ENVIRONMENTAL RESEARCH
CLIMATE

PURPOSE-LED PUBLISHING™ Part of the ENVIRONMENTAL RESEARCH SERIES

PAPER • OPEN ACCESS

Enhancing climate variability in earth system models through paleo-informed nudging of sea surface temperatures

To cite this article: V Skiba *et al* 2026 *Environ. Res.: Climate* **5** 015025

ENVIRONMENTAL RESEARCH
CLIMATE

OPEN ACCESS

RECEIVED
15 September 2025

REVISED
24 December 2025

ACCEPTED FOR PUBLICATION
19 January 2026

PUBLISHED
19 February 2026

PAPER

Enhancing climate variability in earth system models through paleo-informed nudging of sea surface temperatures

V Skiba[1,*], **L Ackermann**[2], **N Hirsch**[1,3], **Y Sun**[2], **A M Dolman**[1], **G Lohmann**[2,4,5] and **T Laepple**[1,3,5]

[1] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[2] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
[3] Faculty of Geosciences, University of Bremen, Bremen, Germany
[4] Faculty of Physics, University of Bremen, Bremen, Germany
[5] MARUM Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany
[*] Author to whom any correspondence should be addressed.

E-mail: vanessa.skiba@awi.de

Keywords: internal variability, model nudging, paleoclimate

Supplementary material for this article is available online

Abstract

Knowledge on natural climate variability is essential for making informed projections about future climate change. Yet, as highlighted by previous studies using paleoclimate reconstructions, climate models (CMs) often underestimate surface temperature variability on multidecadal to millennial timescales, both in the ocean and on land. Here, we enhance the long-term memory of sea surface temperature (SST) in the Alfred Wegener Institute Earth System Model (ESM) by nudging SSTs towards modified values based on estimates from marine paleoclimate archives. SSTs are nudged, while other dynamics and thermodynamics evolve freely, meaning the atmosphere and sea ice adjust to the altered SST variability. We find that enhanced long-term ocean memory leads to increased atmospheric temperature variability, while maintaining realistic decadal to interannual climate variability modes and mean climate state. It further brings model output into closer agreement with ice core–derived temperature reconstructions from Greenland and East Antarctica. These results support the hypothesis that ocean variability shapes long-term land temperature fluctuations and underscores the need for improved representation of long-term ocean memory in CMs and ESMs. Our framework enables exploration of long-term natural variability, likely underestimated by current ESMs, with applications from paleo-data assimilation to detection and attribution and future climate-risk assessment.

1. Introduction

Knowledge of the characteristics of natural climate variability is crucial for assessing the range of plausible future climate trajectories in the next centuries. Previous studies suggest that climate models (CMs) seem to fail to capture long-term regional variability recorded in paleoclimate temperature proxy data. Marine paleoclimate reconstructions and instrumental data show larger amplitudes of regional sea surface temperature (SST) variability on multi-decadal to millennial timescales (hereafter referred to as 'supradecadal') than CM and Earth system model (ESM) simulations, implying a stronger memory (i.e. persistence of anomalies) in ocean surface temperatures on those timescales (Laepple and Huybers 2014a, 2014b, Cheung *et al* 2017, Dee *et al* 2017, Parsons *et al* 2017). Ocean variability is a driver of low-frequency temperature variability over land, as shown in CM simulations (Dommenget 2009), as well as by the spatial pattern of millennial-scale temperature variability reconstructed from proxy data (Hébert *et al* 2022). In fact, it has been also found that land variability on supradecadal timescales is underestimated by CMs and ESMs (Ellerhoff and Rehfeld 2021, Ellerhoff *et al* 2022, Hébert *et al* 2022, Laepple *et al* 2023).

Various causes have been suggested for lower supradecadal variability in CMs and ESMs relative to paleo-data. The models may lack natural forcing or underestimate the dynamical response to that forcing due to limitations in the models (Shindell *et al* 2001). The absence of interactive components such as ice sheets and dynamic vegetation (Bonan *et al* 2008, Bakker *et al* 2017, Braconnot *et al* 2019, Laguë *et al* 2019, Hopcroft and Valdes 2021) may also cause a lack of regional climate variability. Furthermore, it has been suggested that weak or poorly represented ocean eddies (Jüling *et al* 2021), and limitations in the representation and propagation of sub-grid-scale processes may artificially dampen regional, supradecadal temperature variability in models.

The potential underestimation of slow climate variability and long-term memory in models has important consequences for projections of future change, because memory in the climate system is closely linked to predictability. Low-frequency variability can either reduce predictive skill by obscuring forced signals (Hawkins and Sutton 2009) or increase predictive skill when persistence can be exploited for prediction (Laepple *et al* 2008, Hurrell *et al* 2010). It affects prediction of future regional climate trends, extreme event statistics, attribution studies and impact assessments (Calel *et al* 2020, Harrington *et al* 2021, Schwarzwald and Lenssen 2022). For example, larger variability can increase the frequency of extreme weather and climate events (Blanusa *et al* 2023, Laepple *et al* 2023). This is important to consider in climate attribution assessments, which use climate or earth system model (ESM) simulations of historical scenarios to understand the extent to which anthropogenic climate change has altered the probability or magnitude of particular climate events. Additionally, climate variability has been identified as an important driver of polar ice sheet mass balance (e.g. Jenkins *et al* 2018, Holland *et al* 2019, Wood *et al* 2021), which constitutes a large uncertainty source of future sea-level predictions (Tsai *et al* 2017, 2020, Robel *et al* 2019, Caillet *et al* 2025). For example, internal climate variability can trigger ice sheet retreat via processes like marine ice cliff instability, which is the result of the combination of hydrofracturing due to surface melt and failure of vertical ice cliffs at the grounding line (Tsai *et al* 2020).

Here, we introduce and test a method for enhancing long-term internal variability in a fully-coupled ESM according to paleo-observations by nudging the Alfred Wegener Institute (AWI) ESM towards SSTs with enhanced long-term memory. We rescale SSTs of a 1500 year equilibrium simulation with fixed pre-industrial (PI) boundary conditions of the model and rerun it with the modified SST fields nudged at each timestep while other dynamics and thermodynamics evolve freely. This enables us to explore the impact of enhanced long-term ocean memory on atmospheric variability. We compare the results to ice-core-based reconstructions of surface air temperature variability over the polar ice sheets to test the nudged model's ability to reproduce realistic long-term climate variability.

2. Material and methods

2.1. AWI-ESM simulation

The AWI-ESM is a comprehensive ESM that includes the (AWI-CM) which has been widely used for past and future simulations and participated in the latest Coupled Model Intercomparison Project phase (CMIP6) (Sidorenko *et al* 2015, 2019, Shi and Lohmann 2016, Rackow *et al* 2018, Semmler *et al* 2020).

The model version used in this study (AWI-ESM2) includes an updated version of the finite volume sea ice—ocean model (FESOM2) which is described in Danilov *et al* (2017). The model employs an unstructured grid which allows for a spatially varying horizontal resolution (up to 20 km and 120 km in the high latitudes and low latitudes, respectively; vertical discretisation of 47 levels). It has been widely used for past, present and future applications (e.g. Danabasoglu *et al* 2016, Sein *et al* 2016, Lohmann *et al* 2020, Shi *et al* 2022a). The atmosphere component of AWI-ESM2 is the sixth generation of the atmospheric general circulation model ECHAM, developed by the Max Planck Institute for Meteorology in Hamburg, Germany (ECHAM6; T63L47 setup with horizontal resolution of $\sim$1.9° and 47° vertical levels; Stevens *et al* 2013). It is comprised of a spectral dynamical core and includes a hydrological discharge model for river runoff (Hagemann and Dümenil 1997) and a land surface model (Reick *et al* 2021) to account for dynamic vegetation.

For the PI equilibrium simulation, the ocean component was initialised with climatological temperature and salinity fields from the world ocean atlas (period 1950–2000, Levitus *et al* 2010). No flux corrections or additional restoring were applied after initialisation.

We integrated the AWI-ESM2 (from now on AWI-ESM) for a total of 3000 years under fixed PI boundary conditions following the protocols of PMIP4 (Otto-Bliesner *et al* 2017). To provide a well-equilibrated baseline model state for our experiment, we use the final 1500 years, during which the model exhibits near-stationary behaviour. Global mean SST trends remain within quasi-equilibrium

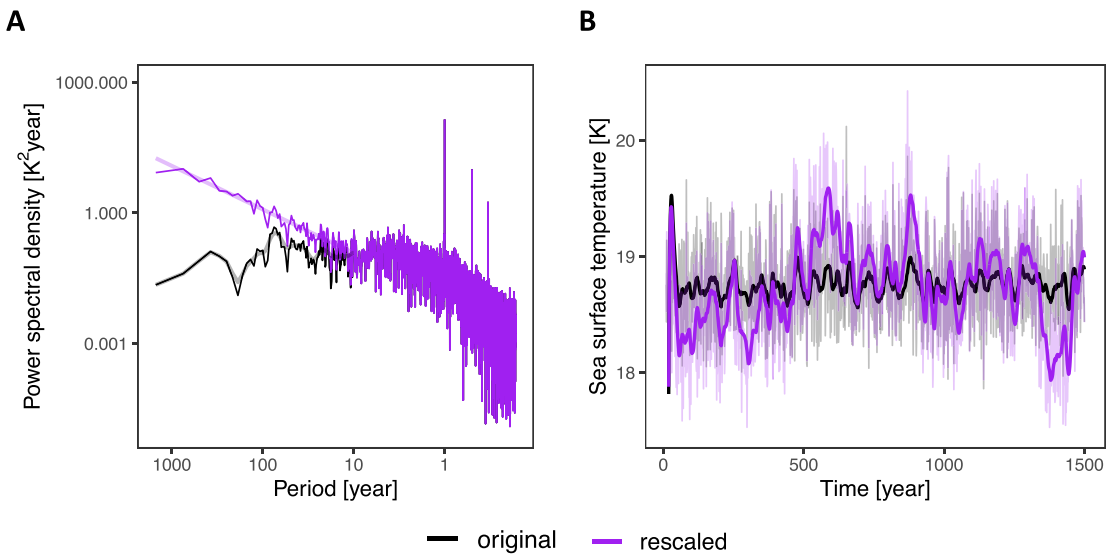

Figure 1. Rescaling of a time series for an example grid cell. (A). For the rescaling, the smoothed spectrum (thick grey line) is calculated from the original spectrum (thin black line) using a log smoother. The rescaled power spectral density (thin purple line) is obtained by multiplying the original spectrum with the ratio between the smoothed original spectrum and the power-law model with exponent of one (thick purple line). (B). Low-pass filtered (cut-off frequency = 1/20 yr^{-1}) time series in the time domain indicated as thick lines. Non-filtered time series are shown as thin lines.

