https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

Paleo-NorCPM: A framework toward fully coupled ocean-atmosphere paleoclimate reanalysis over the past centuries

Quentin Dalaiden[1], François Counillon[1], Lea Svendsen[2], Ingo Bethke[2], and Noel Keenlyside[2,1]

[1]Nansen Environmental and Remote Sensing Center and Bjerknes Centre for Climate Research, Bergen, Norway
[2]Geophysical Institute, Bjerknes Centre for Climate Research, University of Bergen, Bergen, Norway

Correspondence: Quentin Dalaiden (quentin.dalaiden@nersc.no)

Abstract.

We present a new paleo reanalysis (https://zenodo.org/records/19660743; Dalaiden, 2026) based on an adaptation of the Norwegian Climate Prediction Model (NorCPM), in which the Norwegian Earth System Model is equipped with an Ensemble Kalman Filter to assimilate hundreds of annually resolved proxy records, including tree rings, corals, and ice cores, over the last 500 years. First, an offline 30-member reanalysis is produced. This is followed by a dynamically coupled 10-member reanalysis obtained through atmospheric wind nudging from the offline reconstruction, allowing the ocean to respond to reconstructed atmospheric circulation variability. This approach enables an explicit representation of ocean dynamics and associated atmosphere-ocean feedback, which are typically indirectly represented in existing multi-century reanalyses. The offline reanalysis successfully reproduces large-scale atmospheric variability and hydroclimate variability over the 20th century across the tropics, mid-latitudes, and polar regions, and compares well with existing paleo reconstructions. The wind-nudged ensemble further captures ocean variability and exhibits more pronounced multi-decadal variability, highlighting the role of wind-driven ocean adjustment in shaping low-frequency climate variability. Challenges remain, particularly in the Southern Ocean where the wind-nudged ensemble shows strong sensitivity to wind variability. Nevertheless, these results demonstrate a promising pathway toward fully coupled ocean-atmosphere reanalyses spanning the past centuries, opening new opportunities to investigate the mechanisms underlying low-frequency variability and coupled feedback.

1 Introduction

Providing information on past climate variability is essential to place the observed climate changes over the past few decades into a longer-term context (e.g., Mann et al., 2008; Thorne et al., 2026), and to better understand their underlying physical drivers. Climate reconstructions enable quantitative assessment of the relative contributions of internal variability (arising from natural fluctuations of the climate system) and externally forced variability (driven by factors such as volcanic eruptions and anthropogenic greenhouse gas emissions) to observed climate changes. For instance, global mean surface temperature since the 20$^{\text{th}}$ century has been characterized by a long-term anthropogenic warming trend, superimposed with pronounced multi-decadal variability (e.g., Kosaka and Xie, 2013; England et al., 2014; Dai et al., 2015). These low-frequency fluctuations primarily reflect internal climate variability, particularly associated with changes in the state of the Pacific and Atlantic Oceans.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Depending on its phase, this internal variability can either amplify or temporarily offset the anthropogenic warming signal. A prominent example is the early 21st century warming hiatus, which has been linked to anomalously cool conditions in the tropical Pacific Ocean (e.g., England et al., 2014; Dai et al., 2015). Owing to the large thermal inertia of the ocean, such anomalies can persist for years to decades, exerting substantial influence on global and regional climate (e.g., Henley, 2017), including hydroclimate extremes such as the 2019-2020 Australian bushfires (Abram et al., 2021). A robust estimation of internal variability is also critical for accurately attributing observed climate changes to external forcing by quantifying the contribution of unforced variability (Hegerl and Zwiers, 2011). Understanding internal variability is therefore essential not only for interpreting past climate fluctuations, but also for future climate projections. Even under strong anthropogenic forcing expected over the coming decades, internal variability can substantially influence the trajectory of climate change on decadal to multi-decadal timescales, particularly at regional scales (e.g., Deser et al., 2020; Lehner and Deser, 2023; Fischer et al., 2018). However, the instrumental record remains too short to adequately characterize low-frequency internal variability and its underlying mechanisms.

To extend the observational record of climate variability, information about past climate changes relies on indirect evidence, more specifically paleoclimate records. Over the past centuries, a wide range of such records are available from various natural archives, including tree rings (the most common), ice cores, corals, and sediment cores, among others (Emile-Geay et al., 2017). These archives record environmental conditions at the time of formation through measurable proxy variables (e.g., tree-ring width). Thanks to major efforts from the paleoclimate community, particularly through the Past Global Changes (PAGES) activities, several thousand of paleoclimate records have been compiled and organized into publicly available databases (Emile-Geay et al., 2017; Walter et al., 2023; Thomas et al., 2017; Konecky et al., 2020; Breitenmoser et al., 2014). Given the large number of available records, these databases provide a key opportunity to investigate low-frequency internal climate variability. However, paleoclimate records provide information only at specific locations and are typically sensitive to specific climate variables. Historically, Climate Field Reconstructions (CFRs) have been developed using statistical methods to spatially propagate the local information contained in proxy records (e.g., Mann et al., 2008). More recently, data assimilation (DA) frameworks have been used. DA estimate the state of a dynamical model based on observations and statistical information, and was originally developed for numerical weather prediction (e.g., Kalnay, 2003) and have been adapted for paleoclimate applications (e.g., Goosse et al., 2009; Hakim et al., 2016). Paleo Data Assimilation (PDA) combines the information from proxy records with the physical constraints provided by Earth System Models (ESMs), producing spatially complete, multivariate, and dynamically consistent reconstructions of past climate variability (e.g., Hakim et al., 2016; Valler et al., 2024).

In the standard DA framework, the climate model is integrated forward in time to provide a forecast, which serves as the initial estimate (called the prior) of the climate system at the next time step. However, for paleoclimate applications, integrating a fully coupled climate model over the period of interest (e.g., several centuries) is computationally prohibitive with a state-of-the-art ESM. As a result, and because the forecast step may provide limited additional constraint, most paleoclimate DA studies rely on pre-existing model simulations to construct the prior (e.g., Hakim et al., 2016; Steiger et al., 2018; Valler et al., 2024). In these so-called offline DA frameworks, the prior is sampled from ensembles of simulations. Efforts have been made to improve the realism of the prior by using ensembles of simulations forced with transient external forcing (Valler et al., 2024).

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

However, the reanalysis from Valler et al. (2024) employs a prior derived from atmosphere-only simulations, which therefore do not represent coupled ocean-atmosphere interactions. To address this limitation, Perkins and Hakim (2017) introduced a coupled PDA framework incorporating a Linear Inverse Model (LIM) to approximate the linear impact of ocean inertia on climate variability. Although the LIM provides a computationally efficient representation of ocean memory, it remains a linear approximation trained on General Circulation Model simulations, which limits its ability to fully represent nonlinear ocean processes. More recently, Dalaiden et al. (2025) reconstructed ocean variability using a global ocean–sea-ice model forced by a paleo-based atmospheric reanalysis. While this approach allows the ocean to dynamically respond to atmospheric variability, it does not include two-way coupling between the ocean and atmosphere, and therefore omits coupled feedback processes.

The main objective of this work is to establish the foundation for a fully dynamical, coupled ocean–atmosphere reanalysis spanning the past centuries. To this end, we extend the Norwegian Climate Prediction Model (NorCPM) (Counillon et al., 2016; Bethke et al., 2021) to enable the assimilation of paleoclimate records. NorCPM combines the Norwegian Earth System Model version 1 (NorESM1) (Bentsen et al., 2013; Tjiputra et al., 2013) with an Ensemble Kalman Filter (EnKF) (Evensen, 2003), and was originally developed for climate reconstruction and prediction using instrumental observations. Here, we present a new offline paleo reanalysis (hereafter paleo-NorCPM1 offline) based on a state-of-the-art DA framework, using a 30-member ensemble forced with transient external forcing and an extensive proxy database spanning the past five centuries. This approach leverages a fully coupled Earth System Model prior, allowing consistent representation of atmosphere–ocean interactions. As a step toward a fully coupled paleo reanalysis, we further perform wind-nudged NorESM simulations (hereafter paleo-NorCPM1 wind nudging), in which the reconstructed atmospheric circulation from the offline product (i.e., paleo-NorCPM1 offline) is used to constrain the model through atmospheric nudging, thereby enabling a dynamically consistent reconstruction of the full climate system over the past centuries, including ocean variability. The NorCPM framework, proxy database, and the offline reconstruction methodology are described in Sect. 2, and the wind-nudging framework in Sect. 3. The resulting reanalyses are evaluated against independent observational datasets and compared with existing paleoclimate reanalyses in Sect. 4.

2 Norwegian Climate Prediction Model

The Norwegian Climate Prediction Model (NorCPM) was originally developed for climate reconstruction using instrumental oceanic observations (e.g., Counillon et al., 2016; Wang et al., 2017), as well as for seasonal and decadal climate predictions (e.g., Wang et al., 2019). More specifically, NorCPM combines the Norwegian Earth System Model (NorESM) with an Ensemble Kalman Filter (EnKF; Evensen, 2003) to incorporate observational information into the model state. In this section, we describe the main components of NorCPM, the paleoclimate proxy database used in this study, and the specific adaptations implemented into NorCPM to enable the assimilation of paleoclimate records.

2.1 Ensemble Kalman Filter

The fundamental objective of data assimilation is to update the initial state of a system using available observations in an optimal way, accounting for uncertainties in both the initial state and the observations (e.g., Carrassi et al., 2018). In climate

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

applications, the initial state is represented by an ensemble of climate model simulations, while the observations may come from diverse sources, such as near-surface air temperature from weather stations. The data assimilation method implemented in NorCPM is based on the EnKF (e.g., Kalnay, 2003; Evensen, 2003) and can be written as:

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{K}\left[\mathbf{D} - \mathbf{H}(\mathbf{X}^b)\right], \tag{1}$$

where $\mathbf{X}^b$ is the initial estimate of the system state, called the prior, and $\mathbf{X}^a$ is the updated estimate (the posterior, often referred to as the reanalysis). The matrix $\mathbf{D}$ contains the observations assimilated at a given analysis step.

During the assimilation process, the prior (i.e., $\mathbf{X}^b$) is compared with the observations (i.e., $\mathbf{D}$). For this model-observation comparison, the prior $\mathbf{X}^b$ is mapped into observation space using $\mathbf{H}$, the observational operator. In straightforward cases, $\mathbf{H}$ simply extracts the model value at the appropriate geographic location. In paleoclimate applications, $\mathbf{H}$ corresponds to a Proxy System Model (PSM) (Evans et al., 2013), which simulates the proxy quantity being assimilated (e.g., tree-ring width) from standard climate variables such as temperature. Further details on the PSMs are provided in Section 2.4. The difference between $\mathbf{D}$ and $\mathbf{H}(\mathbf{X}^b)$ is the innovation. Large innovations generally lead to larger updates of the prior, depending on $\mathbf{K}$. In Eq. (1), the analysis is thus a linear combination of the prior and the innovation, weighted by the Kalman gain $\mathbf{K}$, defined as:

$$\mathbf{K} = \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}. \tag{2}$$

The matrix $\mathbf{P}$ is the covariance matrix of $\mathbf{X}^b$. $\mathbf{P}$ plays a pivotal role in the final reanalysis as it governs how information from each assimilated observation is propagated spatially and across variables, through both spatial and multivariate covariances computed from the prior ensemble. The matrix $\mathbf{R}$ represents the observational error covariance matrix. In this study it is specified as diagonal, implying that observational errors are assumed independent between observations. The estimation of $\mathbf{R}$ is described in Section 2.4.

We employ the deterministic Ensemble Kalman Filter (DEnKF) of Sakov and Oke (2008), which updates the ensemble mean and the ensemble anomaly (i.e., anomalies relative to the ensemble mean) separately and avoids the need to perturb the observations (Whitaker and Hamill, 2002; Sakov and Oke, 2008). More specifically, the algorithm first solves Eq. (3) for the ensemble mean as follows:

$$\overline{\mathbf{X}^a} = \overline{\mathbf{X}^b} + \mathbf{K}\left[\mathbf{D} - \mathbf{H}(\overline{\mathbf{X}^b})\right]. \tag{3}$$

Eq. (3) minimizes the distance between the ensemble mean and observations, while accounting for observational error. In a second step, the ensemble anomalies are updated as follows:

$$\mathbf{X'}^a = \mathbf{X'}^b - \frac{1}{2}\mathbf{K}\mathbf{H}(\mathbf{X'}^b), \tag{4}$$

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

where $'$ indicate departures from the ensemble mean (e.g., $\mathbf{X}'^{a} = \mathbf{X}^a - \overline{\mathbf{X}^a}$).

Given the high degree of freedom of the climate system and the sampling error inherent to a finite ensemble size, NorCPM employs two main mitigation strategies. The first addresses the well-known issue of ensemble collapse (i.e., reduction of the ensemble spread). Instead of applying a traditional inflation factor (Anderson, 2001), NorCPM uses a moderation approach: the observational error is increased by a factor of two (Eq. 4). This adjustment affects only the update of the ensemble anomalies, not the estimation of the ensemble mean, thereby preserving sufficient ensemble variance while still providing the optimal estimate of the mean state (Sakov et al., 2012). The second strategy is a pre-screening of observations associated with large innovations. When an observation deviates substantially from the ensemble, the observational error for this specific observation is increased such that the innovation falls within two standard deviations of the ensemble. This preserves the information provided by the observation while preventing it from exerting an impact on the analysis that is inconsistent with the physical constraints from the climate model, thereby maintaining the overall physical consistency of the system.

