling assessment indicates that (pre-industrial) external forcings generally affect the variability but not the dynamics associated with the AMV for most of the definitions and simulations. Although our assessment concerns forced climates which differ only slightly from pre-industrial unperturbed conditions in terms of global climatology and variability, it lends further support to the hypothesis that the AMV is, in its different expressions, a feature inherent to simulated climate systems. There are, however, caveats to this general impression concerning especially EOF-based indices and times of large external forcing events (compare correlations from original and reconstructed indices in Fig. 9). Differently-defined AMV indices and oceanic heat transport can display oppositely signed correlations around strong tropical volcanic eruptions, possibly reflecting the different inter-decadal impacts volcanic eruptions produce on simulated North Atlantic oceanic circulation and northward heat transport (Otterå et al. 2010; Zanchettin et al. 2012b). The distinct evolutions of the different indices around strong forcing events complicate the hypothesis of external forcing being “a metronome” of the AMV (Otterå et al. 2010; see also: Booth et al. 2012; Zanchettin et al. 2012a).

Teleconnection patterns of the different AMV definitions have either a regional or a hemispheric character, implying that our interpretation of the AMV's impact depends on our definition of it. Remarkable inter-model consistency between predominant features supports this finding (on inter-model robustness of the AMV signature see also: Ting et al. 2011). In particular, removal of the global SST signal from NASSTs before the evaluation of spatial-average-based AMV indices results in teleconnection patterns changing from being hemispherically to only locally relevant.

The internal/external and hemispheric/regional antagonisms entail strong implications for our understanding and interpretation of paleoclimatic records. In the context of our differently-defined indices, we show the general skill of (2 m) SAT-based reconstructions of the AMV (e.g., Gray et al. 2004). However, the surrogate reconstructions question which aspect of the AMV is captured by real-world reconstructions. Despite the uncertainty in the results, our reconstructions indicate, in unperturbed and forced simulations, comparatively better skills for AMV indices including the global-SST signal. Thus, a linear separation between components of NASST variability is likely too simplistic to successfully distinguish between global/hemispheric and regional as well as externally-driven and

internal signals. Zanchettin et al. (2012a) proposed that different background climate conditions cause differences in the multidecadal large scale climate response to similarly-shaped AMV cooling phases. However, our “reverse” analysis stresses the importance of sub-basin scale NASST variability, which may help to reconcile the spatial and temporal differences in reconstructions of the AMV (e.g., Chylek et al. 2012).

Of course, it remains an open question whether the multicentennial AMV features presented here can be viably extrapolated to the observed and simulated climate over the last century. The background climate state affects inter-annual (e.g., Choi et al. 2011) and multidecadal (e.g., Yoshimori et al. 2010; Zanchettin et al. 2012a) simulated SST variability. It also influences the inter-decadal response of the simulated coupled ocean–atmosphere system to external forcing (e.g., Zhong et al. 2010; Zanchettin et al. 2012b). As we have highlighted, continuous modifications in the processes influencing SST anomalies potentially result in a non-stationary covariance structure of NASSTs, which, in turn, is going to affect the construction of EOF-based AMV indices. On sub-millennial time-scales, the different AMV expressions may be more or less separated, more or less physically relevant and more or less discernible. The strong non-stationarity implied by centennial-scale global warming or cooling complicates the partitioning of the AMV from the global signal (Knight 2009) and results in elusive differences between differently-constructed AMV indices (e.g., Wu et al. 2011b). As outlined here (but see also Zanchettin et al. 2012a), this arises in naturally-forced climate simulations under periods characterized by strong external forcing (e.g., strong volcanic eruptions) when forced signals interfere with rather than simply superpose on the unperturbed multidecadal-to-centennial NASST variability, and, thus, specific processes temporarily dominate the variability.

Our results highlight various further questions:

First, separating the North Atlantic oceanic heat transport into its gyre/horizontal and meridional overturning/vertical components could foster the causal interpretation of the AMV in its different manifestations and consequent dynamical implications. For instance, the magnitude and pattern of simulated AMOC are similar in COSMOS-Mil and MPI-ESM-P/-MR, though peak values at 26.5°N and 1,000 m depth are stronger in MPI-ESM-P (compare Zanchettin et al. 2012b; Jungclaus et al. 2012). As shown here (Fig. 10), whereas the AMOC-AMV lag-correlation profiles in COSMOS-Mil and MPI-ESM-P/-MR are overall congruent for average-based AMV indices, they appreciably differ for the EOF-based AMV indices. The different AMV definitions could therefore assist in clarifying the dynamical relation between simulated multidecadal-to-centennial variability of NASSTs and the overturning

circulation (see also, e.g., Menary et al. 2012). Similarly, the different definitions can also help in disentangling the contribution of the stratosphere to simulated low-frequency NASST variability (e.g., Manzini et al. 2012).