(Shi *et al* 2022b), with a maximum trend not exceeding 0.05 K/century and an average trend of 0.007 K/century.

2.2. Spectral estimation and rescaling of SST

Our rescaling approach (hereafter also referred to as 'spectral rescaling') is informed by long-term surface ocean variability, estimated from SST reconstructions for the Holocene time period from marine paleo-data. Previous studies have quantified variability from those reconstructions by estimating their power spectral density (PSD), y, as a function of frequency, f, the inverse of the timescale τ, i.e. $\tau = f^{-1}$. Their relationship can be statistically approximated by a power-law model described by equation (1).

$$y = \alpha f^{-\beta} \tag{1}$$

where α is a coefficient, β is the power-law exponent which informs about the 'scaling', i.e. the change in PSD occurring with timescale (frequency) and, hence, how much memory the system has.

SSTs reconstructed from marine sediment core proxy records covering the mid- to late Holocene (past 8000 years) and the last millennium have been shown to have a power-law exponent β of approximately one at supradecadal timescales (Laepple and Huybers 2013, Ellerhoff and Rehfeld 2021). Hence, we modify each local grid cell from the SST fields of the 1500 year-long AWI-ESM PI equilibrium simulation so that spectral slopes have a power-law exponent β of approximately one at frequencies below 1/10 yr^{-1}.

To achieve this, after removing the mean and detrending in order to prevent leakage of the trend onto higher frequencies, we apply a transfer function to the Fourier transform of each SST time series so that the re-scaled real components follow a power-law with $\beta = 1$ at frequencies below $f = 1/10$ yr^{-1}. We leave the imaginary component of the Fourier transform unchanged, so that the magnitude and persistence of anomalies are rescaled but their timing (phase information) is preserved. We then apply the inverse Fourier transform to obtain the rescaled time series and re-add the original mean and linear trend (figure 1(B)). In polar regions, rescaled SSTs are truncated at the model's freezing threshold (–1.8 °C) before being prescribed.

To obtain the transfer functions, we estimate the spectrum of SST using Thomson's multitaper method with three tapers (Percival and Walden 1993). Because detrending and tapering bias the lowest frequencies, we omit them from all plots and calculations. We then smooth the spectrum with a Gaussian kernel (width = 0.05 in log-frequency) and parameterise a power-law model with $\beta = 1$ by matching the observed smoothed power at 1/10 yr^{-1} frequency (figure 1(A)). The transfer function for the rescaling is then the ratio of the original spectrum to this power-law model for frequencies below 1/10 yr^{-1}.

2.3. Nudging of SSTs

We nudge the model SSTs towards modified (rescaled) values while other dynamics and thermodynamics evolve freely, allowing the atmosphere and sea ice to adjust to the altered SST variability through their coupling with the ocean. We implement this by including an additional term $hf*_{\mathrm{relax}}$ at the ocean surface, defined as the difference between modelled (SST) and prescribed SST ($\mathrm{SST}_{\mathrm{pre}}$) and multiplied by a relaxation factor (f_{relax}, equation (2).

$$hf*_{\mathrm{relax}} = -f_{\mathrm{relax}} \cdot (\mathrm{SST} - \mathrm{SST}_{\mathrm{pre}}). \tag{2}$$

We add this term to the heat flux at the ocean surface to nudge the model SST towards the prescribed one. We prescribe rescaled monthly SST fields at each (30 min) time step which results in a lack of SST variability on frequencies higher than monthly, i.e. unrealistic model output on these frequencies. To select a relaxation factor value, we perform several 150 year-long nudged simulations using rescaled SSTs with different relaxation factors (figure S1). We choose a relaxation factor of 1.929^{-3} m s^{-1}, corresponding to a relaxation time of 0.3 days over a layer thickness of 50 m. With such a small relaxation time, SSTs are similar to the prescribed (modified) SSTs. This relaxation factor is similar to those used in previous SST nudging experiments (e.g. Keenlyside *et al* 2005, Manda *et al* 2005, Kong *et al* 2024).

2.4. Variability modes

We investigate the impact of nudging the AWI-ESM on the climate variability modes El Niño-Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) and the Atlantic Multi-Decadal Variability (AMV). For this, we quantify ENSO variability with the Niño 3.4 index, the area-weighted SST mean over 5° N–5° S, 170–120° W, detrended with a 30 year climatology, smoothed with a 5 month running mean, and normalised by the climatological standard deviation (SD) (Bamston *et al* 1997, Trenberth and Stepaniak 2001). We calculate the NAO index as the normalised DJF surface pressure difference between Stykkishólmur/Reykjavík and Lisbon (Hurrell 1995). For quantifying the AMV, we calculate the area-weighted spatial average of SSTs over the North Atlantic region (0° N–65° N, 80° W–0° E; Enfield *et al* 2001). We linearly detrend the spatially and annually averaged time series and apply a low-pass filter with cut-off frequency at 1/50 yr^{-1} to investigate the change due to the nudging specifically at low-frequencies. To compare to observations, we use SSTs from two instrumental datasets, the HadISST (Hadley Centre Sea Ice and SST) dataset (Rayner *et al* 2003) and the infilled version of the dynamically consistent ensemble of temperature at the earth surface (DCENT-I) dataset (Chan *et al* 2025).

2.5. Variability estimates from ice cores

We compare air temperature variability from the nudged model simulation to ice-core-based variability reconstructions from Greenland and Antarctica. As stable oxygen isotope records (δ^{18}O) from ice cores are affected by independent noise, the signal can be separated from the noise by using multiple ice cores in close proximity and comparing the PSD of the mean record with the mean PSD calculated for individual records (Münch and Laepple 2018). δ^{18}O records can be further corrected for isotopic diffusion in the firn, and for the effects of time uncertainty which can lead to misalignment of the signal between records and false attribution as noise, by modelling the effects of diffusion and time uncertainty in the frequency domain (Münch and Laepple 2018).

We use the reconstruction for the plateau region of Dronning Maud Land (DML), Antarctica, from Münch and Laepple (2018), which is based on firn core measurements with annual resolution. It was corrected in the described way and is, thus, a robust estimate of the δ^{18}O variability on decadal to multi-centennial timescales in Antarctica for the last ∼1000 years which we convert to air temperature using a regional calibration (Stenni *et al* 2017; 0.93 ± 1 SD of 0.05‰ $°C^{-1}$).

We estimate variability over decadal to multi-centennial timescales in Greenland using two firn and ice core δ^{18}O datasets following the approach of Münch and Laepple (2018). One consists of the firn cores used for the annually resolved North Greenland traverse δ^{18}O stack (hereafter 'NGT' stack; Hörhold *et al* 2023). As the methodology for signal separation relies on a fixed number of records for each time point (Münch and Laepple 2018), we use the 14 of 16 available firn core records which cover the last 474 years (1505–1979) following Hörhold *et al* (2023). The other dataset spans the last 7000 years, which is considered as a stable phase of the Holocene, and consists of three δ^{18}O records in 20 year resolution dated within the Greenland Ice Core Chronology 2005 (Seierstad *et al* 2014, Rasmussen *et al* 2014; hereafter 'GICC' stack). The 20 year resolution of the chronology of these three cores is the result of a decrease in resolution for glacial sections. For the Holocene, the sampling resolution is typically annual, similar to the ice cores of the 'NGT' stack and the Antarctic firn cores, which is then averaged to the lower resolution (Seierstad *et al* 2014). Thus, we do not expect any aliasing effects on PSD estimates as higher frequency variance is removed by the averaging step.

The 'NGT' stack is corrected for isotopic diffusion and time uncertainty (up to 5 years, M. Hörhold, pers. communication; figure S2). As firn diffusion and time uncertainty is only affecting high frequencies, correction of the'GICC' stack is not necessary. For conversion to temperature, we estimate the δ^{18}O-air temperature linear regression slope between the topography-corrected air temperatures averaged across the ice core locations from a regional CM (available for the period from 1871 to 2012; Fettweis *et al* 2017) and the 'NGT' stack as provided by Hörhold *et al* (2023) including all 16 firn cores as varying number of records for each time point can be considered for regression (figure S3). Using an 11 year running mean to focus on (multi-)decadal timescales, we obtain a slope of 0.77 $\pm$ 0.19 ‰°C^{-1} (figure S4), which is within the range of previous estimates, Hörhold *et al* 2023).

We compare the ice-core based temperature spectra with the modelled air temperature, regridded to each ice core location using nearest neighbour remapping. Subsequently, we average modelled air temperature time series from the ice core locations in the same way as described above for the ice core δ^{18}O data.

3. Results

3.1. Mean state

Annual mean local SSTs are similar between nudged and model runs for most regions, with 85% of grid cells differing by less than 0.05 K (figure 2(A)). The largest differences occur in the polar regions (mainly south of 60° S and north of 60° N) where mean SSTs in the nudged model simulation are up to ∼0.3 K higher. The spatial pattern of annual mean 2 m air temperature differences closely follows that of the SSTs (figures 2(A) and (B)). The largest anomalies occur in the polar regions (south of ∼50° S and north of ∼50° N), where nudged simulations are up to ∼4 K warmer, a result of limiting rescaled SSTs to a minimum of—1.8 °C in sea-ice grid cells. Additionally, a cooling of up to ∼1 K is found in parts of the North Atlantic and continental East Asia (figure 2(B)).

Changes in mean local surface pressure due to the nudging are more heterogenous (figure 2(C)). Pressure decreases over the Arctic and Atlantic Ocean, while Greenland shows an increase. Antarctica, the adjacent Southern Ocean, and the Indian Ocean also show higher values compared to the original simulation.

3.2. SST variability

3.2.1. ENSO

The nudged model run is characterised by decades- to century-long periods with higher or lower values of the NINO3.4 index (figure 3 (A)). Thus, higher intensities of El Niño or La Niña events occur, i.e. larger deviations of the index above 0.4 or below −0.4, respectively, following Trenberth and Stepaniak (2001). For example, years 600–620 show a tendency towards stronger El Niño conditions while years 660–700 show a tendency towards stronger La Niña conditions (figure 3 (B)). Since no rescaling was applied at frequencies above 1/10 yr^{-1}, the Niño 3.4 index shows no changes at ENSO's dominant frequencies (1/7–1/2 yr^{-1}; figure 3 (C)). PSD estimates of the Niño 3.4 index start to diverge at multi-decadal timescales. For these timescales, estimates from instrumental datasets vary. While the HadISST-estimated multi-decadal variability is more similar to the original model simulations, multi-decadal SST variability estimated from the DCENT-I dataset is closer to the nudged model simulation.

We find only minor changes in the spatial pattern of the ENSO impact due to the nudging, with differences in regression slopes with local SSTs between −0.08–0.06 K K^{-1} (10 and 90th percentiles, respectively; figure 4). However, local relationships seem to increase for both the negative and the positive ones (figures S5(A) and (B)). This results in a slight increase in the range between 10th and 90th percentiles of absolute slopes from 0.4 K K^{-1} (–0.13–0.27 K K^{-1}, respectively) in the original model simulation to 0.46 K K^{-1} (–0.15–0.31 K K^{-1}, respectively) in the nudged simulation.