This is the operational scheme in NorCPM and has been extensively tested and validated (Sakov et al., 2012; Counillon et al., 2016; Bethke et al., 2021; Wang et al., 2025). Further details on the DEnKF are provided in Sakov and Oke (2008), and the specific NorCPM implementation is described in Counillon et al. (2014). For readability, we use the acronym EnKF to refer to the DEnKF throughout the manuscript.

2.1.1 Offline approach

In this study, we use an offline, or no-cycling, data assimilation approach, in which the assimilation increment (i.e., the difference between $\mathbf{X}^b$ and $\mathbf{X}^a$) is not used to update the prior for the following assimilation step. In other words, $\mathbf{X}^b_t$ is not taken as $\mathbf{X}^a_{t-1}$. This approach is widely used in paleo data assimilation (e.g., Hakim et al., 2016; Steiger and Hakim, 2016; Dalaiden et al., 2023; Valler et al., 2024; Hu et al., 2025). The rationale lies from the temporal resolution and representativeness of the assimilated records. Most proxies used in this study are annual and primarily reflect atmospheric and surface processes (see Section 2.3). Because the atmosphere has a weaker temporal memory than one year, the benefit of a fully "online" sequential approach is limited, as demonstrated by previous studies (Matsikaris et al., 2015; Okazaki et al., 2021). Therefore, by employing an offline approach based on a fixed prior, the temporal variability is entirely constrained by the observations incorporated in the data assimilation system.

A major advantage of the offline method is its substantial computational efficiency: the Earth System Model (ESM) does not need to be integrated forward until the next analysis step. Instead, the key requirement for a successful reanalysis is to build a prior that is as informative as possible about the target climate state. Paleo data assimilation studies typically achieve this by drawing priors from long ESM simulations spanning the past several centuries, using seasonal or annual means as the prior ensemble (e.g., Hakim et al., 2016; Steiger and Hakim, 2016; Dalaiden et al., 2023; Valler et al., 2024; Hu et al., 2025). However, during periods with limited proxy availability (i.e., low innovation), this strategy may provide only modest additional constraint, so that the reanalysis remains close to the prior. To partly address this, we construct the prior using a 30-member ensemble of NorESM simulations in which all members share identical external forcing but differ in initial conditions (see Section 2.2 for experimental details). For each reconstructed year, the prior consists of the 30 ensemble members for

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

that specific year, ensuring consistent forcing across the ensemble. Within our framework, in years with zero innovation, the reanalysis is identical to the prior. This provides a more informative estimate than using annual means from a long control simulation, because it retains the externally forced signal specific to that year. This is especially important for periods with strong external perturbations but sparse observational coverage, for example, years following major volcanic eruptions where limited observations cannot constrain the climate response, but the ESM can supply the missing forced component.

2.1.2 Hybridization of the covariance matrix

Using a 30-member ensemble is insufficient to reliably estimate the covariance matrix $\mathbf{P}$, making the reanalysis prone to the sampling error (e.g., Anderson and Collins, 2007). To further mitigate this issue in our study, we adopt a hybrid version of the EnKF in which $\mathbf{P}$ is constructed from a combination of a dynamical covariance matrix ($\mathbf{P}^d$), estimated from the 30-member ensemble, and a static covariance matrix ($\mathbf{P}^s$), derived from a long climatological simulation (e.g., Counillon et al., 2009; Barthélémy et al., 2024; Valler et al., 2024). Therefore, the static covariances are not used to form the prior but only to infer updates resulting from the observations. As such, $\mathbf{P}$ (Eq. 2) is estimated by a linear combination of $\mathbf{P}^d$ and $\mathbf{P}^s$:

$$\mathbf{P} = \beta\,\mathbf{P}^d + (1-\beta)\,\mathbf{P}^s, \tag{5}$$

where $\beta \in [0,1]$. In this study, $\mathbf{P}^s$ is estimated from a 300-year preindustrial control simulation, as in Barthélémy et al. (2024).

Here, the parameter β is a scalar. To determine the optimal value, we performed 20 DA experiments in which β varied from 0 to 1 in increments of 0.05, assimilating all available paleoclimate observations over the period 1950–2000. A value of $\beta = 1$ corresponds to only relying on the dynamical covariance matrix for weighting the innovation, whereas $\beta = 0$ relies exclusively on the static covariance matrix. For each experiment, we evaluated the skill of the resulting reanalysis against observational datasets and selected the value of β that minimized the root mean square error (RMSE) of sea surface temperature (SST), continental near-surface air temperature, and 500-hPa geopotential height. ERSSTv5 (Huang et al., 2017) was used as the reference for SST, while ERA5 (Hersbach et al., 2020) served as the reference for near-surface air temperature and 500-hPa geopotential height. For ocean grid points covered by sea ice (defined as a concentration exceeding 15%), ERA5 near-surface air temperature was used instead of SST.

We obtain an optimal value of $\beta = 0.14$, indicating that the innovation is weighted predominantly by the static covariance matrix rather than the dynamical covariance. This result highlights the importance of the ensemble size. It is important to note, however, that the optimal value of β strongly depends on the specific DA configuration: the dynamical ensemble size, the characteristics of the assimilated observations, and the assimilation frequency. In our case, because most assimilated records are continental proxies with annual resolution (see Section 2.3), a larger contribution from the static covariance matrix is expected, since the long climatological simulation provides a more robust estimate of the covariance structures than can be obtained from the dynamical 30-member ensemble. Moreover, the weak interannual persistence of annually resolved continental proxies limits the added value of flow-dependent covariances.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

2.1.3 Localization cut-off of the covariance matrix

The covariance matrix $\mathbf{P}$ is a pivotal component of DA because it governs how information from assimilated observations is propagated across space and between variables. However, $\mathbf{P}$ is estimated from a finite ensemble (and that even with hybrid covariance), which makes it prone to sampling error and the emergence of spurious long-range covariances. These artificial covariances can introduce substantial biases in the final reanalysis, particularly in regions with sparse observational coverage, where data assimilation increments may be dominated by spurious remote covariances. To mitigate this issue, we apply localization within a local analysis framework, which restricts the update at each model grid point to observations located within a prescribed radius of influence (Evensen, 2003). In this study, we use the Gaspari and Cohn function (Gaspari and Cohn, 1999) to impose a smooth spatial tapering. The Gaspari and Cohn function generates Gaussian weights that decrease to zero at the cut-off radius. As a result, the EnKF restricts the influence of each observation to a defined spatial domain (i.e., a local analysis).

Similar to the hybrid approach described in Section 2.1.2, covariance localization reduces errors arising from the limited number of ensemble members. In an idealized setting with a very large ensemble (e.g., more than 10,000 members), Miyoshi et al. (2014) showed that localization may no longer be necessary in a data assimilation framework using an intermediate Atmospheric General Circulation Model. However, producing such a large ensemble is unpractical using a fully coupled ESM, especially over multi-century timescale considered here (compared with the three-week experiment of Miyoshi et al. (2014)).

The choice of localization radius depends on several factors, including the density and temporal evolution of the observational network, the physical decorrelation scales of the assimilated variables, the assimilation time step, the ensemble size, and the nature of the prior. Paleo DA studies typically employ cut-off distances between 12,000 and 25,000 km (e.g., Tardif et al., 2019; Osman et al., 2021). Using a too small radius can generate discontinuities in the reanalysis field, a particular concern for paleo reanalyses where the observational network is sparse and unevenly distributed in space. Conversely, a too large radius can lead to unrealistic signal. Based on a series of sensitivity tests, we adopt a cut-off distance of 15,000 km for both $\mathbf{P}^d$ and $\mathbf{P}^s$.

2.2 Norwegian Earth System Model version 1

In this study, we use the medium-resolution configuration of the Norwegian Earth System Model, NorESM1-ME (Bentsen et al., 2013; Tjiputra et al., 2013). NorESM1 is a CMIP-class Earth System Model and has participated in CMIP since phase 5. It is built upon the Community Earth System Model version 1 (CESM1) (Hurrell et al., 2013), but includes several key modifications. The atmospheric component of NorESM1 is based on the Community Atmosphere Model version 4 (CAM4) (Neale et al., 2010), with adjustments including a fully prognostic treatment of the aerosol life cycle and a representation of aerosol-cloud interactions (Kirkevåg et al., 2013). The ocean component is an updated version of the Miami Isopycnic Coordinate Ocean Model, namely the Bergen Layered Ocean Model (BLOM) (Bentsen et al., 2013). Other model components follow those used in CESM1: the Los Alamos Sea Ice Model version 4 (CICE4) (Holland et al., 2012) for sea ice, and the

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Community Land Model version 4 (CLM4) (Lawrence et al., 2011) for land processes. A detailed description of NorESM1 is provided in Bentsen et al. (2013).

NorESM1-ME is run at approximately 2° horizontal resolution (latitude: 1.9°; longitude: 2.5°) for both the atmospheric and land components. CAM4 uses 26 vertical hybrid-sigma levels, with the model top at 3 hPa. The ocean and sea-ice components employ a nominal 1° horizontal resolution. BLOM uses 51 vertical isopycnal layers in combination with a two-layer bulk mixed layer formulation with varying mixed layer depth and density.

2.2.1 Experimental design of NorESM1 simulations

For this study, we first generate a 30-member ensemble for the period 1500–1860 based on transient external forcing (see Section 2.2.2 for details). After 1850, we make use of the already existing historical 30-member ensemble spanning the 1850–2100 period from Bethke et al. (2021), which uses the same model version. The 10-year overlap between the two ensembles allows us to match any potential inconsistencies by aligning their ensemble means. As for the initial conditions in 1500, we initialize each member from different years of a 1153-year NorESM1-ME preindustrial control simulation. Specifically, initial conditions are taken from the final 150 years of the simulation at five-year intervals, providing 30 distinct and dynamically consistent initial conditions. This procedure perturbs not only the atmosphere but also the full ocean, thereby ensuring sufficient ensemble spread in the fully coupled climate system.

2.2.2 External forcing datasets

The external forcing over the period 1500–2010 combines datasets following both the PMIP and CMIP protocols. For the pre-1850 period, volcanic forcing is taken from Gao et al. (2008), who reconstructed volcanic aerosol loadings using 54 ice-core records from Greenland and Antarctica. From 1850 onward, we use the standard CMIP6 volcanic forcing dataset (Revell et al., 2017; Thomason et al., 2018). Solar forcing is similarly merged from two sources. Before 1850, total solar irradiance variations are taken from the Spectral And Total Irradiance REconstruction (SATIRE) dataset (Vieira et al., 2011), including an imposed 11-year sunspot cycle. For the historical period, solar forcing is from the CMIP6 forcing dataset (Matthes et al., 2017). Greenhouse gas concentrations follow the PMIP3 (Schmidt et al., 2012) and CMIP6 protocols (Meinshausen et al., 2017). For all the forcing datasets, we note that the pre-historical and historical datasets show negligible differences over their overlap period (1850–2000). Orbital forcing is prescribed according to the calculations of Berger (1978). For the historical period, aerosol emissions are also prescribed using the CMIP6 forcing (Meinshausen et al., 2017; Hegglin et al., 2016). From 2015, the Shared Socioeconomic Pathway 2.5-4 (SSP2-4.5) is used as forcing.

Over the past 500 years, volcanic activity has exhibited strong variability (Figure S1), with several major eruptions prior to the 20th century, notably the 1815 Tambora eruption, and more recently the 1991 Mount Pinatubo eruption. Solar forcing has also displayed substantial changes (Figure S2), including a prolonged period of reduced solar activity from the mid-17th to early-18th century (known as the Maunder Minimum), followed by a general increase toward present-day levels. Greenhouse gas concentrations remained relatively stable until the 19th century (Figure S3), after which they began to markedly rise, most prominently during the 20th century.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

2.3 Paleoclimate observations

Since our objective is to reconstruct climate variability over the past several centuries, we restrict our observational database to paleoclimate records. This choice ensures long temporal coverage and avoids potential artifacts in the reconstruction that may arise from mixing instrumental and proxy observations, including for instance a jump in the data availability. We consider only sub-annual and annual records. Sub-annual records (some coral records) are averaged to annual resolution to ensure consistency across all records.

Most of the records used here originate from major international community efforts, particularly the Past Global Changes (PAGES) working groups, which have compiled climate proxies spanning the last millennium. The latest PAGES2k database (Emile-Geay et al., 2017) provides the largest collection of multi-archive records covering the past 2,000 years validated by experts. Although this database includes numerous proxy types from a wide range of archives, we limit our selection to tree rings (width [TRW] and mixed latewood density [MXD]), ice cores (δ^{18}O, ^{2}H, snow accumulation and sea salt content), and corals (δ^{18}O and Sr/Ca), owing to their well-established physical and statistical relationships with climate. These three archives represent approximately 98% of available records in the latest PAGES2k database. However, because the PAGES2k database was released in 2017, many recently published records are not included. We therefore supplement this database with additional databases.