Second, the different atmospheric teleconnections encourage to assess how the different definitions of the AMV impact our interpretation of phasing and inter-dependence of decadal-to-multidecadal variability in different ocean basins (e.g., Zhang and Delworth 2007; Park and Latif 2010; Zanchettin et al. 2012a). Different definitions of the AMV can similarly highlight the repercussions local SST biases have on the realism of simulated inter-basin and/or remote interactions. We note, for instance, that average-based AMV indices in COSMOS-Mil incorporate a strong, regular ENSO-type signal (Fig. 2b, see also: Zanchettin et al. 2012a), a feature which affects less MPI-ESM-P/-MR due to the improved representation of SSTs in the equatorial Pacific region (Jungclaus et al. 2012). The effects of an overly strong and regular ENSO on Pacific-Atlantic interactions could explain why differently-defined AMV indices are most clearly separated in COSMOS-Mil (see correlations in Fig. 2a).

Third, by capturing different features of the simulated large-scale coupled climate system, the different representations of the AMV can further improve our understanding of how the simulated SST variability is affected by the differences between numerical models in the representation of key dynamics and phenomena, e.g., ocean-boundary currents (Kwon et al. 2010; Gent et al. 2011). For instance, the large spread of multi-model EOF-based AMV patterns over the subtropical gyre margin contrasts with the small spread of multi-model average-based AMV patterns over the same region (Fig. 3). This distinctive character gathers further relevance since this region commonly features cold simulated SST biases against observations that are originated by the limited performance of coupled systems to reproduce the observed Gulf Stream separation and North Atlantic current orientation, even in eddy-permitting oceanic resolutions as the one featured in MPI-ESM-MR (Jungclaus et al. 2012).

5 Concluding remarks

Different representations of simulated multidecadal NASST variability describe different statistics, patterns and teleconnections although displaying common multidecadal [O(50–90 years)] variability. They further differ in their response to (mostly natural) external forcing and in their reproducibility. Our results imply that the way we define the AMV affects our interpretation of its statistics and physics. Generally, our results support the paradigm of the AMV as a predominantly internally-generated climate

feature, but different indices display different responses to external forcing events. Furthermore, we cannot unambiguously state, which portion of multidecadal NASST variability is resolved by surface air temperature-based reconstructions of past AMV variability, although it is likely a cumulative signal of remote and local origin. Our results emphasize that the AMV is a multifaceted climate feature. As demonstrated here, a multi-perspective approach in the statistical construct employed for its description can foster our physical understanding of multidecadal NASST variability.

Acknowledgments The authors thank two anonymous reviewers whose comments helped to improve the study and Katja Lohmann for useful comments on an early version of the manuscript. This research was supported by the Max Planck Society for the Advancement of Science. This work was funded by the Federal Ministry for Education and Research in Germany (BMBF) through the research program “MiKlip” (FKZ:01LP1158A). O.B. is funded by the DFG through the Cluster of Excellence CliSAP, University of Hamburg. J.B. is partly funded by the DecCen project funded by the research council of Norway. J. H. J. received funding from the European Community 7th framework program under grant agreement GA212643 (THOR: “Thermohaline Overturning—at Risk?”, 2008–2012). We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling and the participating groups for producing and making available the model output.