3.2.2. AMV

We further investigate changes in the modelled spatial impact of the AMV as an example for a low-frequency climate variability mode. Due to the nudging, amplitudes of centennial- to multi-centennial-scale AMV are increasing (figures 5(A) and (B)). These timescales (>100 years) are longer than what can be estimated by observations due to data coverage being limited in time (figure 5(B), vertical lines mark typical AMV frequencies of 1/70–1/50 yr^{-1}).

Spatial regression maps of the AMV index onto local SSTs reveal a more global impact of the AMV in the nudged simulation compared to the original model simulation (figures 5(C) and (D)). While in the original model run regression slopes (positive) are mostly confined to the Northern Hemisphere, more positive relationships (slopes) exist for the Southern Hemisphere in the nudged

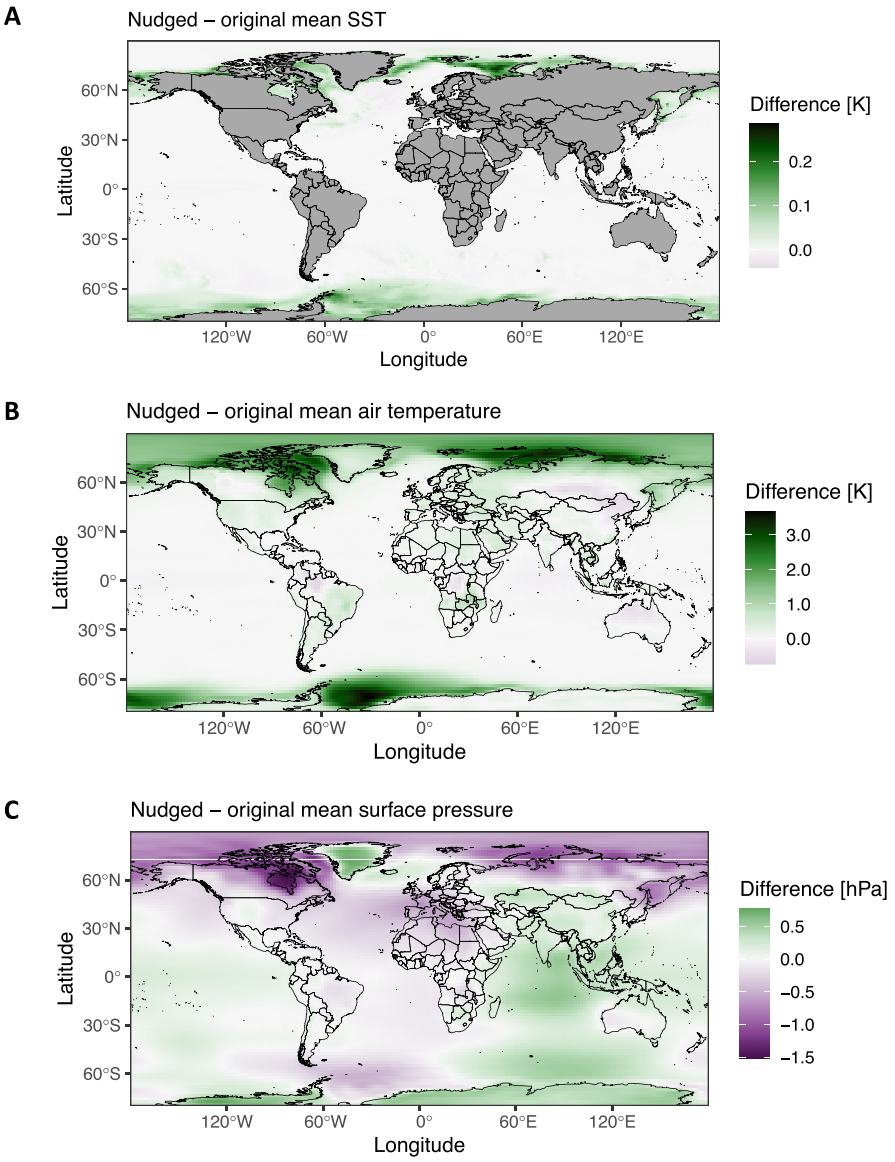

Figure 2. Change in mean annual SSTs, mean annual air temperature and mean annual surface pressure due to the nudging with rescaled SSTs relative to the original PI equilibrium simulation. Mean annual SSTs (A), air temperature (B) and surface pressure (C) have been temporally averaged over the whole length of simulation for each grid cell.

model run. Furthermore, the nudging results in weakening of positive relationships in the Northern North Atlantic and in more negative regression slopes which occur mostly at lower latitudes in the Southern Hemisphere. The strength of relationships in the Northern North Atlantic as well as the occurrence of more extensive areas with positive slopes in the Southern Hemisphere after the nudging are in agreement with observations (HadISST, figure 5(E); DCENT-I, figure S7). Improved model-observation match is also indicated by probability density functions (PDFs) of the slopes of the model runs and the HadISST data (figure 5(F)). The original model run shows a strong peak around slopes of 0.1 K K^{-1} with a maximum probability density of ∼2. After the nudging, the maximum probability density of this peak is more similar to the instrumental data (HadISST, ∼1). The nudged model run shows a fatter tail towards positive slopes than the original model run which is also more in agreement to the observations (HadISST and DCENT-I). However, the increase in negative relationships after the nudging resulting in a fatter tail towards negative slopes that the original model run (figure 5(F)) which is not found in the instrumental datasets. There, the HadISST data indicates barely any slopes lower than −0.6 K K^{-1} similar to the original model run. Regression slopes obtained from the DCENT-I dataset show a PDF which is centred over more positive slopes (∼0.6 K K^{-1}) and which includes much less negative slopes than the HadISST dataset and also the model (original and nudged).

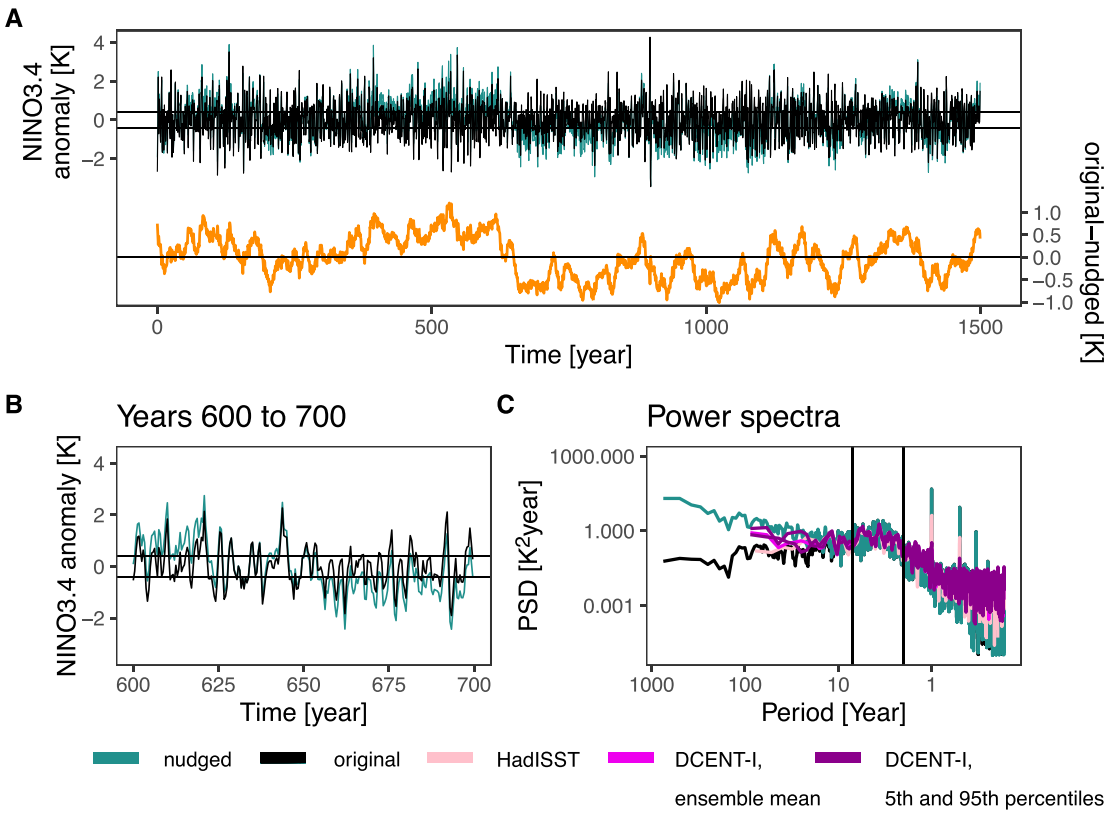

Figure 3. Comparison of NINO3.4 index between the nudged and original model simulation. (A). Time series of the difference in NINO3.4 index between the nudged and the original model simulation. Time series of the NINO3.4 index from the original and nudged model simulation (black and dark cyan, respectively) and the difference between both (orange line, bottom). Black horizontal lines indicate the 0.4 and −0.4 K thresholds for defining El Niño and La Niña events, respectively. (B). Same as top part of (A) but for the years 600–700. (C). Spectra of the monthly non-smoothed and non-normalised NINO3.4 index for the model runs and instrumental data (HadISST, Rayner *et al* 2003, pink; DCENT-I dataset, Chan *et al* 2025; ensemble mean in magenta, 5th and 95th percentiles of the ensemble both in dark magenta). Black vertical lines mark typical ENSO frequencies 1/7 and 1/2 yr^{-1}.

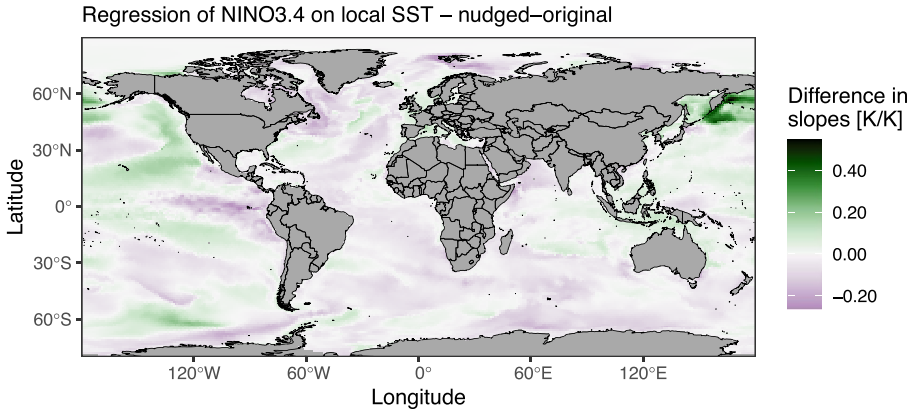

Figure 4. Spatial regression map of the annually averaged NINO3.4 index and local annual SSTs showing the difference in regression slopes between the original and nudged model simulation. We linearly regress the non-smoothed and non-normalised NINO3.4 index onto each local SST time series using least squares fitting. We use the annual mean of NINO3.4 and local SSTs to exclude an impact of the seasonal cycle on the resulting regression slope.