For coral records, we incorporate the CoralHydro2k database (Walter et al., 2023), which provides a compilation of coral records spanning the last two millennia (δ^{18}O and the ratio of Strontium and Calcium [Sr/Ca] from the coral skeleton). For ice-core records, we leverage from several additional sources to address gaps in PAGES2k. First, we use the Iso2k database (Konecky et al., 2020), from which we retain only ice-core isotope records (δ^{18}O and ^{2}H). We further include Antarctic and Greenland snow-accumulation records (Thomas et al., 2017; Medley et al., 2018; Box et al., 2013) as well as the recently published Clivash2k database of sea-salt concentrations in Antarctic ice cores (Thomas et al., 2022). Finally, to expand the tree-ring width archive, we also add the large compilation of tree-ring width time-series from Breitenmoser et al. (2014), who synthesized all the chronologies from the International Tree Ring Data Bank (ITRDB). Detailed information for all record types are provided in Table 1.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Table 1. Proxy observations used in paleo-NorCPM1.

Data	Archives	Proxy variables	References
$\delta^{18}O$ coral records	coral	$\delta^{18}O$	CoralHydro2k (Walter et al., 2023)
Sr/Ca coral records	coral	Sr/Ca	CoralHydro2k (Walter et al., 2023)
PAGES2k tree-ring width tree records	tree	TRW	PAGES2k (Emile-Geay et al., 2017)
PAGES2k mixed latewood density tree records	tree	MXD	PAGES2k (Emile-Geay et al., 2017)
Breitenmoser tree-ring width tree records	tree	TRW	Breitenmoser et al. (2014)
$\delta^{18}O$ ice-core records	ice core	$\delta^{18}O$	Iso2k (Konecky et al., 2020)
^{2}H ice-core records	ice core	^{2}H	Iso2k (Konecky et al., 2020)
Antarctic snow accumulation ice-core records	ice core	snow accumulation	Ant2k (Thomas et al., 2017; Medley et al., 2018)
Greenland snow accumulation ice-core records	ice core	snow accumulation	Box et al. (2013)
Antarctic sea salt ice-core records	ice core	sea salt (Na^{+})	Clivash2k (Thomas et al., 2022)

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

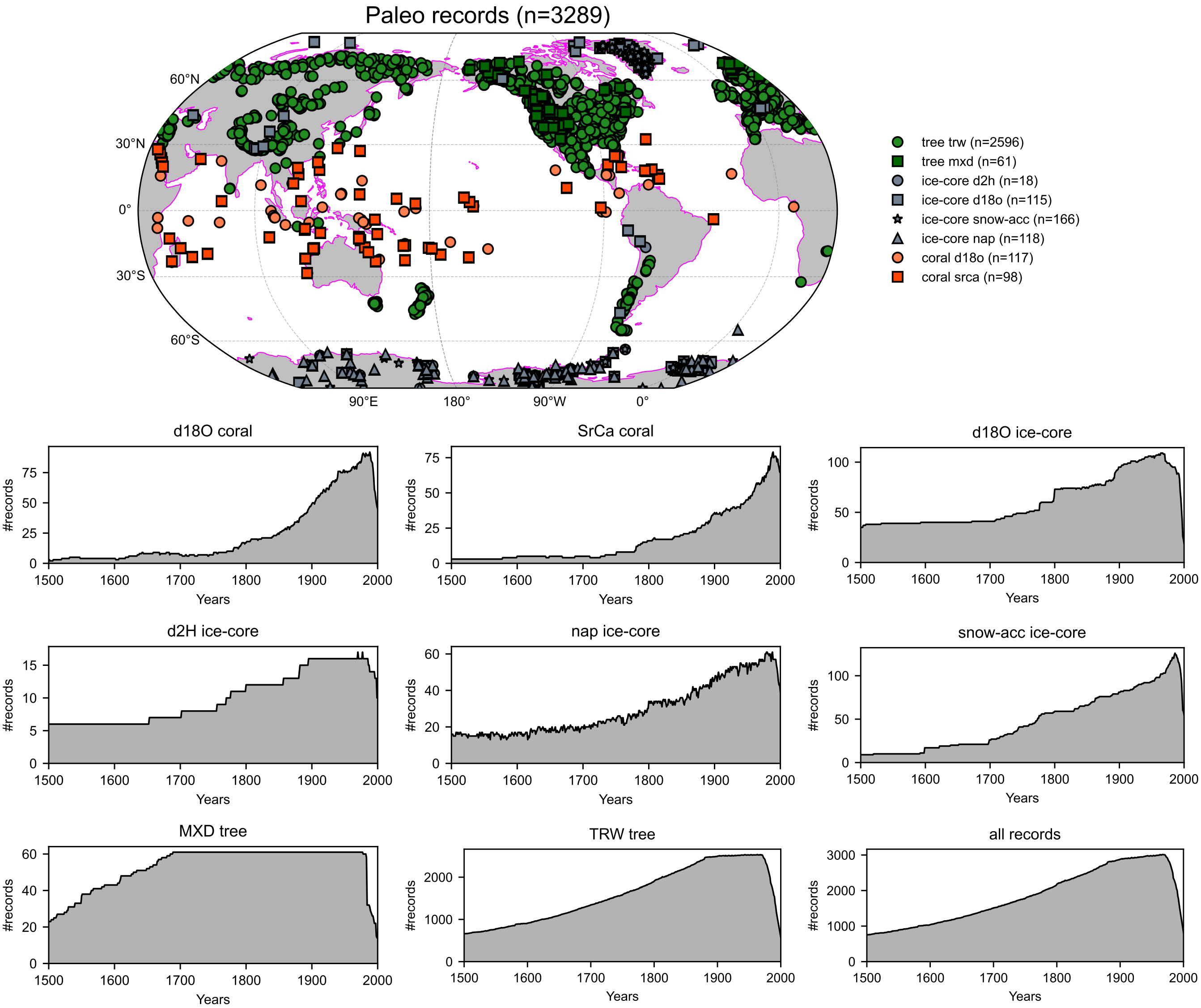

Figure 1. Geographical locations of all paleo records along with the temporal availability for the different types of records.

In total, our observational database comprises more than 3,000 records (Figure 1). The number of available records increases gradually through time, with the highest coverage occurring in the late 20th century. The observational network is dominated by continental proxies, primarily tree-ring width records, reflecting the substantial contribution from Breitenmoser et al. (2014). In contrast, oceanic information is limited to coral records, which represent the only proxy type providing direct historical information for the ocean. Spatial coverage is uneven. Several regions, in particular the African continent and the extratropical oceans, remain sparsely sampled. Furthermore, despite the large number of Antarctic ice-core records, the Southern Hemisphere as a whole is under-represented.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

2.4 Proxy System Models

The Proxy System Model (PSM) serves as the forward operator ($\mathbf{H}$ in Eq. (2)) that maps climate variables from the ESM outputs (i.e., the prior) onto the observational proxy space. Following the definition of Evans et al. (2013), a PSM represents "*the complete set of processes by which the response of a sensor (e.g., a tree) to environmental forcing is recorded and observed in a material archive (e.g., a tree ring) as a measurable variable (e.g., tree-ring width).*" Historically, paleoclimate studies relied on inverse modeling (e.g., Mann et al., 2008), translating proxy observations into climate variables. With the emerging development of process-based PSMs (e.g., Evans et al., 2013; Dee et al., 2015), which explicitly simulate the physical, biological and chemical processes governing proxy recording, forward modeling has become the standard, as it provides a natural and flexible framework for integrating diverse proxy types into DA systems.

In theory, process-based PSMs are ideal for mapping the prior onto proxy space because, by explicitly resolving the processes leading to the proxy archiving, these PSMs explicitly encode the associated environmental conditions to the proxy archiving. However, these PSMs often involve a high degree of freedom, making calibration challenging (e.g., Rezsöhazy et al., 2022). Their complexity requires extensive observational constraints, which are rarely available for most proxy types and locations. This strongly limits global applications. For these reasons, and consistent with many paleoclimate DA studies (e.g., Hakim et al., 2016; Steiger et al., 2017; Valler et al., 2024; Dalaiden et al., 2025), we adopt simple statistical PSMs in paleo-NorCPM1, specifically linear uni- or bivariate models. In these models, the proxy value is simulated using one or two explanatory climate variables. Besides being easier to calibrate over process-based PSMs, these simple PSMs offer considerable practical advantages, including straightforward implementation across all the paleoclimate archives and types. Nevertheless, a limitation of this approach is that simple statistical PSMs do not explicitly represent all processes involved in proxy recording. Consequently, by using such simple PSMs, we do not fully exploit the information from proxy observations.

For each assimilated proxy time-series, the following linear uni- or bivarite models are considered:

$$y_i = \alpha_i + \beta_{1i} X_{1i} + \epsilon_i \tag{6}$$

$$y_i = \alpha_i + \beta_{1i} X_{1i} + \beta_{2i} X_{2i} + \epsilon_i \tag{7}$$

where i corresponds the ith proxy time-series and y_i denotes the simulated time series for proxy i. The term α_i is the intercept, while β_{1i} and β_{2i} are the linear coefficients (i.e., slopes) associated with the explanatory variables X_{1i} and X_{2i}, respectively. The term ϵ_i represents Gaussian noise with zero mean and unit variance. The parameters α_i, β_{1i} and β_{2i} are estimated using ordinary least squares.

The definitions of the PSMs, together with their explanatory variables and calibration datasets for each proxy type, are summarized in Table 2.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Table 2. Information regarding the Proxy System Models.

Paleo records	PSMs	Calibration datasets	Calibration period	Min. length
Ice core				
δ^{18}O, ^{2}H (AIS)	f(tas)	Bromwich et al. (2024)	1958–	21
δ^{18}O, ^{2}H (GrIS)	f(tas)	ERA5 (Hersbach et al., 2020)	1950–	21
Snow accumulation (AIS)	f(prcp)	ERA5 (Hersbach et al., 2020)	1979–	9
Snow accumulation (GrIS)	f(prcp)	ERA5 (Hersbach et al., 2020)	1958–	9
Sea salt	f(slp)	Dalaiden et al. (2025)	1958–	21
Tree				
TRW, MXD	$f(\mathrm{tas_s},\mathrm{prcp_s})$	ERA5 (Hersbach et al., 2020)	1950–	21
Coral				
δ^{18}O	f(sst,sss)	NorCPM1 (Bethke et al., 2021)	1950–	21
Sr/Ca	f(sst)	NorCPM1 (Bethke et al., 2021)	1950–	21

sst: sea surface temperature, sss: sea surface salinity, tas: near-surface air temperature, prcp: precipitation, slp: sea-level pressure, $_s$ denotes seasonal mean (detailed in Section 2.4.2).

The PSMs are calibrated over the second half of the 20$^{\mathrm{th}}$ century, when instrumental observations are most reliable and spatially complete. Consequently, restricting the calibration to this period minimizes biases associated with data sparsity and inhomogeneities and ensures that the statistical relationships between proxies and climate variables are derived from high-quality climate datasets.

2.4.1 Ice-core records

For ice-core water isotope records (i.e., δ^{18}O and ^{2}H) records, we assume that the isotopic signal reflects annual-mean local near-surface air temperature variability, following the well-established isotope–temperature relationship (e.g., Jouzel et al., 1997). For records located on the Greenland Ice Sheet, ERA5 is used for calibration, as it provides a realistic representation of surface processes from 1950 onward when compared with state-of-the-art polar regional atmospheric models (Delhasse et al., 2020). For Antarctic ice cores, we instead employ the recently published dataset of Bromwich et al. (2024), because ERA5 exhibits a cold bias over Antarctica prior to 1979. The Bromwich et al. (2024) dataset is based on a kriging approach that interpolates surface temperature across the ice sheet using long-term weather-station observations available from 1958 (Turner et al., 2004).

Ice-core snow-accumulation records are calibrated against local precipitation. ERA5 precipitation is used for Greenland from 1958 onward, while for Antarctica the calibration period is restricted to 1979 onward due to the pre-1979 cold bias in ERA5 Bromwich et al. (2024). Although precipitation in ERA5 is not directly assimilated but generated by the model physics (Hersbach et al., 2020), it is constrained indirectly by the assimilation of atmospheric temperature, humidity, and circulation-related fields. It therefore provides one of the most spatially consistent large-scale precipitation products available over the ice sheets, where in situ observations are sparse (Gossart et al., 2019; Delhasse et al., 2020). Because snow-accumulation records are directly compared with precipitation, the minimum length required for calibration is relaxed to nine data points, following Medley and Thomas (2018). For both water isotope and snow-accumulation records, proxies are excluded when the sign of the relationship with the explanatory variable is inconsistent with physical expectations. This screening strategy has also been adopted in previous paleoclimate reconstructions (e.g., Emile-Geay et al., 2017; Hakim et al., 2016; Valler et al., 2024). A specific PSM is calibrated for each ice-core records, depending on the local climatic inputs.