References

- Booth BBB, Dunstone NJ, Halloran PR, Andrews T, Bellouin N (2012) Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*. doi:[10.1038/nature10946](https://doi.org/10.1038/nature10946)
- Bothe O, Jungclaus JH, Zanchettin D, Zorita E (2012) Climate of the last millennium: ensemble consistency of simulations and reconstructions. *Clim Past Discuss* 8:2409–2444. doi:[10.5194/cpd-8-2409-2012](https://doi.org/10.5194/cpd-8-2409-2012)
- Brovkin V, Lorenz SJ, Jungclaus J, Raddatz T, Timmreck C, Reick CH, Segsneider J, Six K (2010) Sensitivity of a coupled climate-carbon cycle model to large volcanic eruptions during the last millennium. *Tellus B* 62:674–681. doi:[10.1111/j.1600-0889.2010.00471.x](https://doi.org/10.1111/j.1600-0889.2010.00471.x)
- Choi J, An S-I, Kug J-S, Yeh S-W (2011) The role of mean state on changes in El Niño’s flavours. *Clim Dyn* 37:1205–1215
- Chylek P, Folland CK, Dijkstra HA, Lesins G, Dubey MK (2011) Icecore data evidence for a prominent near 20 year time-scale of the Atlantic multidecadal oscillation. *Geophys Res Lett* 38:L13704. doi:[10.1029/2011GL047501](https://doi.org/10.1029/2011GL047501)
- Chylek P, Folland C, Frankcombe L, Dijkstra H, Lesins G, Dubey M (2012) Greenland ice core evidence for spatial and temporal variability of the Atlantic Multidecadal Oscillation. *Geophys Res Lett* 39:L09705. doi:[10.1029/2012GL051241](https://doi.org/10.1029/2012GL051241)
- Cook ER, Briffa KR, Jones PD (1994) Spatial regression methods in dendroclimatology—a review and comparison of 2 techniques. *Int J Climatol* 14:379–402
- Crowley TJ, Unterman MB (2012) Technical details concerning development of a 1200-yr proxy index for global volcanism. *Earth Syst Sci Data Discuss* 5:1–28. doi:[10.5194/essdd-5-1-2012](https://doi.org/10.5194/essdd-5-1-2012)
- Crowley TJ et al (2008) Volcanism and the little ice age. *Pages News* 16:22–23
- DelSole T, Tippett MK, Shukla J (2011) A significant component of unperturbed multidecadal variability in the recent acceleration of global warming. *J Clim* 24:909–926. doi:[10.1175/2010JCLI3659.1](https://doi.org/10.1175/2010JCLI3659.1)
- Deser C et al (2012) ENSO and Pacific decadal variability in the community climate system model version 4. *J. Clim* 25:2622–2651
- Dima M, Lohmann G (2009) Evidence for two distinct modes of large-scale ocean circulation changes over the last century. *J Clim* 23:5–16. doi:[10.1175/2009JCLI2867.1](https://doi.org/10.1175/2009JCLI2867.1)
- Dommenget D, Latif M (2008) Generation of hyper-climate modes. *Geophys Res Lett* 35:L02706. doi:[10.1029/2007GL031087](https://doi.org/10.1029/2007GL031087)
- Enfield DB, Cid-Serrano L (2006) Projecting the risk of future climate shifts. *Int J Climatol* 26:885–895
- Enfield DB, Cid-Serrano L (2010) Secular and multidecadal warmings in the North Atlantic and their relationships with major hurricane activity. *Int J Climatol* 30(2):174–184
- Enfield DB, Mestas-Núñez AM (1999) Multiscale variabilities in global sea surface temperatures and their relationships with tropospheric climate patterns. *J Clim* 12:2719–2733
- 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
- Fernández-Donado L et al (2012) Temperature response to external forcing in simulations and reconstructions of the last millennium. *Clim Past Discuss* 8:4003–4073. doi:[10.5194/cpd-8-4003-2012](https://doi.org/10.5194/cpd-8-4003-2012)
- Folland CK, Parker DE, Palmer TN (1986) Sahel rainfall and worldwide sea temperatures, 1901–1985. *Nature* 320:602–607
- Gao C, Robock A, Ammann C (2008) Volcanic forcing of climate over the last 1500 years: an improved ice-core based index for climate models. *J Geophys Res* 113:D2311. doi:[10.1029/2008JD010239](https://doi.org/10.1029/2008JD010239)
- Gent PR et al (2011) The community climate system model version 4. *J Clim* 24:4973–4991. doi:[10.1175/2011JCLI4083.1](https://doi.org/10.1175/2011JCLI4083.1)
- Giorgetta MA et al (2012) Climate change from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project 5. Submitted to JAMES, special issue. The Max Planck Institute for Meteorology Earth System Model
- Gray ST, Graumlich LJ, Betancourt JL, Pederson GT (2004) A tree-ring based reconstruction of the Atlantic multidecadal oscillation since 1567 AD. *Geophys Res Lett* 31:L12205
- Grosfeld K, Lohmann G, Rimbu N (2008) The impact of Atlantic and Pacific Ocean sea surface temperature anomalies on the North Atlantic multidecadal variability. *Tellus*. doi:[10.1111/j.1600-0870.2008.00304.x](https://doi.org/10.1111/j.1600-0870.2008.00304.x)
- Henriksson SV, Räisänen P, Silén J, Laaksonen A (2012) Quasiperiodic climate variability with a period of 50–80 years: fourier analysis of measurements and earth system model simulations. *Clim Dyn*. doi:[10.1007/s00382-012-1341-0](https://doi.org/10.1007/s00382-012-1341-0)
- Jungclaus JH et al (2010) Climate and carbon-cycle variability over the last millennium. *Clim Past* 6:723–737. doi:[10.5194/cp-6-723-2010](https://doi.org/10.5194/cp-6-723-2010)
- Jungclaus JH et al (2012) Characteristics of the ocean simulations in MPIOM, the ocean component of the Max Planck Institute Earth System Model. Submitted to JAMES, special issue. The Max Planck Institute for Meteorology Earth System Model
- Kerr RA (2000) A North Atlantic climate pacemaker for the centuries. *Science* 288:1984–1986
- Knight JF (2009) The Atlantic multidecadal oscillation inferred from the forced climate response in coupled general circulation models. *J Clim* 22:1610–1625
- Knudsen MF, Seidenkrantz M-S, Jacobsen BH, Kuijpers A (2011) Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years. *Nature Comm* 2:178. doi:[10.1038/ncomms1186](https://doi.org/10.1038/ncomms1186)
- Kwon Y-O et al (2010) Role of Gulf Stream and Kuroshio-Oyashio systems in large-scale atmosphere-ocean interaction: a review. *J Clim* 23:3249–3281