3.2.3. Centennial-scale variability

Nudging the AWI-ESM with the rescaled SSTs results in an increase of centennial-scale SST variability (between 1/600–1/100 yr^{-1}) in the nudged model run compared to the original run (>99.9% of grid points, figures 6(A) and (B)). Changes are heterogenous, with larger variability increases in areas showing lower variability in the original model run, and vice versa. This leads to a latitude-dependent increase, as we impose a relative larger change in centennial-scale variability at low-latitudes in the

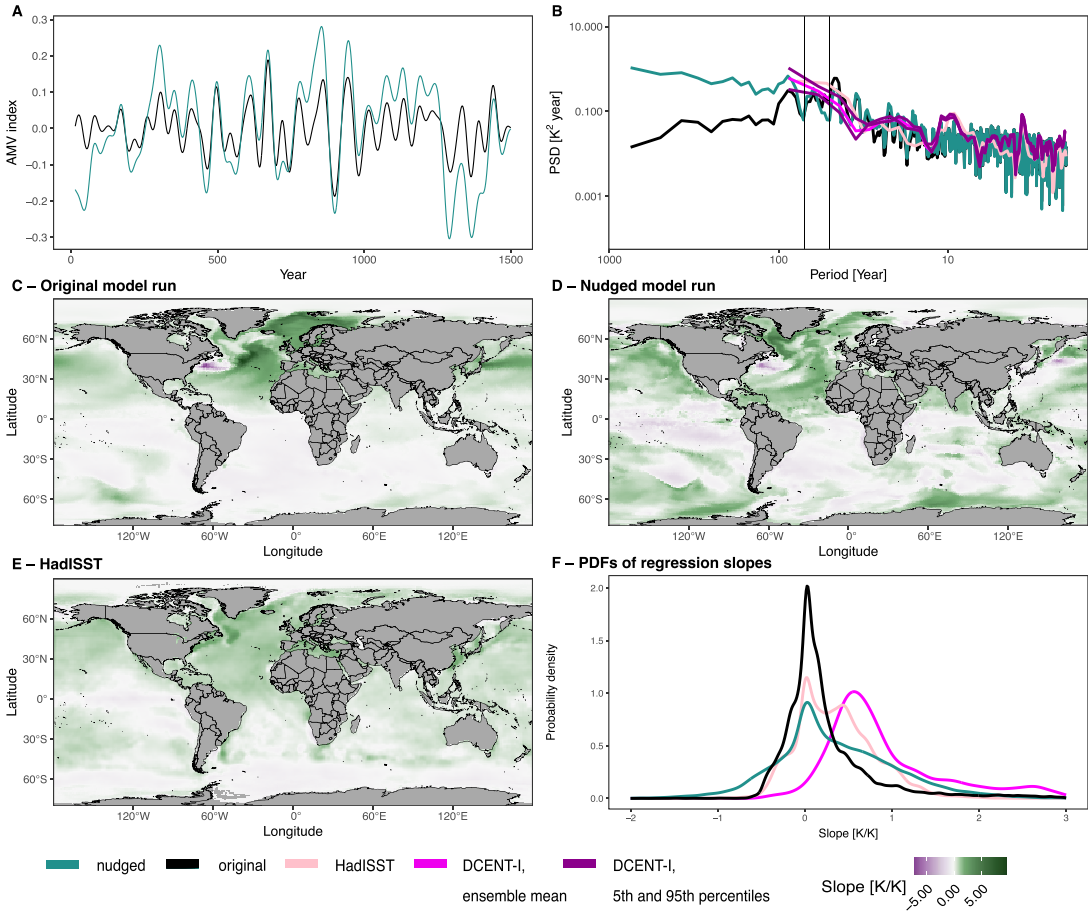

Figure 5. Comparison of the AMV index between the nudged (cyan) and original model simulation (black) and HadISST (pink, Rayner *et al* 2003). (A). Time series of the AMV index of the nudged and the original model simulation (linearly detrended and low-pass-filtered as described in section 2.4). (B). Spectra of the linearly detrended annual AMV index. Black vertical lines mark typical AMV frequencies 1/70 and 1/50 yr^{-1}. (C)–(E). Maps showing the slopes of linear regression (using least squares fitting) between the AMV index and local SSTs (annually averaged, linearly detrended and low-pass-filtered with cut-off frequency = 1/50 yr^{-1} similar to the AMV index). Maps for the original (C) and nudged (D) model simulation (1500 years of data) and the HadISST dataset ((E), 156 years of data). (F). Probability density functions (PDFs) of the local regression slopes shown in the maps in (C)–(E).

nudged compared to the original model simulation leading to a flatter latitudinal gradient (figure 6(C)). For example, we obtain large variability changes in the central Pacific ENSO region and also other tropical regions like the Indian Ocean but relatively small changes in centennial-scale variability occur in the northern North Atlantic (Gulf Stream and subpolar gyre) and Southern Ocean regions (North of Amundsen Sea and the Ross Gyre, figure 6).

3.3. Atmospheric variability

3.3.1. Surface pressure variability

Nudging the AWI-ESM with rescaled SSTs alters the (NAO, figure 7). The NAO time series from the nudged and original runs differ in their temporal evolution and show no significant correlation ($\tau = -0.006$, Kendall's tau test, $p > 0.1$; $r = -0.01$, Pearson's product-moment correlation, $p > 0.1$), consistent with internally generated atmospheric variability in a freely dynamically evolving atmosphere (figure 7(A)). Spectra of the NAO index from the original and nudged model simulations are similar for subcentennial timescales but diverge at frequencies below 1/200 yr^{-1} (figure 7(B)). This leads to multi-century persistent phases of more positive or more negative NAO states (figure 7(A)).

We find only minor changes in the spatial pattern of the relationship between NAO and local winter air temperature with the nudging with differences in local regression slopes between −0.03–0.07 hPa K^{-1} (10th and 90th percentiles, respectively; figure 8). Between Greenland and North America, both simulations show a positive relationship that strengthens in the nudged simulation, particularly over Baffin Island (figures S8(A) and (B)). In the eastern North Atlantic and northern Eurasia, a negative relationship is found, which intensifies with nudging over the Kara Sea near 60°E. However, overall

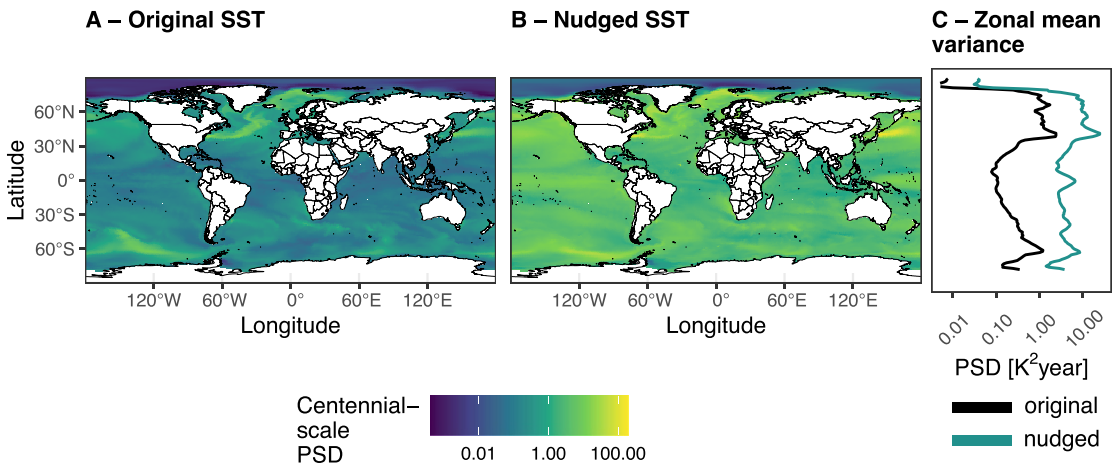

Figure 6. Change in centennial-scale variability (PSD in K^2 year, mean over frequency band 1/600–1/100 yr^{-1}) in the nudged modelled sea surface temperatures (SSTs) relative to the originally modelled SSTs due to the spectral rescaling. (A). Original AWI-ESM SST centennial-scale variability. (B). Nudged AWI-ESM SST centennial-scale variability. (C). Centennial-scale SST variability of nudged and original model simulations over latitudes.

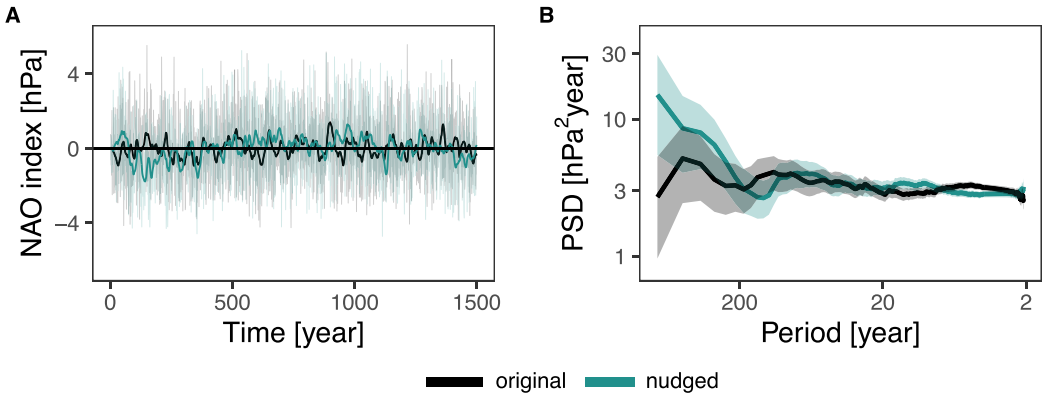

Figure 7. NAO index from the original (black) and nudged (dark cyan) model simulation. (A). Raw time series (thin lines) and low-pass filtered time series (cut-off frequency = 1/20 yr^{-1}). (B). Multitaper spectra (3 tapers) of raw time series.