Finally, for sea-salt records, we employ the PSM based on the relationship with the large-scale atmospheric circulation developed by Dalaiden et al. (2023), using the surrounding, offshore sea-level pressure as the explanatory variable. Calibration relies on the recent atmospheric reanalysis of Dalaiden et al. (2025), which is constrained by long-term surface temperature and atmospheric pressure observations available from 1958 (Turner et al., 2004). This approach extends the calibration period by 21 years compared with what is possible using ERA5 alone as ERA5 does not provide accurate variability in the high-latitude Southern Hemisphere regions before 1979 (Bromwich et al., 2024; Dalaiden and Bethke, 2026).

2.4.2 Tree records

For tree-ring records, the PSMs are based on local near-surface temperature and precipitation. We tested multiple combinations of explanatory variables for both tree-ring width and wood-density records, exploring a range of seasonal averages to identify the most appropriate PSM for each tree chronology. Tree-ring records may preferentially reflect temperature, precipitation, or a combination of both, and their sensitivity depends on specific seasons, such as the growing season or summer conditions. For that reason, temperature and precipitation were averaged over several seasonal windows and tested within uni- and bivariate linear models. For Northern Hemisphere records, the tested seasons include AMJJAS (April to September), JJA (June to August), JAS (July to September), AMJJA (April to August), and JJASON (June to November), while for Southern Hemisphere records the tested seasons are ONDJFM (October to March), DJF (December to February), JFM (January to March), SONDJF (September to February), and DJFMAM (December to May). All possible combinations of explanatory variables and seasonal windows were evaluated. The optimal PSM for each record was selected using the Bayesian Information Criterion (BIC), which penalizes model complexity (Schwarz, 1978); the PSM with the lowest BIC was retained. Therefore, like for ice-core records, for each tree-ring records, a specific PSM is built.

ERA5 was used as the calibration dataset to ensure dynamical consistency between temperature and precipitation fields. Sensitivity tests using the Global Meteorological Forcing Dataset for land surface modeling (Sheffield et al., 2006) yielded broadly similar but slightly inferior PSM performance. Additionally, some previous studies have instead relied on drought indices to represent moisture availability in tree-based PSMs (e.g., Steiger et al., 2018). Accordingly, we tested replacing precipitation with the Standardized Precipitation–Evapotranspiration Index (SPEI). However, PSMs based on SPEI exhibited no substantial improvement regarding the skill of the reanalysis. Therefore, precipitation was retained as the indicator of moisture availability in the final PSM configuration.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

2.4.3 Coral records

The coral proxies used in this study are δ^{18}O and Sr/Ca. Coral δ^{18}O reflects variations in both sea surface temperature (SST) and sea surface salinity (SSS) (Epstein et al., 1953; Thompson et al., 2011; Conroy et al., 2017), whereas Sr/Ca primarily records SST variability (Beck et al., 1992). Accordingly, both SST and SSS are used as explanatory variables in the PSMs for coral δ^{18}O records, while only SST is employed for Sr/Ca records.

Like the other proxy types, records are excluded from the assimilation system when the sign of the relationship between the proxy and the explanatory variables is inconsistent with the theory (i.e., a negative slope with SST and a positive slope with SSS for δ^{18}O). Coral PSMs are calibrated using the oceanic reanalysis NorCPM1 (Bethke et al., 2021), which is based on the Norwegian Climate Prediction Model and assimilates gridded SST from HadSST2 (Rayner et al., 2003) along with subsurface temperature profiles (Good et al., 2013). Because NorCPM1 is built upon NorESM1, this choice ensures dynamical consistency with the paleo reanalysis framework used in this study.

The coral PSMs (specific to each records) are built using annual means defined over the tropical year (April-March). Although some records from CoralHydro2k have sub-annual resolution, all coral records are averaged to annual resolution to ensure consistency across the full proxy database.

2.5 Observational error

The observational error ($\mathbf{R}$ in Eq. 2) plays a central role in the reanalysis, as it determines the extent to which each proxy record influences the update of the prior. In paleo data assimilation, the observational error is commonly assumed to arise primarily from two sources (Badgeley et al., 2020; Dalaiden et al., 2022): (i) unresolved processes in the climate model due to the coarse spatial resolution, and missing physical processes in the climate model; and (ii) errors associated with the PSMs, which do not explicitly represent all processes involved in proxy archiving. Following previous DA-based reconstructions (Tardif et al., 2019; Valler et al., 2024), we define the observational error as the variance of the residuals from the linear PSMs. Because all climate variables are bi-linearly interpolated onto the CAM grid prior to PSM calibration, representativeness errors are implicitly accounted for.

To improve the robustness of the observational error estimate, we apply a Monte Carlo approach during PSM calibration. For each proxy record, 1,000 realizations are generated by randomly selecting 80% of the available data for calibration and using the remaining 20% to estimate the residual error. The mean error variance across the 1,000 realizations is retained as the final observational error. Compared to estimates based on the full calibration period alone, this procedure increases the standard deviation of the residuals by approximately 20%, reflecting the additional uncertainty associated with limited calibration data. This approach, using a split between calibration and validation periods for error estimation while retaining the full dataset to calibrate the regression coefficients, provides a balanced compromise given the relatively short overlap between proxy records and the available calibration datasets.

2.6 Data assimilation experimental design

The offline paleo reanalysis spans the period 1501–2000, at monthly resolution. The assimilation window corresponds to the April–March tropical year, in order to preserve Pacific tropical variability, which represents the major driver of global internal climate variability (e.g., Allan, 2000; England et al., 2014). Because our objective is to use paleoclimate observations to constrain the temporal variability of the ESM, the assimilation is performed on anomaly fields. Both the prior and proxy records are expressed as anomalies relative to the 1961–2000 mean, thereby removing the mean state differences between the model and the observations.

Only proxy records for which a local PSM has been successfully calibrated are assimilated; specifically following the criteria of the regression slope(s) being statistically significant at the 95% confidence level. This pre-screening procedure has been shown to improve reconstruction skill compared with assimilating all available records (Franke et al., 2020), as it retains only proxies that exhibit a robust relationship with climate variables and thereby maximizes the signal-to-noise ratio, i.e., the strength of the climatic signal relative to non-climatic variability in the proxy records. Proxy records exhibiting values exceeding 3.5 standard deviations from their temporal mean are excluded from the assimilation at the corresponding time step (pre-screening). This procedure aims to avoid the assimilation of extreme values that are inconsistent with the model variability or may reflect proxy-specific noise or errors. Such large innovations can lead to unrealistic corrections and degrade the stability of the DA system. Similar approaches are employed in previous studies (e.g., Dalaiden et al., 2023; Valler et al., 2024).

Ice-core and coral records are available as annual signals, whereas tree-ring records primarily reflect seasonal conditions, typically associated with the growing season or summer mean. As a result, reconstruction skill is highest for annual means and during the summer seasons. To propagate this seasonal information to monthly variability, we exploit the temporal covariance structure. Therefore, the state vector (i.e., $\mathbf{X}^{\mathbf{b}}$ in Eq. 1) includes all variables required by the PSMs as well as the reconstructed climate fields, at monthly resolution. The innovation $(\mathbf{D} - \mathbf{H}(\mathbf{X}^b))$ is at the resolution of the proxy record (annually or seasonally). In the case of a tree-ring record whose selected PSM depends on local summer temperature, the innovation is computed from the mismatch between the observed proxy value and the simulated proxy value derived from summer temperature. If the prior covariance between summer temperature at that site and other variables or locations is zero, then no update is applied to those variables or locations, even though the innovation itself is non-zero. The resulting update is therefore governed by the modeled covariance structure. In theory, all the variables simulated by the ESM can be reconstructed. However, throughout the manuscript, we mainly analyse surface variables for which we expect the largest impact, such as near-surface air temperature, SST, precipitation, as well as variables related to the atmospheric circulation, like 500-hPa geopotential heights and low to mid-tropospheric winds.

3 Wind nudged NorESM simulations

In offline DA frameworks, the prior at time $t+1$ does not depend on the posterior at time t. In other words, each time step is reconstructed independently. Because the DA window in paleo-NorCPM1 offline is annual, this approach is not well suited for reconstructing subsurface ocean variability, owing to the longer intrinsic memory of the ocean compared with the atmosphere.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

To address this limitation, we perform additional NorESM simulations using atmospheric nudging that dynamically propagate benefits from atmospheric assimilation into the ocean.

The primary objective of these nudging experiments is to reconstruct climate variability consistent with the reconstructed atmospheric circulation from paleo-NorCPM1 offline, while allowing the ocean and sea-ice components to respond dynamically. This approach provides a physically consistent reconstruction of the coupled climate system, under the assumption that wind variability and external forcing are the dominant drivers of large-scale climate variability. Furthermore, constraining only the atmospheric circulation provides a framework for future fully coupled online data assimilation, in which oceanic proxy records could be assimilated directly, as demonstrated in Garcia-Oliva et al. (2024).

Atmospheric nudging in ESMs allows reconstruction of the full climate state while constraining specific atmospheric variables. As described in Roach et al. (2022), the implementation of atmospheric nudging in CAM modifies the temporal evolution of the atmospheric model state $\mathbf{x}$. Without nudging, the model evolves freely according to its internal dynamics, and the temporal tendency of the system is given by:

$$F(x) = \frac{d\mathbf{x}}{dt} \tag{8}$$

where t stands for time. When atmospheric nudging is applied, a restoring term is added in the model tendency for the next analysis time step that relaxes the model state toward prescribed target conditions $\mathbf{T}$ over a specified restoring timescale τ. The nudging tendency is defined as:

$$F_{nudge} = \frac{\mathbf{T} - \mathbf{x}(t)}{\tau} \tag{9}$$

Therefore, the temporal evolution of the model state can be rewritten as follows:

$$\frac{d\mathbf{x}}{dt} = F(\mathbf{x}) + \frac{\mathbf{T} - \mathbf{x}(t)}{\tau} \tag{10}$$

In this formulation, the model state $\mathbf{x}$ is continuously relaxed toward the prescribed atmospheric target $\mathbf{T}$, while remaining dynamically consistent with the coupled model physics.

In our experiments, the three-dimensional zonal and meridional wind fields from paleo-NorCPM1 offline are globally nudged over the period 1600–2000. More specifically, we perform ten simulations using the wind fields from ten ensemble members of paleo-NorCPM1 offline. We consider that this ensemble size represents a reasonable trade-off between computational cost and a sufficient number of members to capture the ensemble spread of paleo-NorCPM1 offline. The experimental setup is otherwise identical to the transient simulations (see Section 2.2.1). These simulations therefore reflect the climate evolution constrained by the reconstructed atmospheric variability from paleo-NorCPM1 offline and external forcing. In this configuration, while the atmospheric circulation is constrained, the thermodynamics predominantly evolve freely through coupled ocean-atmosphere interactions. To preserve air-sea coupling and allow surface fluxes to adjust dynamically, the atmospheric boundary layer is excluded from nudging (i.e., the lowest five atmospheric layers are not nudged). The wind fields are relaxed toward the target state at every model time step (30 minutes), using a restoring timescale τ of 6 hours.

Paleo-NorCPM1 offline provides wind fields only at monthly resolution. As an initial step, we therefore set the restoring timescale τ to one month. However, this configuration yielded weak synchronization between the model and the target wind fields, including an increasing time lag. This is partly related to differences between the intrinsic model tendency and the target fields, which become more pronounced when using longer restoring timescales, as the nudging correction is applied more gradually compared with a short restoring timescale. Similar results were obtained when using restoring timescales of 10 days and 1 week, consistent with previous findings (e.g., Kruse et al., 2022). Effective nudging requires a frequent restoring towards the target fields. Applying strong restoring directly toward monthly-mean wind fields without including sub-monthly variability will reduce the atmospheric variability, leading to an unrealistic climate variability. Therefore, following earlier work (Naughten et al., 2018; Dalaiden et al., 2025), we generated 6-hourly wind fields by superimposing generic high-frequency variability onto the monthly variability of paleo-NorCPM1 offline. The sub-monthly variability was derived from a pre-industrial control simulation performed with NorESM1-ME. Specifically, the monthly mean was removed from the original 6-hourly wind fields to isolate the high-frequency component. A 30-year segment was selected to construct this high-frequency variability, which was then cyclically reused throughout the experiment. To reduce sensitivity to the specific realization of the prescribed high-frequency variability, each ensemble member is initialized from a different starting point within the 30-year high-frequency variability segment (e.g., year 0 for member 1 and year 9 for member 10), after which the sequence is applied cyclically throughout the simulation. This approach minimizes the imprint of the selected high-frequency variability on the simulations and ensures that the reconstructed climate variability remains primarily constrained by the monthly wind variability from paleo-NorCPM1 offline, rather than by the particular realization of sub-monthly variability.

Before applying this workflow to the real experiments with the target wind fields from paleo-NorCPM1 offline, we evaluated the nudging configuration using a twin experiment setup (i.e., based on free NorESM1 simulations). These tests demonstrate that the nudging setup successfully synchronizes the monthly wind variability and enables realistic reconstruction of major modes of climate variability, including among others El Niño-Southern Oscillation (ENSO), North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO), Southern Annular Mode (SAM), Indian Ocean Dipole (IOD) and the Interdecadal Pacific Oscillation (IPO; Figure S5).