- Landrum L, Otto-Bliesner BL, Wahl ER, Conley A, Lawrence PJ, Teng H (2011) Last millennium climate and its variability in CCSM4. *J Clim*. doi:[10.1175/JCLI-D-11-00326.1](https://doi.org/10.1175/JCLI-D-11-00326.1)
- Liu Z (2012) Dynamics of interdecadal climate variability: a historical perspective. *J. Clim* 25:1963–1995. doi:[10.1175/2011JCLI3980.1](https://doi.org/10.1175/2011JCLI3980.1)
- Mann ME, Emanuel KA (2006) Atlantic hurricane trends linked to climate change. *Eos Trans AGU* 87(24):233–244
- Mann ME, Zhang Z, Rutherford S, Bradley RS, Hughes MK, Shindell D, Ammann C, Faluvegi G, Ni F (2009) Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science* 326:1256–1260
- Manzini E, Cagnazzo C, Fogli PG, Bellucci A, Müller WA (2012) Stratosphere-troposphere coupling at inter-decadal time scales: implications for the North Atlantic Ocean. *Geophys Res Lett* 39:L05801. doi:[10.1029/2011JGL050771](https://doi.org/10.1029/2011JGL050771)
- Marshall J et al (2001) North Atlantic climate variability: phenomena, impacts and mechanisms. *Int J Climatol* 21:1863–1898. doi:[10.1002/joc.693](https://doi.org/10.1002/joc.693)
- Menary MB, Park W, Lohman K, Vellinga M, Palmer MD, Latif M, Jungclaus JH (2012) A multimodel comparison of centennial Atlantic meridional overturning circulation variability. *Clim Dyn* 38:2377–2388
- Neumaier A, Schneider T (2001) Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans Math Softw* 27:27–57
- Otterå OH, Bentsen M, Drange H, Suo L (2010) External forcing as a metronome for Atlantic multidecadal variability. *Nat Geosci*. doi:[10.1038/NGEO995](https://doi.org/10.1038/NGEO995)
- Park W, Latif M (2010) Pacific and Atlantic multidecadal variability in the Kiel Climate Model. *Geophys Res Lett* 37:L24702. doi:[10.1029/2010JGL045560](https://doi.org/10.1029/2010JGL045560)
- Saenger C, Cohen AL, Oppo DW, Halley RB, Carilli JE (2009) Surface temperature trends and variability in the low-latitude North Atlantic since 1552. *Nat Geosci* 2:492–495. doi:[10.1038/ngeo552](https://doi.org/10.1038/ngeo552)
- Schlesinger ME, Ramankutty N (1994) An oscillation in the global climate system of period 65–70 years. *Nature* 367:723–726. doi:[10.1038/367723a0](https://doi.org/10.1038/367723a0)
- Schmidt GA et al (2011) Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). *Geosci Model Dev* 4:33–45. doi:[10.5194/gmd-4-33-2011](https://doi.org/10.5194/gmd-4-33-2011)
- Ting M, Kushnir Y, Seager R, Li C (2009) Forced and internal twentieth century SST in the North Atlantic. *J Clim* 22:1469–1481. doi:[10.1175/2008JCLI2561.1](https://doi.org/10.1175/2008JCLI2561.1)
- Ting M, Kushnir Y, Seager R, Li C (2011) Robust features of Atlantic multi-decadal variability and its climate impacts. *Geophys Res Lett* 38:L17705. doi:[10.1029/2011JGL048712](https://doi.org/10.1029/2011JGL048712)
- 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)