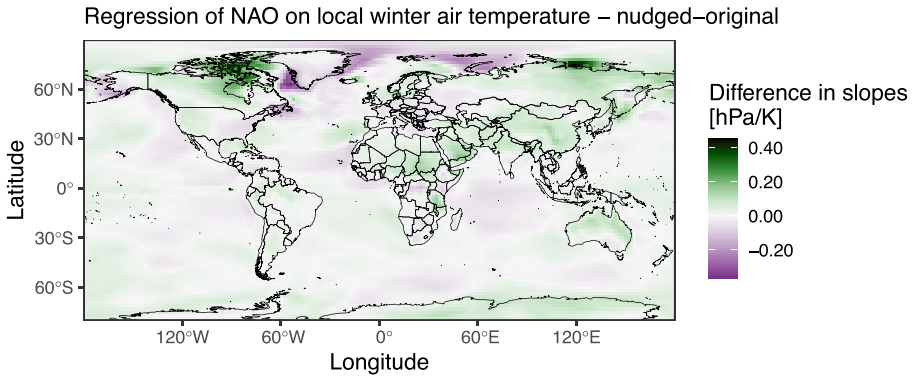

Figure 8. Spatial regression map of the NAO index and local winter air temperature showing the difference in regression slopes between the original and nudged model simulation. The NAO index is calculated as normalised winter (DJF) surface pressure difference between coordinates Stykkishólmur/Reykjavík (65.05° N, 22.74° W) and Lisbon, Portugal (38.71° N, 9.14° W) following Hurrell (1995).

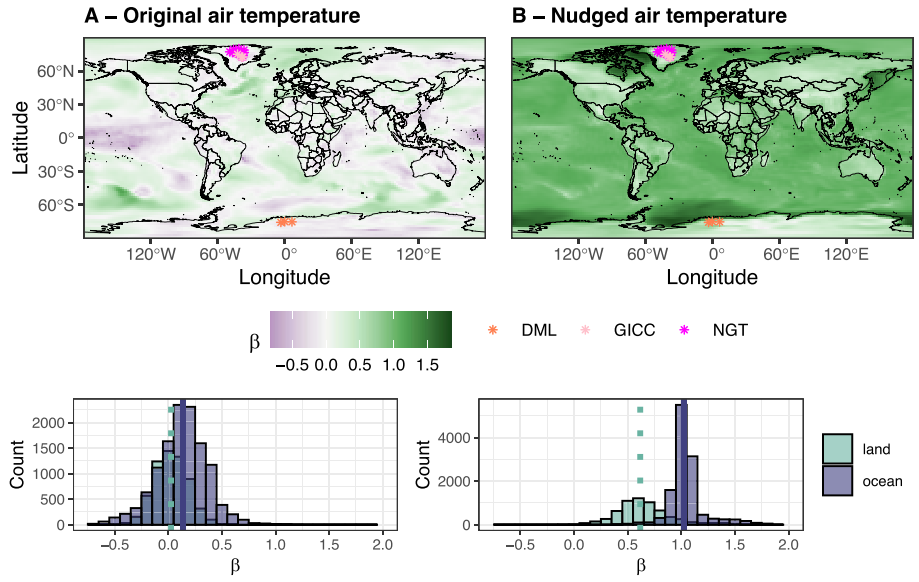

Figure 9. Spectral scaling of air temperature over multi-centennial to decadal timescales represented by a linear regression slope fitted over frequency bands 1/600–1/10 yr^{-1} in the original (A) and nudged (B) model simulation. Asterisks mark the ice core locations used for the 'DML', 'GICC' stack and the 'NGT' stack, in orange, pink and magenta colours, respectively. Histograms at the bottom show the distribution of β for the original (left) and nudged (right) model simulation over land and ocean areas with their medians indicated by vertical lines. To estimate β for the spectra, we fit a linear regression model using least squares fitting to the logarithm of the PSD and frequency, where a linear regression slope of —1 corresponds to a power-law exponent β of 1.

change in the range between 10% and 90% quantiles of absolute slopes due to the nudging is negligible (0.27 hPa K^{-1} and 0.28 hPa K^{-1} in the original and nudged simulation, respectively).

3.3.2. Surface air temperature variability

Nudging affects not only SST variability but also surface air temperature variability over both ocean and land. Nearly all grid cells (>99.9%) show stronger long-term persistence in their air temperature records, reflected in larger power-law exponents β (1/600–1/10 yr^{-1}; figure 9) for nearly all grid cells (>99.9%). While in the original model run, the distributions of β have similar median values of 0.1 and zero for ocean and land regions, respectively, (figure 9(A)), in the nudged model simulation they cluster around medians of 1 and 0.6, respectively (figure 9(B)). The β of 1 for local air temperature spectra in the ocean regions is due to the strong direct influence of the prescribed SSTs (with β of 1) on air temperatures in those regions, while the increase of β value from 0.1 to 0.6 over land regions demonstrates the influence of ocean variability on long-term terrestrial climate variability. Hence, the nudging results in a clear land-sea contrast in β which is lacking in the original model run (figures 9(A) and (B)).

We compare air temperature variability in the original and nudged model simulations to available ice-core-based temperature variability reconstructions from Greenland and Antarctica (locations of cores marked in figure 9).

In East Antarctica, the spectrum of annual air temperatures in the non-nudged model remains fairly flat, with values near 1 K^2 year across most timescales (black line in figure 10(A)). In contrast, the variability of the nudged model run shows higher values starting from multi-decadal timescales onward rising to values up to 4 K^2 year at multi-centennial timescales. As a result, the nudged-model-based spectrum more closely matches the ice-core-based variability estimate at multi-decadal to multi-centennial timescales (figure 10(A)).

This is also reflected in the distribution of the power-law exponents (fitted over frequencies 1/600–1/10 yr^{-1}) for realisations of the ice core spectrum available from Münch and Laepple (2018) and of the modelled air temperature spectra calculated using the same procedure (see 2.5 and figure caption for details, note that slopes are unaffected by the $\delta\,^{18}O$-temperature calibration). For the original model run, estimates of the exponent β of air temperature variability scatter around a median of zero, these increase to a median of 0.3 for the nudged model run. Thus, the spectral slope of the nudged model run is closer to the distribution of ice-core-based air temperature spectral slopes which has a median value of 0.5.

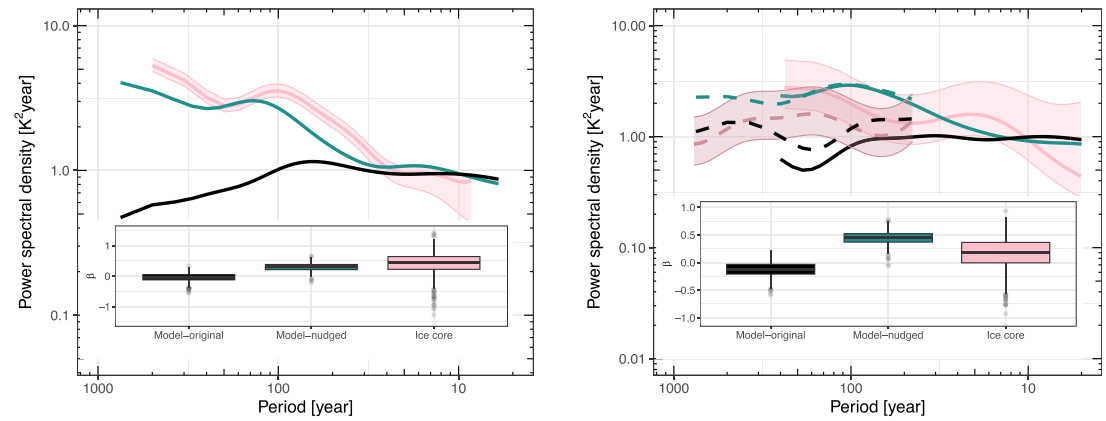

Figure 10. Comparison of modelled annual air temperature in East Antarctica (A) and Greenland (B) with ice-core-based variability reconstructions. Ice-core-based spectra are shown in pink (solid lines and dashed lines in (B) for the spectrum based on 'NGT' and 'GICC' locations, respectively). All spectra are log-smoothed with a smoother width of 0.1 for visualisation. The inset shows the distribution of linear regression slopes fitted over frequencies 1/600–1/10 year^{-1} (A) and 1/237–1/10 year^{-1} (B) of each spectrum realisation as boxplots (in (B) at 'NGT' locations). The confidence intervals (10th and 90th quantiles) are estimated from 1000 Monte Carlo realisations of each spectrum where a power-law fit (equation (1) is applied to the actual proxy signal and noise spectra, and the resulting power-law coefficients are used to generate surrogate signal and noise series. To estimate β for the spectra, we fit a linear regression model using least squares fitting to the logarithm of the PSD and frequency. (A). Shown with $\pm$ 1 SD uncertainty of the δ^{18}O-temperature conversion and time uncertainty by using the upper uncertainty bound from the time uncertainty corrected spectrum and the lower bound from the non-time uncertainty corrected spectrum. Highest frequencies are removed from the ice core spectra at 1/8.5 yr^{-1} due to too large uncertainties in the diffusion correction following Münch and Laepple (2018). (B). The 1 SD uncertainty is shown as shading and consists of the δ^{18}O-temperature conversion uncertainty for the 'GICC' spectrum and the δ^{18}O-temperature conversion uncertainty together with the time uncertainty for the 'NGT' spectrum (by using the upper uncertainty bound from the time-uncertainty-corrected spectrum and the lower bound from the non-time-uncertainty-corrected-spectrum).

For East Antarctica, the coastal regions in the DML sector and in the Wilkes Land sector show the highest increase in air temperature variability due to the nudging, where surrounding ocean regions also show the strongest increase in centennial-scale air temperature variability (figure S9).

For the Greenland 'NGT' stack locations, the air temperature spectrum from the nudging experiment (dark cyan solid line) shows stronger variability for multi-decadal to multi-centennial timescales than the original model simulation (black solid line, figure 10(B)). This brings it in closer agreement with the ice-core-based temperature variability (figure 10(B)). The ice-core-based temperature variability estimate from the 'GICC' stack (pink dashed line) indicates relatively stable variability for centennial to multi-centennial timescales. This is consistent with both the nudged-model-based estimate (dark cyan dashed line) and the one from the original model run (black dashed line). However, the mean variability in the nudged model run is higher than that inferred from the ice cores which is closer to the mean PSD level of the original model run.

In the original model run, estimated scaling exponents β of air temperature variability over frequencies 1/237–1/10 yr^{-1} cluster around -0.1, while in the nudged run they increase to about 0.4. This indicates a steeper spectrum in the nudged simulation and a positive relationship between variability and timescale, which is more similar to the ice-core-based estimate (median slope of about 0.2).

In Greenland, nudging leads to a stronger increase in air temperature variability in coastal regions and less increase in air temperature variability in central regions more distant to the coast (figure S9).