The initial wind nudging experiments exhibit a small cooling drift in the ocean. Such drift has also been reported in wind nudging experiments with CESM (e.g., Gilbert et al., 2025). This drift primarily arises from small imbalances in air-sea heat fluxes introduced when the atmospheric circulation is externally imposed. Although the drift is predominantly confined to deep ocean layers, its cumulative effect over the 400-year simulation can influence the reconstructed variability. Therefore, to mitigate this long-term drift, we apply an ad hoc surface heat flux correction. Specifically, a spatially uniform, constant heat flux is added to the ocean at each coupling time step between the atmosphere and ocean. The magnitude of the correction is estimated using a twin experiment framework by evaluating the drift in a 100-year control simulation with monthly wind nudging, following the nudging setup described above. This turns into a required heat flux correction of 0.22 W m^{-2}. Although the heat flux correction compensates for most of the oceanic cooling drift, a small residual drift remains due to nonlinear

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

processes. Nevertheless, the magnitude of the remaining drift is modest and comparable to the intrinsic drift present in long free-running model simulations.

4 Evaluation

In order to evaluate the performance of paleo-NorCPM1, we rely on three complementary metrics, following Dalaiden et al. (2022). The first metric is the Pearson correlation coefficient (r), which measures the strength of the linear relationship between two time series and quantifies the ability of the reconstruction to reproduce the temporal or spatial change. It is defined as:

$$r = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - \overline{y})(x_i - \overline{x})}{\sigma_y \sigma_x} \tag{11}$$

where N is the length of the time series, x denotes the reference time series (i.e., the observational time-series), and y represents the reconstructed time series (i.e., paleo-NorCPM1). σ_x and σ_y correspond to the standard deviations of the reference and reconstructed time series, respectively, and overbars indicate temporal means. In addition to r, we also use the Coefficient of Efficiency (CE) (Nash and Sutcliffe, 1970), , also known as the mean square skill score (MSSS), which evaluates both the timing and amplitude of variability relative to a reference climatology. CE is defined as:

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{12}$$

CE ranges from $-\infty$ to 1. A value of 1 indicates perfect agreement between reconstruction and reference, while values greater than 0 indicate that the reconstruction performs better than the reference climatology. Finally, we assess the reliability of the ensemble spread using the Ensemble Calibration Ratio (ECR), defined as:

$$\mathrm{ECR} = \mathrm{median}\left(\frac{(x_i - y_i)^2}{\sigma^2_{ens,i}}\right) \tag{13}$$

The ECR quantifies the consistency between ensemble spread and reconstruction error. An ECR value greater than unity indicates an underdispersive ensemble, meaning that the ensemble spread underestimates the reconstruction error. In contrast, an ECR value less than unity indicates an overdispersive ensemble, where the ensemble spread exceeds the reconstruction error. Ideally, an ECR value close to 1 indicates a well-calibrated ensemble. From a conservative perspective, an underdispersive ensemble is less desirable than an overdispersive ensemble, as it underestimates the true reconstruction uncertainty and may lead to overconfident interpretation of the results.

4.1 Evaluation over the instrumental period

To assess the performance of paleo-NorCPM1 in reproducing year-to-year variability, we compare key climate variables against instrumental and reanalysis datasets over the 20$^{\text{th}}$ century. Specifically, we evaluate the following variables and their corre-

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

sponding reference datasets: 500-hPa geopotential height from the 20$^{\mathrm{th}}$ Century Reanalysis version 3 (20CRv3) (Slivinski et al., 2021), near-surface air temperature from the Met Office Hadley Centre/Climatic Research Unit global surface temperature dataset (HadCRUT5) (Morice et al., 2021), SST from NOAA Extended Reconstructed SST version 5 (ERSSTv5) (Huang et al., 2017), precipitation from the Global Soil Wetness Project phase 3 dataset (GSWP3) (Lange et al., 2021), SPEI derived from near-surface temperature from HadCRUT5 and precipitation from GSWP3-W5E5, and sea-ice concentration from HadISST4 (Kennedy et al., 2019). With the exception of sea-ice concentration, which is evaluated over the period 1979–2000 owing to data availability, all variables are assessed over the full 20$^{\mathrm{th}}$ century. The evaluation relies on datasets independent of those used during the PSM calibration, thereby ensuring an objective assessment of the reanalysis performance, although common observations may be shared between the datasets for the PSM calibration and reanalysis evaluation.

Figure 2 shows maps of the temporal correlation coefficient, computed from annual averages over the calendar year, for the six evaluated variables. Overall, paleo-NorCPM1 offline reproduces large-scale atmospheric circulation, near-surface air temperature, and SST well, with spatially averaged correlation coefficients of 0.54, 0.55, and 0.52, respectively. The highest correlations are found in the tropics, particularly over the oceans, reflecting an accurate representation of tropical variability through the information provided by the coral records. Regions characterized by dense observational networks also exhibit higher skill, including North America, Europe, the Greenland Ice Sheet, and the West Antarctic Ice Sheet. For precipitation, paleo-NorCPM1 offline performs best in the tropical Pacific, indicating a good representation of Inter Tropical Convergence Zone (ITCZ) dynamics, as well as in North America and Europe, where tree-ring and ice-core snow-accumulation records provide strong constraints. However, skill is lower in data-sparse regions, particularly over parts of Africa, reflecting the limited availability of proxy constraints. It is also important to note that instrument-based gridded precipitation datasets contain substantial uncertainties (e.g., Sun et al., 2018; Lange et al., 2021), which complicates the evaluation of reconstructed precipitation. Consistent with this, the agreement is much closer when comparing paleo-NorCPM1 offline with ERA5 (Figure S7). Finally, the reconstructed sea-ice concentration shows substantial skill, particularly in the Baffin and Hudson Bays, the Beaufort and Labrador Seas in the Arctic, and the Amundsen–Bellingshausen and Weddell Seas in the Southern Ocean. As no direct sea-ice proxies are assimilated in paleo-NorCPM1 offline, this highlights the role of the ensemble covariance to update variables other than those assimilated. Seasonally, reconstructed near-surface air temperature over North America and Europe shows stronger agreement with HadCRUT5 in summer than in annual means, highlighting the seasonal information provided by tree-ring records.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

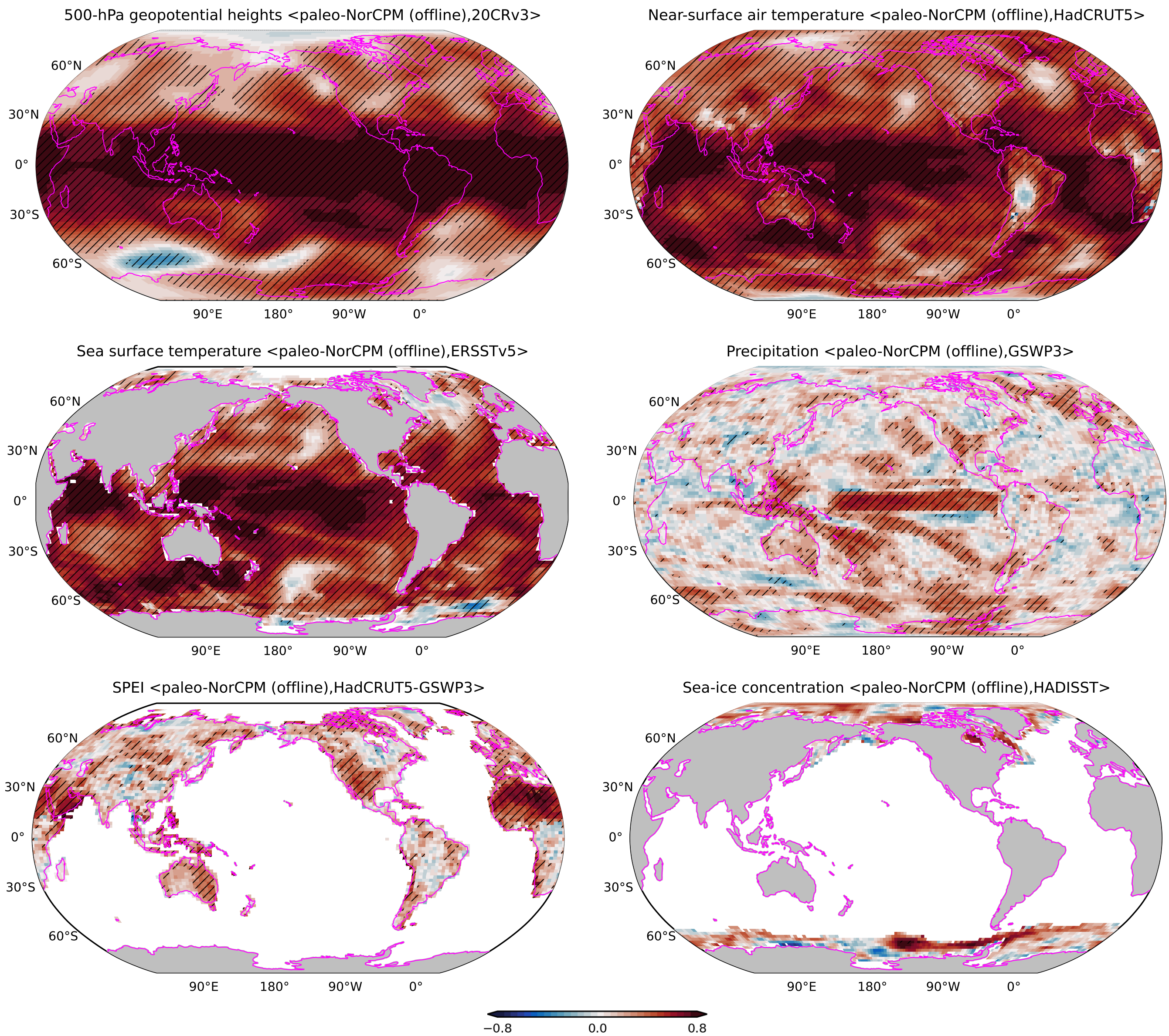

Figure 2. Maps of the correlation coefficients for paleo-NorCPM1 offline (ensemble mean), evaluated against reference datasets for six variables: 500-hPa geopotential height from the 20th Century Reanalysis version 3 (20CRv3; Slivinski et al., 2021); near-surface air temperature from the Met Office Hadley Centre/Climatic Research Unit global surface temperature dataset (HadCRUT5; Morice et al., 2021); SST from the NOAA Extended Reconstructed SST version 5 (ERSSTv5; Huang et al., 2017); precipitation from the Global Soil Wetness Project phase 3 dataset (GSWP3) (Lange et al., 2021); SPEI derived from HadCRUT5 temperature and GSWP3 precipitation; and sea-ice concentration from HadISST4 (Kennedy et al., 2019). Correlations are computed using annual time-series and over the 20th century, except for sea-ice concentration, which is evaluated over 1979–2000. Stippling indicates statistically significant correlations at the 95% confidence level (taking into account the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995)).

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

As for paleo-NorCPM1 wind nudging (Figure S6), the overall spatial patterns of correlation with independent datasets remain broadly similar to those obtained with paleo-NorCPM1 offline – albeit the correlation coefficients are generally reduced compared with the offline reconstruction (Figures 2 and S7). This reduced agreement likely reflects several factors: i) paleo-NorCPM1 wind nudging reflects the climate variability as the dynamical response of the climate system to the reconstructed atmospheric circulation and external forcing, thereby neglecting processes that are not directly associated with the imposed atmospheric variability (e.g. oceanic processes); ii) unlike paleo-NorCPM1 offline, the wind nudging simulations are integrated continuously over 400 years without corrections from the DA system, making paleo-NorCPM1 wind nudging prone to the accumulation of model errors; iii) the wind nudging ensemble consists of 10 members, compared with 30 members for paleo-NorCPM1 offline, which also contributes to slightly reduced agreement with independent datasets; ; iv) in contrast, paleo-NorCPM1 offline directly assimilates proxy information at each time step, which may lead to a stronger agreement with reference datasets partly due to the assimilation procedure itself. Despite these limitations, paleo-NorCPM1 wind nudging maintains a comparable level of skill across most variables, indicating that constraining the atmospheric circulation provides a strong large-scale constraint on climate variability. Moreover, by allowing the ocean to dynamically adjust to the reconstructed wind variability, along with the associated atmosphere-ocean feedback, paleo-NorCPM1 wind nudging improves the representation of SST variability in regions characterized by strong ocean circulation. In particular, SST variability in the North Pacific and North Atlantic Oceans is slightly better reproduced compared with paleo-NorCPM1 offline (Figures 2 and S6), illustrating the added value of the wind nudging framework for capturing climate variability in dynamically active ocean regions (e.g., Deser et al., 2010). To note, we observe in paleo-NorCPM1 wind nudging an important degradation in the high-latitude regions of the Southern Hemisphere, due to an overly strong dynamical response of the Southern Ocean to the imposed atmospheric circulation. This highlights the combined influence of uncertainties in the reconstructed wind fields and the model sensitivity in this region (e.g., Sallée et al., 2025). As a consequence, paleo-NorCPM1 wind nudging exhibits reduced skill in representing hydroclimate variability over Antarctica (this is discussed in more detail in Sections 4.2.2 and 4.2.5).