- Vieira LEA, Solanki SK, Krivova NA, Usoskin I (2011) Evolution of the solar irradiance during the holocene. *Astron Astrophys* 531:A6. doi:[10.1051/0004-6361/201015843](https://doi.org/10.1051/0004-6361/201015843)
- Vincze M, Jánosi IM (2011) Is the Atlantic Multidecadal Oscillation (AMO) a statistical phantom? *Nonlin Proc Geophys* 18:469–475. doi:[10.5194/npg-18-469-2011](https://doi.org/10.5194/npg-18-469-2011)
- Wei W, Lohmann G (2012) Simulated atlantic multidecadal oscillation during the holocene. *J Clim* 25:6989–7002. <http://dx.doi.org/10.1175/JCLI-D-11-00667.1>
- Wu Z, Huang NE, Wallace JM, Smoliak BV, Chen X (2011a) On the time-varying trend in global-mean surface temperature. *Clim Dyn* 37:759–773. doi:[10.1007/s00382-011-1128-8](https://doi.org/10.1007/s00382-011-1128-8)
- Wu S, Liu Z, Zhang R, Delworth TL (2011b) On the observed relationship between the Pacific decadal oscillation and the Atlantic Multi-decadal Oscillation. *J Oceanogr* 67:27–35. doi:[10.1007/s10872-011-0003-x](https://doi.org/10.1007/s10872-011-0003-x)
- Wyatt MG, Kravtsov S, Tsonis AA (2011) Atlantic multidecadal oscillation and Northern Hemisphere's climate variability. *Clim Dyn*. doi:[10.1007/s00382-011-1071-8](https://doi.org/10.1007/s00382-011-1071-8)
- Yoshimori M, Raible CC, Stocker TF, Renold M (2010) Simulated decadal oscillations of the Atlantic meridional overturning circulation in a cold climate state. *Clim Dyn* 34:101–121. doi:[10.1007/s00382-009-0540-9](https://doi.org/10.1007/s00382-009-0540-9)
- Zanchettin D, Rubino A, Traverso P, Tomasino M (2008) Impact of variations in solar activity on hydrological decadal patterns in northern Italy. *J Geophys Res* 113:D12102. doi:[10.1029/2007JD009157](https://doi.org/10.1029/2007JD009157)
- Zanchettin D, Rubino A, Jungclaus JH (2010) Intermittent multidecadal-to-centennial fluctuations dominate global temperature evolution over the last millennium. *Geophys Res Lett* 37:L14702. doi:[10.1029/2010JGL043717](https://doi.org/10.1029/2010JGL043717)
- Zanchettin D, Rubino A, Bothe O, Matei D, Jungclaus JH (2012a) Multidecadal-to-centennial SST variability in the MPI-ESM simulation ensemble for the last millennium. *Clim Dyn*. doi:[10.1007/s00382-012-1361-9](https://doi.org/10.1007/s00382-012-1361-9)
- Zanchettin D, Timmreck C, Graf H-F, Rubino A, Lorenz S, Krueger K, Lohmann K, Jungclaus JH (2012b) Bi-decadal variability excited in the coupled ocean–atmosphere system by strong tropical volcanic eruptions. *Clim Dyn* 39(1–2):419–444. doi:[10.1007/s00382-011-1167-1](https://doi.org/10.1007/s00382-011-1167-1)
- Zhang R, Delworth TL (2007) Impact of the Atlantic Multidecadal Oscillation on North Pacific climate variability. *Geophys Res Lett* 34:L23708. doi:[10.1029/2007GL031601](https://doi.org/10.1029/2007GL031601)
- Zhong Y et al (2010) Centennial-scale climate change from decadal-paced explosive volcanism: a coupled sea ice-ocean mechanism. *Clim Dyn* 23:5–7. doi:[10.1007/s00382-010-0967-z](https://doi.org/10.1007/s00382-010-0967-z)