4. Discussion

We developed and tested a new framework to increase marine surface temperature variability in CMs by using paleo-proxy evidence to guide spectral rescaling of SSTs. This approach enhances long-term ocean memory while leaving short-term variability largely unaffected and retains realistic spatial coherence of low-frequency climate modes (section 3.2.2), providing a controlled way to test how persistent SST anomalies shape atmosphere dynamics. By doing so, we present an approach to directly address the suggested mismatch between paleoclimate reconstructions and ESMs, which tend to underestimate variability on multi-decadal to millennial timescales (Laepple and Huybers 2014a, 2014b, Cheung *et al* 2017, Dee *et al* 2017, Parsons *et al* 2017, Ellerhoff and Rehfeld 2021, Ellerhoff *et al* 2022, Hébert *et al* 2022, Laepple *et al* 2023).

Applying this framework in AWI-ESM, we investigated the impact of enhanced decadal- to centennial-scale ocean variability on the atmosphere in a fully coupled setting. The nudging increases low-frequency variability but preserves higher-frequency processes. For instance, the Niño 3.4 index shows unchanged amplitudes at typical ENSO frequencies (1/7–1/2 yr^{-1}) but diverges from the original simulation at longer timescales (>1/10 yr^{-1}). This leads to multi-decadal to multi-centennial periods dominated by stronger El Niño- or La Niña-like conditions. We compared the nudging results to observed variability estimated from instrumental datasets for up to multi-decadal timescales. Inconsistency in multi-decadal variability estimated from two utilised instrumental datasets, however, prevents conclusions on which model simulation agrees better with the observations. While enhanced ENSO amplitudes on long timescales would be qualitatively consistent with paleoclimate evidence of centennial-scale ENSO variability (e.g. Jiang *et al* 2023), the validity of the nudging result should be tested in a future study which carefully investigates variability in instrumental and paleo-data on multi-decadal to -centennial timescales. This will be especially important due to the global-scale impact of ENSO variability and could offer valuable insights for the ongoing debate on the role of internal variability versus forced response in driving the discrepancy between model-predicted and observed state of the equatorial Pacific (e.g. Watanabe *et al* 2021, Seager *et al* 2022).

The NAO also shows changes in low-frequency variability after the nudging. Both the original and nudged simulations exhibit white-noise–like behaviour from interannual to centennial scales, but the nudged run displays enhanced variability from ∼1/200 yr^{-1} onwards, resulting in multi-century periods of predominantly positive or negative NAO states. Similar long-lived NAO regimes have been reported in reconstructions, for example during the medieval climate anomaly (Trouet *et al* 2009, Wassenburg *et al* 2013, Baker *et al* 2015, Deininger *et al* 2017) or mid-to-late Holocene (Lamy *et al* 2006). However, different characteristics in the frequency domain (spectral scaling) of the NAO have been found. instrumental NAO indices (Jones *et al* 1997, 1999) and some paleo-reconstructions (Cook *et al* 2019) show a flat power spectrum which is more similar to the original AWI-ESM PI equilibrium simulation. Other reconstructions are characterised by a power spectrum with a negative slope that is more similar to our nudged simulation (Wunsch 1999, Luterbacher *et al* 2001, Ortega *et al* 2015; see comparison in Cook *et al* 2019). Furthermore, since differences between original and nudged model run only occur from ∼1/200 yr^{-1} onwards, testing agreement with instrumental datasets is not feasible due to limited coverage in time. Thus, we cannot infer whether the nudging results in a more realistic NAO but rather point out that it is not diverging from the original model simulation for sub-centennial timescale (period) where robust information from observations is available.

Nudging increases the amplitudes of the low-frequency climate variability mode AMV on centennial- to multi-centennial timescales. This timescale is not sufficiently covered by observations and the validity of the result needs to be tested with paleo-data from which variability has been carefully estimated. Nevertheless, spatial regressions show that the nudging method retains a realistic modelled spatial impact of the AMV although comparison to the instrumental SST datasets is difficult due to inconsistency in the distribution of spatial regression slopes obtained from them.

Nudging ocean surface temperatures with enhanced long-term memory increases atmospheric memory over land on long timescales. In Greenland, variability increases mainly at multi-decadal timescales and remains steady at lower-than-centennial timescales, while in East Antarctica the largest increases occur at centennial and longer timescales. Generally, the nudging not only increases variability but also alters its spatial distribution, with larger changes in coastal regions where ocean influence is strongest. Comparisons with ice-core-based reconstructions show improved agreement in both Greenland and Antarctica. In both cases, spectral slopes (β) shift closer to ice core estimates, demonstrating that enhanced ocean persistence can improve the match between modelled and reconstructed air temperature variability. However, the variability estimate from the nudged model run does not show an improved match to the ice-core-based estimate on multi-centennial timescale, which is based on the GICC05 Greenland ice cores. This could potentially be due to an overestimation of ocean influences at the GICC05 ice core locations, i.e. at the Greenland ice divide. Alternatively, only three GICC05 ice cores are available for estimation of the climate signal on these long timescales and we cannot exclude that uncertainty might lead to an underestimation of the variability, e.g. due to the spectral estimation error. Nevertheless, our results suggest that underrepresented ocean memory is a key driver of the model–data gap on multi-decadal to multi-centennial timescales in Antarctica and until centennial timescale in Greenland.

4.1. Limitations and next steps

Our spectral rescaling of the SST fields assumes a power-law relationship between variability (PSD) and timescale at each ocean grid cell. While this is consistent with paleo-data evidence (Laepple and Huybers

2013, Ellerhoff and Rehfeld 2021), it oversimplifies regional differences, particularly in sea ice covered regions or open ocean areas where proxy constraints are sparse. To refine the use of nudged-model-based variability estimates, future work could map spatial patterns of long-term variability or spectral scaling directly from paleo-observations using statistical methods to better inform model experiments. Another approach could be to utilise empirical orthogonal functions to decompose SST variability into uncorrelated components representing statistically independent patterns. In that way, the changes over time of an independent spatial pattern (principal component), e.g. one which has been identified to represent a certain variability mode, could be rescaled based on information from paleo-data and the change then projected back onto each grid cell using its eigenvectors.

Despite this simplification, our approach already improves large-scale patterns of low-frequency variability. In particular, it reduces the latitudinal gradient of centennial-scale SST variability (figure 6), addressing a mismatch identified in paleo-data: models underestimate tropical centennial-scale SST variability by one to two orders of magnitude (Laepple and Huybers 2014b), which is similar in magnitude to the increase we impose in tropical regions through rescaling and nudging. The underestimation in mid to high latitudes is smaller, as is the rescaling we impose. Targeted nudging experiments that increase the spectral slope β only in selected regions, such as the central Pacific ENSO domain or the Southern Ocean, could help to isolate and test the role of specific ocean basins in driving long-term variability across the climate system. While the nudging would be regionally constrained in such an approach, the model would be enabled to evolve freely elsewhere. To avoid shocks and discontinuities, this necessitates a transition zone in which the relaxation factor decays to zero. This method has been applied in other studies, e.g. Martin and Rodríguez (2024) and Ortega *et al* (2017).

The spectral rescaling methodology utilised here raises mean SSTs in sea-ice regions, since positive long-term fluctuations are enhanced while negative ones are truncated at the freezing threshold. As long-term SST variability in these regions is poorly constrained, future paleo-data will be key to refining the approach, and our current results in the sea ice regions should be interpreted with caution. Furthermore, we consider the response of the subsurface ocean to the nudging as being unrealistic as we impose large artificial changes to SST not based on physical principles. This also limits the exploitation of the nudged model simulation in order to investigate physical origins of low-frequency variability modes driven by ocean-atmosphere interaction.

A next step would be to validate the nudged simulation results by comparing to hydroclimate proxy reconstructions, e.g. related to the Pacific Walker circulation strength (Falster *et al* 2023) or ITCZ mean position (Haug *et al* 2001). The AWI-ESM has been also enhanced with a comprehensive water isotopes module. Conducting a nudging experiment while simulating water isotopes would allow for a direct comparison of the simulated isotope variability changes due to nudging with corresponding paleoclimate data. This would enable evaluation of the results with a larger amount of data, for example, isotope records representative of rainfall amount. Similarly, AWI-ESM versions with coupling to biogeochemistry (Ye *et al* 2025) and ice sheets have been developed (Ackermann *et al* 2020) and could be utilised in future studies as soon as long control runs from these models are available.

In this study, we apply our method under PI boundary conditions without external forcing. Because of the lack of external forcing, the nudged simulation cannot be directly compared with observed climate time series in the time domain. Future applications could combine paleo-informed nudging with forced simulations of past or future climate, for example by superimposing the nudged variability onto the forced response estimated from a model ensemble mean. However, the paleo-observations, which inform our nudging approach, record local climate variability resulting from both internal variability and response to external forcing. Hence, future work needs to be conducted to refine the imposed long-term variability by accounting for the response to forcing as captured by the paleo-data, e.g. using empirical-orthogonal-function-based approaches (Ayache *et al* 2018).

Beyond the pilot implementation presented here, the nudging framework offers a range of applications across paleoclimate and future climate research. In the paleoclimate context, simulations with enhanced and potentially more realistic variability provide a powerful testbed for pseudo-proxy experiments, where reconstruction methods can be benchmarked against model output with variability levels closer to those inferred from proxy records. The framework can also serve to create ensemble members as priors for paleo-data assimilation frameworks, helping to reconstruct the past evolution of the climate system including uncertainties related to natural variability.

The approach is equally relevant for detection and attribution studies. Current attribution methods rely on separating forced responses from natural variability, yet CMs may underestimate low-frequency variability. By explicitly increasing long-term variability, nudged simulations can be used to test the sensitivity of attribution results to different assumptions about natural variability, providing a more

robust basis for assessing anthropogenic influence for example in the attribution of weather extremes (Harrington *et al* 2021). This will require performing a similar nudging simulation as presented here but on daily timescale in order to investigate effects of the enhanced low-frequency ocean surface temperature variability on extreme events.

Finally, applications extend to projections of future climate risks. Existing IPCC-class models may underestimate the persistence of natural variability, which could affect the frequency and duration of extreme events, multi-decadal trends, or prolonged anomalies. Incorporating paleo-informed nudged simulations into future scenario ensembles could yield a wider, more realistic, envelope of outcomes, particularly for regional risks linked to extremes and for long-term processes such as ice-sheet–ocean interactions or multi-decadal droughts.

5. Conclusions and outlook

We introduced a paleo-informed nudging approach that enhances long-term variability in a coupled ESM. This framework addresses the underestimation of low-frequency variability in models as suggested by previous studies (Laepple and Huybers 2014a, 2014b, Cheung *et al* 2017, Dee *et al* 2017, Parsons *et al* 2017, Ellerhoff and Rehfeld 2021, Ellerhoff *et al* 2022, Hébert *et al* 2022, Laepple *et al* 2023). By rescaling SSTs to match spectral characteristics from paleoclimate data, the nudging experiment produces atmospheric variability that aligns more closely with ice-core reconstructions, highlighting the central role of ocean memory in driving land temperatures at multi-decadal to millennial timescales (Dommenget 2009, Hébert *et al* 2022).