As a further detailed, and independent evaluation over the 20th century, we assess the performance of paleo-NorCPM1 offline in reproducing key integrated variables, including continental temperature, precipitation, SPEI over key regions as well as major modes of atmospheric and climate variability (Table 3). It should be noted that the ECR calculation does not account for uncertainties in the reference datasets. Therefore, ECR values are likely overestimated, and our assessment of the ensemble spread is thus conservative.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Table 3. Evaluation of the paleo-NorCPM1 offline (ensemble mean) for selected integrated climate variables and large-scale climate modes. The temporal correlation coefficients (stars indicate statistically significant under the 95% confidence level), coefficients of efficiency and ensemble calibration ratio are displayed. The annual time-series are considered, otherwise specified. The corresponding reference dataset and the evaluation period are also indicated.

	paleo-NorCPM1 offline			Validation dataset
	corr	ce	ecr	
Global mean surface temperature	0.92*	0.82	0.62	HadCRUT5 (1901-2000)
Northern Hemisphere mean surface temperature	0.88*	0.75	0.75	HadCRUT5 (1901-2000)
Southern Hemisphere mean surface temperature	0.92*	0.68	0.68	HadCRUT5 (1901-2000)
Tropical mean surface air temperature	0.91*	0.80	0.61	HadCRUT5 (1901-2000)
Arctic mean surface temperature	0.59*	0.18	0.63	HadCRUT5 (1958-2000)
Nino3.4	0.88*	0.71	0.11	ERSSTv5 (1901-2000)
Interdecadal Pacific Oscillation	0.85*	0.69	0.31	ERSSTv5 (1901-2000)
Walker Circulation	0.48*	0.22	1.36	20CRv3 (1901-2000)
Indian Ocean Dipole	0.41*	0.16	0.46	ERSSTv5 (1901-2000)
Atlantic Multi-Decadal Variability	0.73*	0.37	1.10	ERSSTv5 (1901-2000)
Winter North Atlantic Oscillation	0.41*	0.16	0.49	20CRv3 (1901-2000)
Summer North Atlantic Oscillation	0.44*	0.19	0.20	20CRv3 (1901-2000)
Summer European mean surface temperature	0.62*	0.38	0.28	HadCRUT5 (1901-2000)
Summer European SPEI	0.48*	0.08	0.95	HadCRUT5-gswp3 (1950-2000)
Southern Annular Mode	0.64*	0.40	0.43	Marshall2003 (1958-2000)
Antarctic mean surface temperature	0.53*	0.20	0.47	Dalaiden2025 (1958-2000)
Antarctic mean snow accumulation	0.64*	0.35	0.86	gswp3 (1958-2000)
Greenland mean surface temperature	0.50*	0.20	0.30	ERA5 (1958-2000)
Greenland mean snow accumulation	0.64*	0.35	1.11	gswp3 (1958-2000)
Antarctic sea-ice extent	0.54*	0.23	0.10	HadISST (1979-2000)
Arctic sea-ice extent	0.80*	0.63	1.11	HadISST (1958-2000)

Based on these criteria, paleo-NorCPM1 offline demonstrates substantial skill in reproducing large-scale temperature variability and major climate modes, including Pacific modes such as ENSO (e.g., Niño3.4 index), IPO, and IOD, the Walker circulation, and Atlantic Multidecadal Variability (AMV). The reconstructed winter and summer North Atlantic Oscillation (NAO), as well as European continental temperature and SPEI, show agreement with reference datasets, highlighting the potential of paleo-NorCPM1 offline for investigating European hydroclimate variability. The reanalysis also performs well in polar regions, with evaluation metrics indicating a realistic representation of near-surface air temperature over the Arctic,

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Greenland and Antarctic ice sheets, and total integrated precipitation over the two ice sheets. Furthermore, the reconstructed Southern Annular Mode is in good agreement with the station-based index of Marshall (2003), indicating that paleo-NorCPM1 offline accurately captures variability in Southern Hemisphere westerly winds. Finally, sea-ice concentration is well reproduced in both the Arctic and Antarctic regions when compared with satellite-based products.

Several caveats regarding the assessment of the performance of the paleo-NorCPM1 offline should be noted. Although the evaluation is designed to be independent, partial overlap may exist between datasets used for the calibration of the PSM and those employed here for the validation. For instance, the datasets from Dalaiden et al. (2025) and Marshall (2003) are based on the temperature and atmospheric pressure records from Antarctic weather stations, which are also used in the datasets for the calibration of the PSMs for Antarctic ice-core records. In addition, reference datasets themselves are subject to uncertainties, and the evaluation should therefore be interpreted as a comparison rather than a strict evaluation. Finally, agreement between paleo-NorCPM1 and reference datasets is generally higher during the second half of the 20$^{\text{th}}$ century (not shown). This likely reflects (i) increased proxy availability over that period (Figure 1), (ii) improved reliability and spatial coverage of instrumental observations, and (iii) partial overlap with the calibration period of the PSMs, which may lead to some degree of overfitting.

4.2 Comparison with other reconstructions

4.2.1 Large-scale temperatures and atmospheric circulation since 1500

In addition to the evaluation using instrumental datasets, we compare paleo-NorCPM1 with other paleo-based reconstructions. More specifically, we consider three widely used products in the paleo community to study climate over the past centuries. The first is the Modern Era Reanalysis (ModE-RA) (Valler et al., 2024), which comprises 20 atmospheric reconstructions constrained by both long-term instrumental time series and paleoclimate records, using an offline data assimilation framework. The overall approach of ModE-RA is therefore similar to that of paleo-NorCPM1. However, several important differences should be noted: i) paleo-NorCPM1 relies exclusively on paleoclimate records, whereas ModE-RA also incorporates instrumental observations; ii) the ModE-RA prior is based solely on atmospheric model simulations, while paleo-NorCPM1 employs a fully coupled ESM; iii) paleo-NorCPM1 explicitly assimilates a broader set of ice-core records, whereas ModE-RA includes only limited ice-core information (mainly a few records from Greenland).

The second reconstruction considered is the Last Millennium Reanalysis version 2 (LMR; Tardif et al., 2019), hereafter referred to as LMR Offline. This product also uses an offline EnKF DA framework to combine model simulations with a paleoclimate database. The proxy database is broadly similar to that used in paleo-NorCPM1, but paleo-NorCPM1 includes a more extensive set of ice-core records, particularly expanded isotopic datasets (Konecky et al., 2020), as well as snow accumulation records (Thomas et al., 2017; Box et al., 2013) and Antarctic sea-salt records (Thomas et al., 2022). To construct the prior, LMR Offline relies on model simulations spanning the period 850–1850, from which individual years are randomly sampled. As noted by Perkins and Hakim (2017), this approach may bias the reanalysis toward the preindustrial climatology and damp low-frequency variability during periods with sparse observational coverage. To partly address this limitation, Perkins and Hakim (2017) developed a Linear Inverse Model (LIM) trained on simulations spanning the last millennium to represent

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

oceanic influences in the reconstruction. The LIM is trained using several oceanic variables, including SST, sea surface salinity, sea surface height, and 0-700m ocean heat content. The LIM is coupled to LMR to represent linear atmosphere–ocean interactions, yielding a third product referred to here as LMR Online. The proxy database used in LMR Online is similar to that of LMR Offline, but excludes the tree-ring records from Breitenmoser et al. (2014).

Finally, all three reconstructions differ from paleo-NorCPM1 in how they treat observational uncertainty. As described in sec. 2.5, paleo-NorCPM1 estimates observational error using 1,000 realizations based on subsampling data points during PSM calibration. This procedure yields a more conservative representation of observational uncertainty than in the other reconstructions and leads to a more cautious weighting of proxy information within the DA system.

Figure 3 shows the evolution of global, Northern Hemisphere, Southern Hemisphere, and tropical (30°) near-surface air temperature over the period 1500–2000 from paleo-NorCPM1 offline, paleo-NorCPM1 wind nudging, ModE-RA, LMR Offline, and LMR Online, together with HadCRUT5 for comparison.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

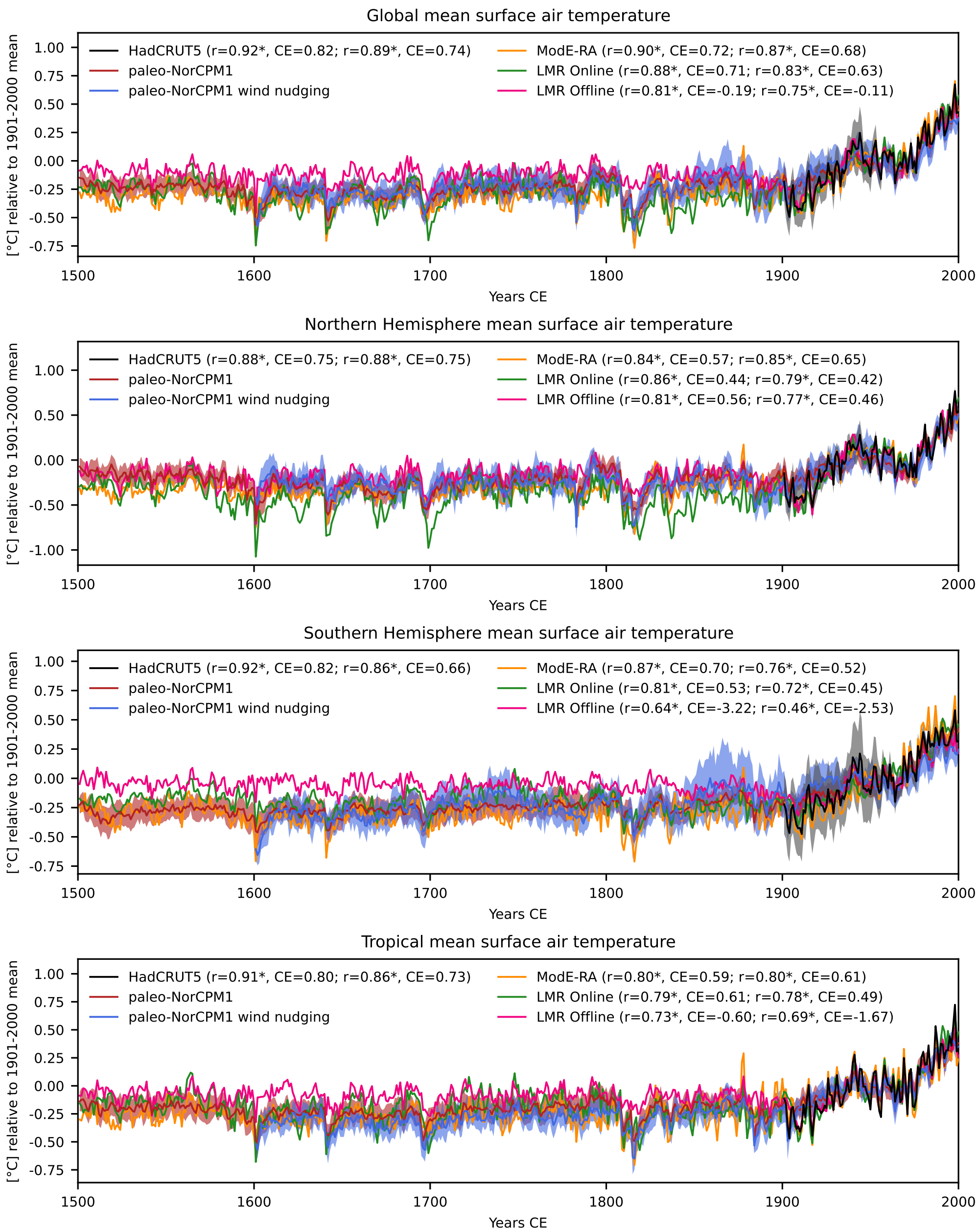

Figure 3. Global, Northern Hemisphere, Southern Hemisphere, and tropical (30°S-30°N) near-surface air temperature anomalies from paleo-NorCPM1 offline, paleo-NorCPM1 wind nudging, ModE-RA, LMR Offline, and LMR Online for the period 1500–2000. HadCRUT5 is also shown over the 20th century, including its 95% uncertainty range (except for the tropical region). For paleo-NorCPM1 offline and paleo-NorCPM1 wind nudging, the ±1 standard deviation ensemble spread is plotted. All time series are expressed relative to the 20th century mean. The temporal correlation coefficient and coefficient of efficiency between paleo-NorCPM1 offline (and paleo-NorCPM1 wind nudging) and the reference datasets are reported in the legend (first and second entries, respectively).

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Overall, all reconstructions capture the major features of 20th century temperature evolution, including the pronounced warming trend, the mid-20th century cooling in the early 1950s, and the strong warming afterwards beginning around 1970. The close agreement between paleo-NorCPM1 wind nudging and the other reconstructions indicates that wind-driven variability combined with external forcing is sufficient to reproduce the main features of large-scale temperature variability.

Prior to the 20th century, differences among the reconstructions are more pronounced. For global mean near-surface air temperature, both paleo-NorCPM1 offline and paleo-NorCPM1 wind nudging exhibit high correlation with the other reconstructions (mean correlations exceeding 0.8), indicating strong overall agreement. The highest consistency is found in the Northern Hemisphere, likely reflecting the dense proxy network dominated by tree-ring records. As a note, LMR Online slightly deviates from the other reconstructions following major volcanic eruptions, exhibiting a stronger cooling response, and the associated impact on the low-frequency variability thereafter. In the Southern Hemisphere, although temporal variability is generally similar across reconstructions, LMR Offline appears systematically warmer than the others. One possible explanation is the relatively sparse observational network in the Southern Hemisphere, together with methodological differences that may favor the pre-industrial climatology where observational constraints are weak. This can ultimately cause an underestimation of the anthropogenic signal during the 20th century. A similar, though weaker, behavior is also observed in the tropics. The reduced temperature differences in LMR Online with the other reconstructions suggest that explicitly accounting for ocean dynamics helps mitigate this potential bias, consistent with the findings of Perkins and Hakim (2017). Although ModE-RA does not assimilate ice-core records, its prior is based on an ensemble forced by external forcing and historical sea surface conditions, which may explain its closer agreement with LMR Online and paleo-NorCPM1.

A distinctive feature of paleo-NorCPM1 wind nudging is its more pronounced low-frequency variability in the Southern Hemisphere, particularly during the 18th century and the late 19th century. This enhanced variability likely arises from wind-driven ocean dynamics – known to be especially strong in the Southern Ocean (Sallée et al., 2025) – which modulate surface temperature on multi-decadal timescales. This behavior is further discussed in Sections 4.2.2 and 4.2.5.

Finally, Figure 4 shows maps of temporal correlation coefficients between paleo-NorCPM1 offline and the three other reconstructions for near-surface air temperature and 500-hPa geopotential height over the period 1601–2000. Overall, paleo-NorCPM1 offline exhibits strong agreement with all three reconstructions, with nearly all grid points displaying positive correlations for both variables. This indicates that paleo-NorCPM1 offline provides a reconstruction of comparable quality to these reconstructions. The highest agreement is found over the Pacific, particularly for 500-hPa geopotential height, highlighting the consistent representation of large-scale atmospheric circulation, including Hadley cell dynamics. To note, paleo-NorCPM1 offline shows the strongest agreement with LMR Online, with a spatially averaged correlation coefficient of 0.58, compared with 0.49 and 0.46 for LMR Offline and ModE-RA, respectively, for near-surface air temperature. Because LMR Online incorporates a LIM to represent the effect of the slow response of the climate system (especially the ocean), this closer agreement suggests that paleo-NorCPM1 offline also captures a component related to oceanic processes. This result is consistent with the prior constructed from a fully coupled ESM and a hybrid covariance approach in paleo-NorCPM1, which allows ocean–atmosphere interactions to be implicitly represented even within an offline assimilation framework (with the linear approximation of the data assimilation step).

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

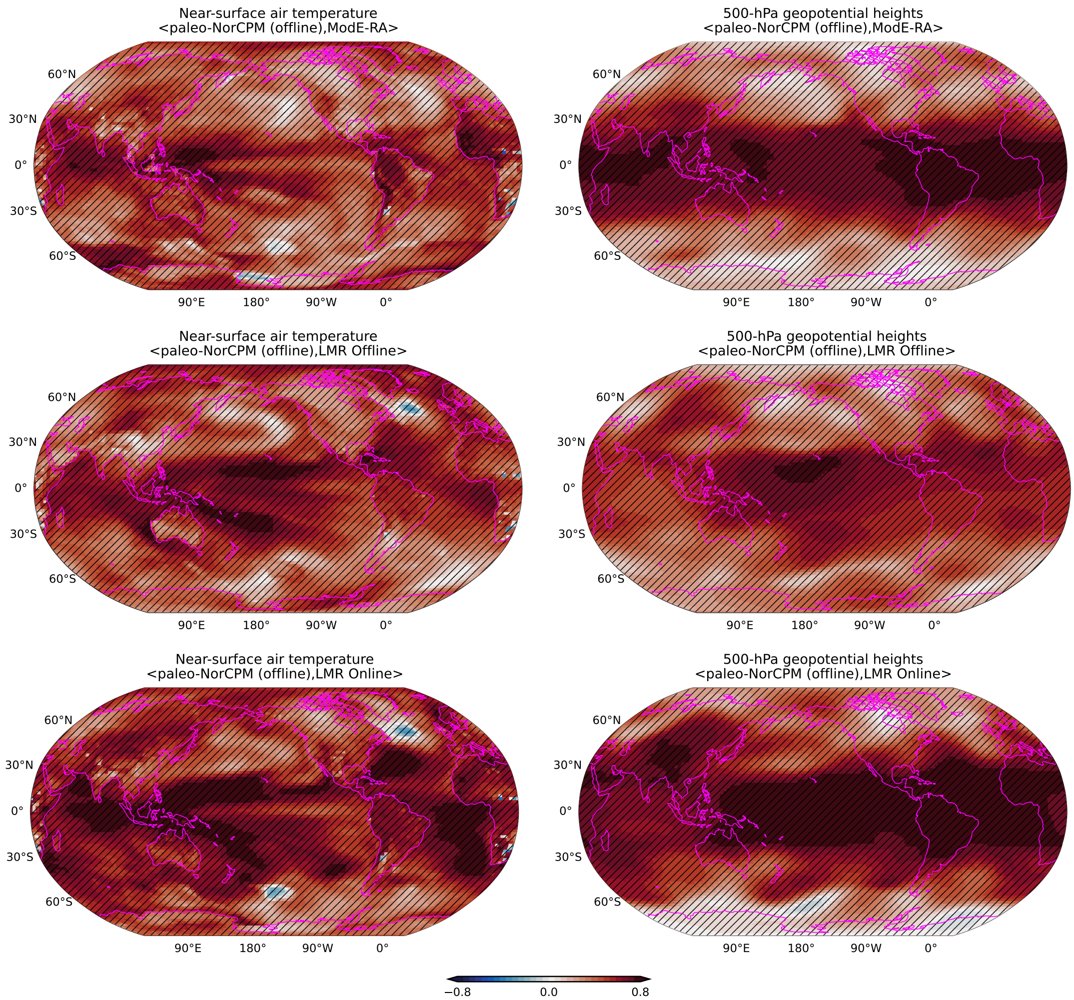

Figure 4. Maps of the correlation coefficients for paleo-NorCPM1 offline (ensemble mean), evaluated against ModE-RA, LMR Offline and LMR Online for annual near-surface air temperature and 500-hPa geopotential height over 1601–2000. Stippling indicates statistically significant correlations at the 95% confidence level (taking into account the False Discovery Rate; FDR).

4.2.2 Ocean Heat Content

After evaluating the performance of paleo-NorCPM1 for the atmosphere and surface, we next assess its ability to reproduce subsurface ocean variability by analyzing the Ocean Heat Content (OHC) integrated over the top 700 m (Figure 5). Direct observations of OHC are sparse prior to the deployment of the Argo float network in the late 1990s (Riser et al., 2016), making the evaluation of OHC from paleo-NorCPM1 highly challenging. To place paleo-NorCPM1 in context, we compare it with three products. The first is the CoRea1860+ climate reanalysis (Wang et al., 2025), which is based on NorCPM and assimilates instrumental SST observations from 1860. CoRea1860+ has been shown to provide reliable estimates of historical ocean variability. For the earlier period, we use the ocean reconstruction of Dalaiden et al. (2025), which is driven by hybrid atmospheric forcing derived from ModE-RA and an Antarctic atmospheric reconstruction. This reconstruction, therefore, reflects the oceanic response to historical atmospheric variability. As such, the two-way interactions between the ocean and atmosphere are not explicitly represented, potentially affecting the amplitude and phase of simulated variability. Finally, LMR Online is also included in the assessment as it provides a reconstruction for the OHC.

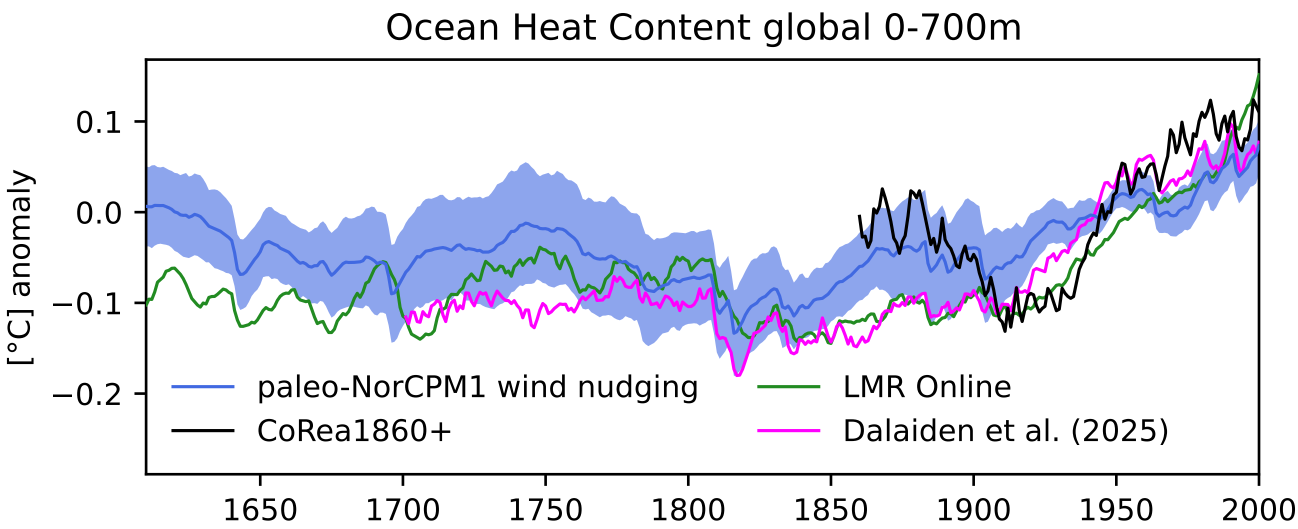

Figure 5. Temporal evolution of the anomalies of global Ocean Heat Content integrated over the top 700 m from 1700 to 2000 (expressed in °C). Anomalies are calculated from the 20th century mean.

Only paleo-NorCPM1 wind nudging is evaluated here, as paleo-NorCPM1 offline does not explicitly take into account the ocean dynamical adjustment to atmospheric variability. Over the 20th century, all reconstructions consistently show a pronounced increase in global OHC (Figure 5), confirming the robust signature of long-term ocean warming. Although paleo-NorCPM1 exhibits a weaker warming trend (0.11 ±0.09 °C per century) compared with CoRea1860+ (0.27 °C per century), Dalaiden et al. (2025) (0.20 °C per century), and LMR Online (0.24 °C per century), it reproduces both the overall warming and the decadal-scale modulation, including the slowdown in warming during the early 1970s.

Prior to the 20th century, paleo-NorCPM1 displays more pronounced low-frequency variability in global ocean heat content than the other reconstructions, although similar variability is also present in the reconstruction from Dalaiden et al. (2025) and LMR Online, but in a lesser extent. This difference likely reflects methodological differences in the representation of ocean dynamics, which might ultimately impact the climate sensitivity. In particular, the wind-driven ocean adjustment can generate substantial multi-decadal variability through changes in ocean circulation and heat uptake, which may not be fully captured in reconstructions where the ocean responds to atmospheric variability without feedback (e.g., Dalaiden et al., 2025). In the same vein, the LIM framework used in LMR Online provides a statistical representation of oceanic processes, in particular the slow response of the ocean, which may underestimate the magnitude of low-frequency variability compared with fully dynamical simulations. Moreover, ESMs tend to underestimate ocean variability when compared with observations (Hébert and Laepple,

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

2025), in particular for periods characterized by weak external forcing. Yet, these simulations are used for training the LIM. This may contribute to the comparatively smaller low-frequency variability seen in LMR Online relative to paleo-NorCPM1.

Despite the noted differences, paleo-NorCPM1 remains broadly consistent with the other reconstructions, particularly after 1700, when the two reconstructions fall within the ensemble envelope of paleo-NorCPM1. The larger amplitude variability simulated by paleo-NorCPM1 likely reflects the explicit representation of the coupling between the ocean and atmosphere. In this respect, the reconstruction may help bridge part of the gap between proxy-based estimates of low-frequency variability and free model simulations, which often exhibit weaker multi-decadal variability (Laepple et al., 2023; Hébert and Laepple, 2025; Skiba et al., 2026). These results indicate that paleo-NorCPM1 provides a credible and dynamically consistent reconstruction of historical ocean heat content, capturing both the long-term warming trend and low-frequency variability excited by the atmospheric dynamics.

As an important feature, paleo-NorCPM1 simulates pronounced low-frequency variability in Southern Ocean heat content (Figure S8). The dynamics of the Southern Ocean – including Ekman transport, vertical upwelling, and dynamics of the Antarctic Circumpolar Current – is primarily driven by surface wind variability (e.g., Frölicher et al., 2015; Armour et al., 2016). Variations in wind stress directly modulate ocean circulation and vertical heat redistribution, making Southern Ocean heat content particularly sensitive to atmospheric forcing (e.g., Frölicher et al., 2015). Because NorESM1 is dynamically constrained by historical wind variability from paleo-NorCPM1 offline, it allows the ocean to adjust consistently to the prescribed wind variability, resulting in pronounced multi-decadal variability. The reconstruction from Dalaiden et al. (2025) also exhibits substantial low-frequency variability, although with differences in phase, while LMR Online shows much weaker variability and only a modest warming trend over the 20th century. These differences illustrate the large impact of wind variability on the OHC, which is highly pronounced in paleo-NorCPM1, most probably because of the active feedback between atmosphere and ocean. Moreover, the resulting historical variability might depend on the climate model used in the reconstruction.