It remains unclear whether the underestimated supradecadal temperature variability in CMs, inferred from paleo-data, arises from missing external forcing, an inadequate simulated response to that forcing, or insufficient internal variability. Until these biases of model simulations are resolved, the here presented nudging approach offers a tractable way to generate physically consistent scenarios with natural variability levels suggested by proxies. This provides a valuable tool for climate attribution studies and for constraining regional climate projections that account for the full spectrum of internal variability. In this way, the approach complements ongoing efforts to improve coupled model physics and data assimilation of low-frequency ocean variability.

Ocean variability is also an important driver of polar ice sheet mass balance (e.g. Jenkins *et al* 2018, Holland *et al* 2019, Wood *et al* 2021) which is a large source of uncertainty in future sea-level predictions (Tsai *et al* 2017, Robel *et al* 2019). Our approach with paleo-informed natural ocean and climate variability can therefore serve as input for ice sheet models to provide realistic future scenarios and probabilities for a future crossing of thresholds (e.g. the melt-elevation feedback; Levermann and Winkelmann 2016).

Our study shows that paleo-informed surface ocean temperature nudging improves consistency with air temperature proxy evidence and provides a framework to explore how potentially underestimated long-term natural variability may shape future climate risks.

Data and code availability statement

The 'NGT' stack can be obtained using a package provided in Hörhold *et al* (2023): https://zenodo.org/records/7178657. 'GICC' ice core data can be found under the following link: www.iceandclimate.nbi.ku.dk/data/. The 'DML' stack can be obtained using a package provided in Münch and Laepple (2018): https://github.com/EarthSystemDiagnostics/proxysnr. The HadISST dataset is available from the Met Office (www.metoffice.gov.uk/hadobs/hadisst/) and the DCENT-I ensemble can be downloaded under this link: https://duochanatharvard.github.io/#DCENT. All model output and code to reproduce the analysis and figures is available under the following DOI: https://https://doi.org/10.5281/zenodo.17120542 (Skiba 2026).

Supporting information is available at https://doi.org/10.1088/2752-5295/ae3a08/data1.

Acknowledgment

The research of VS, LA, YS and GL was supported by Federal Ministry of Research, Technology and Space (BMFTR) as Research for Sustainability initiative (FONA); www.fona.de through the Palmod Project (FKZ: 01LP2310B, 01LP2313A, 01LP2309A and 01LP1917A, respectively). GL is supported by the ERC synergy Grant 'i2B' (Grant No. 101118519). AMD was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation); Project Number 468685498 (A.M.D.);

SPP 2299/Project Number 441832482. We further acknowledge the extensive efforts of the ice core research community and thank the respective project teams for making the data publicly accessible. We acknowledge the use of ChatGPT (OpenAI) for language polishing and production of code for figure design. The authors are solely responsible for the content and interpretations presented in this article. We thank the three reviewers for their constructive comments.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

V Skiba 0000-0002-0496-120X
Conceptualization (equal), Data curation (lead), Formal analysis (lead), Investigation (lead), Methodology (lead), Software (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (lead)

L Ackermann 0000-0001-6643-0714
Methodology (supporting), Writing – original draft (supporting), Writing – review & editing (supporting)

N Hirsch 0000-0001-6079-4699
Data curation (supporting), Formal analysis (supporting), Methodology (supporting), Software (supporting), Visualization (supporting), Writing – original draft (supporting), Writing – review & editing (supporting)

Y Sun
Data curation (supporting), Resources (supporting), Writing – original draft (supporting), Writing – review & editing (supporting)

A M Dolman 0000-0002-6481-966X
Conceptualization (supporting), Investigation (supporting), Methodology (supporting), Software (equal), Writing – review & editing (supporting)

G Lohmann 0000-0003-2089-733X
Conceptualization (supporting), Methodology (supporting), Resources (supporting), Writing – review & editing (supporting)

T Laepple
Conceptualization (lead), Investigation (supporting), Methodology (supporting), Resources (supporting), Visualization (supporting), Writing – original draft (supporting), Writing – review & editing (supporting)

References

Ackermann L, Danek C, Gierz P and Lohmann G 2020 AMOC recovery in a multicentennial scenario using a coupled atmosphere-ocean-ice sheet model *Geophys. Res. Lett.* **47** e2019GL086810

Ayache M, Swingedouw D, Mary Y, Eynaud F and Colin C 2018 Multi-centennial variability of the AMOC over the Holocene: a new reconstruction based on multiple proxy-derived SST records *Glob. Planet. Change* **170** 172–89

Baker A C, Hellstrom J C, Kelly B F J, Mariethoz G and Trouet V 2015 A composite annual-resolution stalagmite record of North Atlantic climate over the last three millennia *Sci. Rep.* **5** 10307

Bakker P, Clark P U, Golledge N R, Schmittner A and Weber M E 2017 Centennial-scale Holocene climate variations amplified by Antarctic Ice Sheet discharge *Nature* **541** 72–76

Bamston A G, Chelliah M and Goldenberg S B 1997 Documentation of a highly ENSO-related SST region in the equatorial Pacific: research note *Atmos.-Ocean* **35** 367–83

Blanusa M L, López-Zurita C J and Rasp S 2023 Internal variability plays a dominant role in global climate projections of temperature and precipitation extremes *Clim. Dyn.* **61** 1931–45

Bonan G B 2008 Forests and climate change: forcings, feedbacks, and the climate benefits of forests *Science* **320** 1444–9

Braconnot P, Zhu D, Marti O and Servonnat J 2019 Strengths and challenges for transient Mid-to Late Holocene simulations with dynamical vegetation *Clim. Past* **15** 997–1024

Caillet J, Jourdain N C, Mathiot P, Gillet-Chaulet F, Urruty B, Burgard C, Amory C, Chekki M and Kittel C 2025 Uncertainty in the projected Antarctic contribution to sea level due to internal climate variability *Earth Syst. Dyn.* **16** 293–315

Calel R, Chapman S C, Stainforth D A and Watkins N W 2020 Temperature variability implies greater economic damages from climate change *Nat. Commun.* **11** 5028

Chan D, Chan S, Siddons J, Cable A, Faulkner A, Cornes R and Huybers P 2025 DCENT-I: a globally infilled extension of the dynamically consistent ensemble of temperature dataset (*eartharxiv)* (https://doi.org/10.7910/DVN/ZY0WM8)

Cheung A H, Mann M E, Steinman B A, Frankcombe L M, England M H and Miller S K 2017 Comparison of low-frequency internal climate variability in CMIP5 models and observations *J. Clim.* **30** 4763–76
Cook E R, Kushnir Y, Smerdon J E, Williams A P, Anchukaitis K J and Wahl E R 2019 A Euro-Mediterranean tree-ring reconstruction of the winter NAO index since 910 CE *Clim. Dyn.* **53** 1567–80
Danabasoglu G, Yeager S G, Bailey D, Behrens E, Bentsen M, Bi D and Wang Q 2016 North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part II: inter-annual to decadal variability *Ocean Model.* **97** 65–90
Danilov S, Sidorenko D, Wang Q and Jung T 2017 The finite-volumE sea ice–ocean model (fesom2) *Geosci. Model Dev.* **10** 765–89
Dee S G, Parsons L A, Loope G R, Overpeck J T, Ault T R and Emile-Geay J 2017 Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: implications for multi-decadal variability *Earth Planet. Sci. Lett.* **476** 34–46
Deininger M, McDermott F, Mudelsee M, Werner M, Frank N and Mangini A 2017 Coherency of late Holocene European speleothem $\delta^{18}O$ records linked to North Atlantic Ocean circulation *Clim. Dyn.* **49** 595–618
Dommenget D 2009 The ocean's role in continental climate variability and change *J. Clim.* **22** 4939–52
Ellerhoff B, Kirschner M J, Ziegler E, Holloway M D, Sime L and Rehfeld K 2022 Contrasting state-dependent effects of natural forcing on global and local climate variability *Geophys. Res. Lett.* **49** e2022GL098335
Ellerhoff B and Rehfeld K 2021 Probing the timescale dependency of local and global variations in surface air temperature from climate simulations and reconstructions of the last millennia *Phys. Rev. E* **104** 064136
Enfield D B, Mestas-Nuñez A M and Trimble P J 2001 The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US *Geophys. Res. Lett.* **28** 2077–80
Falster G, Konecky B, Coats S and Stevenson S 2023 Forced changes in the Pacific Walker circulation over the past millennium *Nature* **622** 93–100
Fettweis X, Box J E, Agosta C, Amory C, Kittel C, Lang C, van As D, Machguth H and Gallée H 2017 Reconstructions of the 1900–2015 Greenland ice sheet surface mass balance using the regional climate MAR model *Cryosphere* **11** 1015–33
Hagemann S and Dümenil L 1997 A parametrization of the lateral waterflow for the global scale *Clim. Dyn.* **14** 17–31
Harrington L J, Schleussner C-F and Otto F E L 2021 Quantifying uncertainty in aggregated climate change risk assessments *Nat. Commun.* **12** 7140
Haug G H, Hughen K A, Sigman D M, Peterson L C and Röhl U 2001 Southward migration of the intertropical convergence zone through the Holocene *Science* **293** 1304–8
Hawkins E and Sutton R 2009 The potential to narrow uncertainty in regional climate predictions *Bull. Am. Meteorol. Soc.* **90** 1095–108
Hébert R, Herzschuh U and Laepple T 2022 Millennial-scale climate variability over land overprinted by ocean temperature fluctuations *Nat. Geosci.* **15** 899–905
Holland P R, Bracegirdle T J, Dutrieux P, Jenkins A and Steig E J 2019 West Antarctic ice loss influenced by internal climate variability and anthropogenic forcing *Nat. Geosci.* **12** 718–24
Hopcroft P O and Valdes P J 2021 Paleoclimate-conditioning reveals a North Africa land–atmosphere tipping point *PNAS* **118** e2108783118
Hörhold M, Münch T, Weißbach S, Kipfstuhl S, Freitag J, Sasgen I, Lohmann G, Vinther B and Laepple T 2023 Modern temperatures in central–north Greenland warmest in past millennium *Nature* **613** 503–7
Hurrell J W 1995 Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation *Science* **269** 676–9
Hurrell J W, Delworth T, Danabasoglu G, Drange H, Griffies S, Holbrook N and Tribbia J 2010 Decadal climate prediction: opportunities and challenges *Proc. OceanObs' 09: Sustained Ocean Observations and Information for Society* pp 521–33
Jenkins A, Shoosmith D, Dutrieux P, Jacobs S, Kim T W, Lee S H, Ha H K and Stammerjohn S 2018 West Antarctic Ice Sheet retreat in the Amundsen Sea driven by decadal oceanic variability *Nat. Geosci.* **11** 733–8
Jiang S, Zhou X, Sachs J P, Li Z, Tu L, Lin Y, Liu X, Chen A and Shen Y 2023 The mean state of the tropical Pacific Ocean differed between the Medieval Warm Period and the Industrial Era *Commun. Earth Environ.* **4** 74
Jones P D, Davies T D, Lister D H, Slonosky V, Jónsson T, Bärring L and Beck C 1999 Monthly mean pressure reconstructions for Europe for the 1780–1995 period *Int. J. Climatol.* **19** 347–64
Jones P D, Jónsson T and Wheeler D 1997 Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland *Int. J. Climatol.* **17** 1433–50
Jüling A, Von Der Heydt A and Dijkstra H A 2021 Effects of strongly eddying oceans on multidecadal climate variability in the Community Earth System Model *Ocean Sci.* **17** 1251–71
Keenlyside N, Latif M, Botzet M, Jungclaus J and Schulzweida U 2005 A coupled method for initializing El Niño–Southern Oscillation forecasts using sea surface temperature *Tellus A* **57** 340–56
Kong X, Jin J, Wang A, Dong X, Bi X and Zhang H 2024 Impacts of nudged sea surface temperature on tropical precipitation, moisture, and vertical velocity in an earth system model *J. Clim.* **37** 457–73
Laepple T and Huybers P 2013 Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability *Earth Planet. Sci. Lett.* **375** 418–29
Laepple T and Huybers P 2014a Global and regional variability in marine surface temperatures *Geophys. Res. Lett.* **41** 2528–34
Laepple T and Huybers P 2014b Ocean surface temperature variability: large model–data differences at decadal and longer periods *Proc. Natl Acad. Sci. USA* **111** 16682–7
Laepple T, Jewson S and Coughlin K 2008 Interannual temperature predictions using the CMIP3 multi-model ensemble mean *Geophys. Res. Lett.* **35** L10701
Laepple T, Ziegler E, Weitzel N, Hébert R, Ellerhoff B, Schoch P and Rehfeld K 2023 Regional but not global temperature variability underestimated by climate models at supradecadal timescales *Nat. Geosci.* **16** 958–66
Laguë M M, Bonan G B and Swann A L 2019 Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and uncoupled land–atmosphere system *J. Climate* **32** 5725–44
Lamy F, Arz H W, Bond G C, Bahr A and Pätzold J 2006 Multicentennial-scale hydrological changes in the Black Sea and northern Red Sea during the Holocene and the Arctic/North Atlantic Oscillation *Paleoceanography* **21** PA1008
Levermann A and Winkelmann R 2016 A simple equation for the melt elevation feedback of ice sheets *Cryosphere* **10** 1799–807
Levitus S, Locarnini R A, Boyer T P, Mishonov A V, Antonov J I, Garcia H E and Seidov D 2010 *World Ocean Atlas 2009* (National oceanographic data center, ocean climate laboratory)
Lohmann G, Butzin M, Eissner N, Shi X and Stepanek C 2020 Abrupt climate and weather changes across timescales *Paleoceanogr. Paleoclimatol.* **35** e2019PA003782
Luterbacher J *et al* 2001 Extending North Atlantic oscillation reconstructions back to 1500 *Atmos. Sci. Lett.* **2** 114–24

Manda A, Hirose N and Yanagi T 2005 Feasible method for the assimilation of satellite-derived SST with an ocean circulation model *J. Atmos. Ocean. Technol.* **22** 746–56
Martin G M and Rodríguez J M 2024 Using regional relaxation experiments to understand the development of errors in the Asian Summer Monsoon *Weather Clim. Dyn.* **5** 711–31
Münch T and Laepple T 2018 What climate signal is contained in decadal-to centennial-scale isotope variations from Antarctic ice cores? *Clim. Past.* **14** 2053–70
Ortega P, Guilyardi E, Swingedouw D, Mignot J and Nguyen S 2017 Reconstructing extreme AMOC events through nudging of the ocean surface: a perfect model approach *Clim. Dyn.* **49** 3425–41
Ortega P, Lehner F, Swingedouw D, Masson-Delmotte V, Raible C C, Casado M and Yiou P 2015 A model-tested North Atlantic Oscillation reconstruction for the past millennium *Nature* **523** 71–74
Otto-Bliesner B L, Braconnot P, Harrison S P, Lunt D J, Abe-Ouchi A, Albani S and Zhang Q 2017 The PMIP4 contribution to CMIP6—Part 2: two interglacials, scientific objective and experimental design for Holocene and last interglacial simulations *Geosci. Model Dev.* **10** 3979–4003
Parsons L A, Loope G R, Overpeck J T, Ault T R, Stouffer R and Cole J E 2017 Temperature and precipitation variance in CMIP5 simulations and paleoclimate records of the last millennium *J. Clim.* **30** 8885–912
Percival D B and Walden A T 1993 *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge University Press)
Rackow T, Goessling H F, Jung T, Sidorenko D, Semmler T, Barbi D and Handorf D 2018 Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part II: climate variability *Clim. Dyn.* **50** 2369–94
Rasmussen S O, Bigler M, Blockley S P, Blunier T, Buchardt S L, Clausen H B and Winstrup M 2014 A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy *Quat. Sci. Rev.* **106** 14–28
Rayner N A, Parker D E, Horton E B, Folland C K, Alexander L V, Rowell D P, Kent E C and Kaplan A 2003 Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century *J. Geophys. Res. Atmos.* **108** 4407
Reick C H, Gayler V, Goll D, Hagemann S, Heidkamp M, Nabel J E and Wilkenskjeld S 2021 JSBACH 3—The land component of the MPI Earth System Model: documentation of version 3.2 (available at: https://pure.mpg.de/rest/items/item_3279802_26/component/file_3316522/content) (Accessed 2 July 2025)
Robel A A, Seroussi H and Roe G H 2019 Marine ice sheet instability amplifies and skews uncertainty in projections of future sea-level rise *Proc. Natl Acad. Sci. USA* **116** 14887–92
Schwarzwald K and Lenssen N 2022 The importance of internal climate variability in climate impact projections *Proc. Natl Acad. Sci. USA* **119** e2208095119
Seager R, Henderson N and Cane M 2022 Persistent discrepancies between observed and modeled trends in the tropical Pacific Ocean *J. Clim.* **35** 4571–84
Seierstad I K, Abbott P M, Bigler M, Blunier T, Bourne A J, Brook E and Vinther B M 2014 Consistently dated records from the Greenland GRIP, GISP2 and NGRIP ice cores for the past 104 ka reveal regional millennial-scale $\delta^{18}O$ gradients with possible Heinrich event imprint *Quat. Sci. Rev.* **106** 29–46
Sein D V, Danilov S, Biastoch A, Durgadoo J V, Sidorenko D, Harig S and Wang Q 2016 Designing variable ocean model resolution based on the observed ocean variability *J. Adv. Model. Earth Syst.* **8** 904–16
Semmler T, Danilov S, Gierz P, Goessling H F, Hegewald J, Hinrichs C and Jung T 2020 Simulations for CMIP6 with the AWI climate model AWI-CM-1-1 *J. Adv. Model. Earth Syst.* **12** e2019MS002009
Shi X and Lohmann G 2016 Simulated response of the mid-Holocene Atlantic meridional overturning circulation in ECHAM6-FESOM/MPIOM *J. Geophys. Res.* **121** 6444–69
Shi X, Werner M, Krug C, Brierley C M, Zhao A, Igbinosa E and Lohmann G 2022a Calendar effects on surface air temperature and precipitation based on model-ensemble equilibrium and transient simulations from PMIP4 and PACMEDY *Clim. Past.* **18** 1047–70
Shi X, Werner M, Wang Q, Yang H and Lohmann G 2022b Simulated mid-Holocene and last interglacial climate using two generations of AWI-ESM *J. Clim.* **35** 7811–31
Shindell D T, Schmidt G A, Mann M E, Rind D and Waple A 2001 Solar forcing of regional climate change during the Maunder Minimum *Science* **294** 2149–52
Sidorenko D, Goessling H, Koldunov N, Scholz P, Danilov S, Barbi D and Jung T 2019 Evaluation of FESOM2.0 coupled to ECHAM6.3: preindustrial and HighResMIP simulations *J. Adv. Model. Earth Syst.* **11** 3794–815
Sidorenko D, Rackow T, Jung T, Semmler T, Barbi D, Danilov S and Wang Q 2015 Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate *Clim. Dyn.* **44** 757–80
Skiba V 2026 Code and data to produce the figures for the manuscript "Enhancing climate variability in earth system models through paleo-informed nudging of sea surface temperatures" *Zenodo* (https://doi.org/10.5281/zenodo.17120542)
Stenni B, Curran M A, Abram N J, Orsi A, Goursaud S, Masson-Delmotte V and Frezzotti M 2017 Antarctic climate variability on regional and continental scales over the last 2000 years *Clim. Past.* **13** 1609–34
Stevens B, Giorgetta M, Esch M, Mauritsen T, Crueger T, Rast S and Roeckner E 2013 Atmospheric component of the MPI-M Earth System Model: ECHAM6 *J. Adv. Model. Earth Syst.* **5** 146–72
Trenberth K E and Stepaniak D P 2001 Indices of El Niño evolution *J. Clim.* **14** 1697–701
Trouet V, Esper J, Graham N E, Baker A, Scourse J D and Frank D C 2009 Persistent positive North Atlantic Oscillation mode dominated the medieval climate anomaly *Science* **324** 78–80
Tsai C-Y, Forest C E and Pollard D 2017 Assessing the contribution of internal climate variability to anthropogenic changes in ice sheet volume *Geophys. Res. Lett.* **44** 6261–8
Tsai C-Y, Forest C E and Pollard D 2020 The role of internal climate variability in projecting Antarctica's contribution to future sea-level rise *Clim. Dyn.* **55** 1875–92
Wassenburg J A, Immenhauser A, Richter D K, Niedermayr A, Riechelmann S, Fietzke J and Esper J 2013 Moroccan speleothem and tree ring records suggest a variable positive state of the North Atlantic Oscillation during the medieval warm period *Earth Planet Sci. Lett.* **375** 291–302
Watanabe M, Dufresne J-L, Kosaka Y, Mauritsen T and Tatebe H 2021 Enhanced warming constrained by past trends in equatorial Pacific sea surface temperature gradient *Nat. Clim. Change* **11** 33–37

Wood M, Rignot E, Fenty I, An L, Bjørk A, van den Broeke M and Zhang H 2021 Ocean forcing drives glacier retreat in Greenland *Sci. Adv.* **7** eaba7282

Wunsch C 1999 The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations *Bull. Am. Meteorol. Soc.* **80** 245–56

Ye Y, Munhoven G, Köhler P, Butzin M, Hauck J, Gürses Ö and Völker C 2025 FESOM2. 1-REcoM3-MEDUSA2: an ocean–sea ice–biogeochemistry model coupled to a sediment model *Geosci. Model Dev.* **18** 977–1000