4.2.3 Multi-decadal variability in the Atlantic and Pacific basins

Modes of climate variability operating on multi-decadal timescales play a critical role in the global climate variability, including their influence on high-frequency modes such as the El Niño Southern Oscillation (ENSO) and NAO. Additionally, variability in the Atlantic and Pacific Oceans has been linked to periods of accelerated and reduced global surface warming during the 20th century (Kosaka and Xie, 2013; England et al., 2014; Dai et al., 2015; DelSole et al., 2011). Accurately reproducing these modes is therefore essential for assessing the ability of paleo-NorCPM1 to represent internally generated climate variability. To this end, we evaluate the Atlantic Multidecadal Variability (AMV) and the Interdecadal Pacific Oscillation (IPO) by comparing reconstructed indices from paleo-NorCPM1 with instrumental estimates derived from HadSST2 and ERSSTv5 over the period 1850–2000 (Figure 6).

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

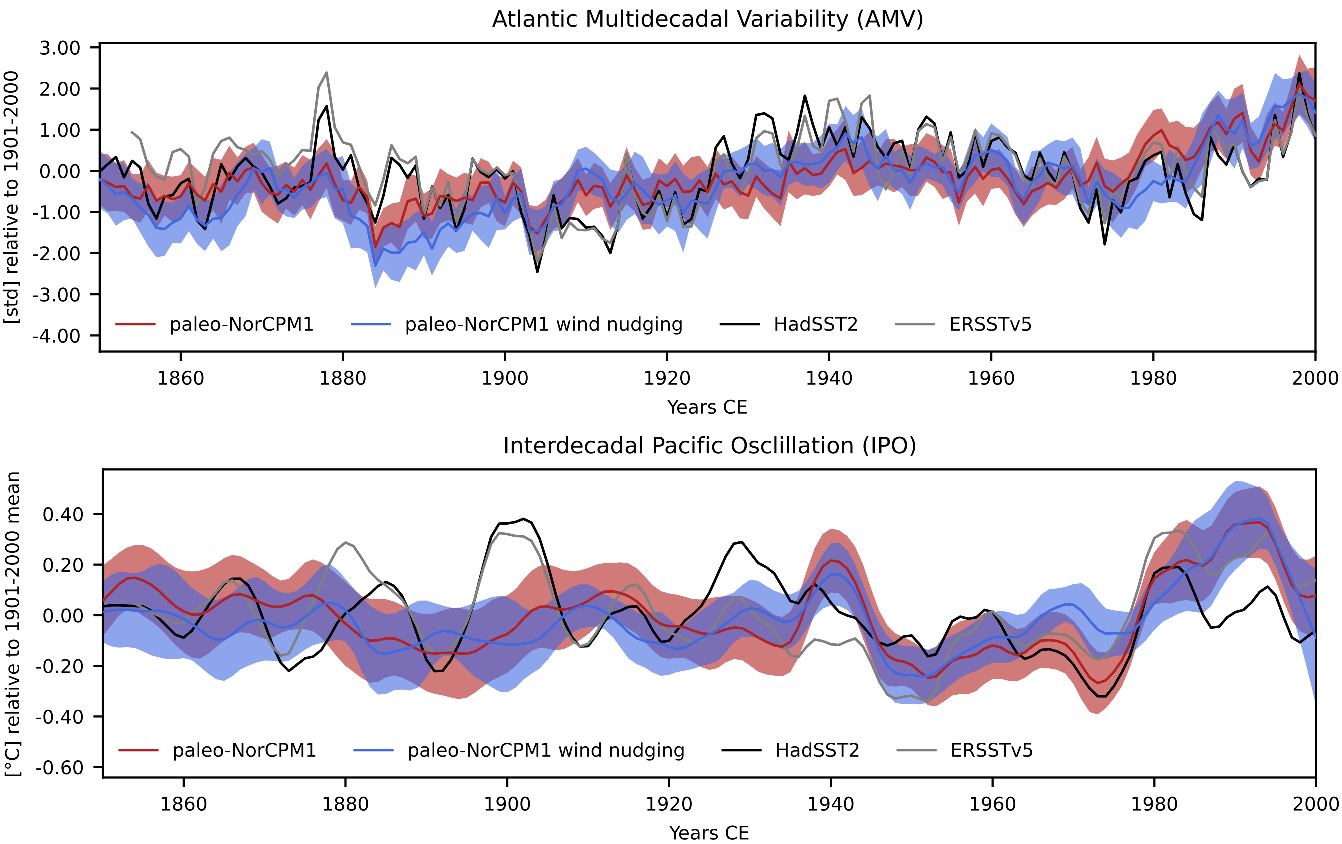

Figure 6. Temporal evolution of the Atlantic Multidecadal Variability (AMV) and Interdecadal Pacific Oscillation (IPO) indices over the period 1850–2000 for paleo-NorCPM1 offline, paleo-NorCPM1 wind nudging, HadSST2, and ERSSTv5. Indices are expressed as anomalies relative to the 20thcentury mean

The AMV index is defined as the spatial average of SST anomalies over the North Atlantic (0-60°N, 80°W-0°). To isolate the internal component of the AMV, the externally forced component is removed using the mean of the 30-member NorESM1-ME ensemble, which provides an estimate of the forced response (to remove any residual internal variability in this estimate, a 100-year low-pass filter is applied to obtain the final forced estimation). This approach avoids potential artefacts associated with removing the global mean temperature, which is commonly used as an estimate for the forced response but may introduce biases in the estimated AMV signal (Deser and Phillips, 2021). The resulting AMV index is normalized relative to the 20th century mean and standard deviation.

Overall, both paleo-NorCPM1 offline and wind nudging reproduce the observed AMV variability derived from HadSST2 and ERSSTv5 (Figure 6). In particular, the two ensembles of paleo-NorCPM1 capture the major multi-decadal phases of the AMV, including the positive phase from the early 20th century to about 1950, the subsequent negative phase until the 1970s, and the renewed positive phase thereafter. While paleo-NorCPM1 successfully reproduces the timing of these phase transitions, the amplitude of the variability is slightly underestimated relative to the instrumental estimates. This reduced amplitude likely

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

reflects the lack of oceanic proxy records directly constraining North Atlantic SST variability, such that the reconstruction primarily relies on continental proxy information and atmospheric teleconnections. Additionally, paleo-NorCPM1 wind nudging solely reflects the response of wind variability, without directly constraining the ocean state through ocean observations. As an important note, the paleo-NorCPM1 wind nudging exhibits stronger decadal variability than paleo-NorCPM1 offline (standard deviation of 0.66°C versus 0.56°C over the 1601–2000 period), highlighting the role of dynamical adjustment to atmospheric circulation variability. By explicitly constraining wind variability, the wind-nudged configuration allows the coupled ocean–atmosphere system to respond dynamically, thereby enhancing the representation of internally generated multi-decadal variability. This result underscores the importance of ocean dynamical processes in shaping AMV variability.

As for the IPO, we use the Tripole Index developed by Henley et al. (2015), defined as the mean SST anomaly over the central Pacific (10°S-10°N, 170°E-90°W) minus the average SST anomalies over the North Pacific (25°-45°N, 140°E-145°W) and South Pacific (50°-15°S, 150°E-160°W). Following Henley et al. (2015), a 13-year low-pass filter is applied to isolate decadal variability. Overall, paleo-NorCPM1 offline and wind nudging are in good agreement with each other (Figure 6), reproducing similar variability, including the positive anomalies during the early 1940s associated with a sequence of strong El Niño events between 1940 and 1942 (Brönnimann et al., 2004), as well as the positive phase during the late 20th century. Compared with AMV, however, agreement between the IPO indices from paleo-NorCPM1 and instrumental IPO indices is weaker, and substantial differences are also evident between HadSST2 and ERSSTv5, even during the recent decades. In particular, the IPO index derived from ERSSTv5 exhibits a weak positive anomaly in the early 1940s compared with paleo-NorCPM1 and HadSST2. Further investigation indicates that this discrepancy does not arise from differences in ENSO variability, as the Nino3.4 index shows excellent agreement between paleo-NorCPM1 and instrumental datasets (Figure S9). Instead, the differences originate from more widespread positive SST anomalies in the central Pacific in paleo-NorCPM1 compared with the instrumental datasets (Figure S10), suggesting differences in the large-scale spatial structure of Pacific variability. NorESM1 exhibits a known bias in the representation of tropical Pacific variability, whereby the SST anomaly patterns associated with ENSO are too strongly confined to the central Pacific (Bentsen et al., 2013), thereby underestimating their remote influence in the North and South Pacific. It is important to mention that instrumental SST products prior to 1950 are subject to large uncertainties due to sparse observational coverage, which is even more pronounced in the mid-latitude regions of the Southern Hemisphere Pacific (Deser et al., 2010; Yasunaka and Hanawa, 2011). These uncertainties limit the reliability of instrumental IPO estimates during the early 20th century and might partly explain the discrepancies between datasets. Similarly to AMV, paleo-NorCPM1 wind nudging exhibits enhanced SST variability in the North Pacific (Figure S10), reflecting the dynamical response of the ocean to wind variability, particularly here through modulation of the North Pacific Gyre.

4.2.4 Atlantic Meridional Overturning Circulation

The Atlantic Meridional Overturning Circulation (AMOC) is a fundamental component of the climate system, characterized by the northward transport of warm, saline surface waters and the southward return flow of cold, fresh waters at depth (e.g., Trenberth and Solomon, 1994). Through this overturning circulation, the AMOC plays a central role in redistributing heat within the Atlantic basin and modulating regional and global climate variability (e.g., Zhang et al., 2019). The AMOC is

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

commonly quantified as the zonally integrated meridional overturning streamfunction, with the maximum transport at 26°N serving as a standard metric of its strength. Although continuous observations from the RAPID array provide direct estimates of AMOC variability since 2004 (McCarthy et al., 2015), the paleo-NorCPM1 reconstruction ends in 2000, preventing direct comparison.

Figure 7 shows the temporal evolution of AMOC anomalies in paleo-NorCPM1 wind nudging, compared with the CoRea1860+ reanalysis and the oceanic reconstruction of Dalaiden et al. (2025). These reconstructions rely on different methodologies that influence the results. In paleo-NorCPM1 wind nudging, AMOC variability emerges dynamically in response to historical wind variability, which is known to be a primary driver of AMOC variability through its influence on momentum and buoyancy forcing (Yeager and Danabasoglu, 2014; Buckley and Marshall, 2016). In contrast, the reconstruction of Dalaiden et al. (2025) is based on an ocean–sea-ice model forced by prescribed atmospheric forcing, therefore without atmosphere–ocean feedback. CoRea1860+, meanwhile, reconstructs the climate state through assimilation of SST observations. Because AMOC variations strongly influence North Atlantic SST through changes in ocean heat transport (e.g., Zhang et al., 2019), AMOC variability in CoRea1860+ is constrained through its imprint on SST.

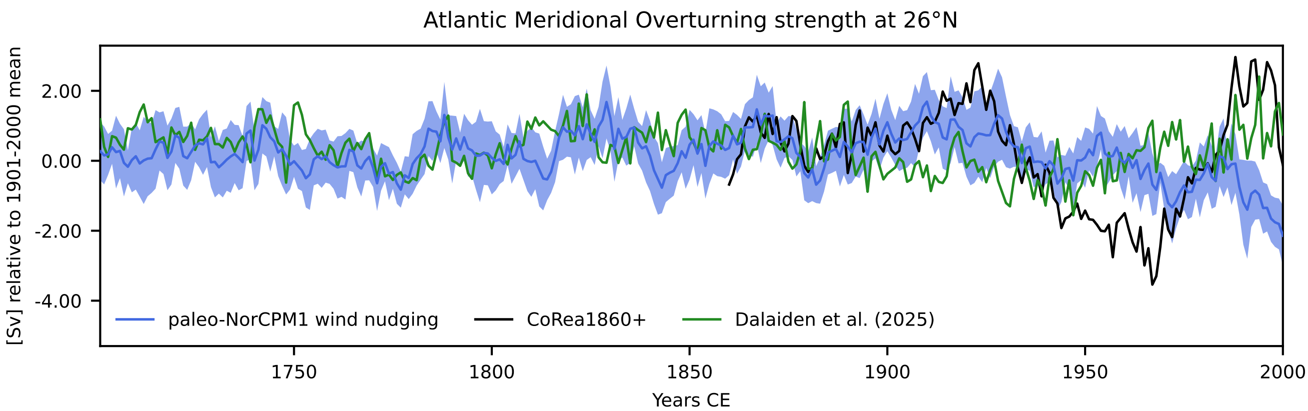

Figure 7. Time-series of the anomalies of the Atlantic Meridional Overturning Circulation strength at 26°N in paleo-NorCPM1 wind nudging, CoRea1860+ and in the reconstruction of Dalaiden et al. (2025). Anomalies are computed relative to the 20th century reference.

Over the 20th century, paleo-NorCPM1 wind nudging simulates a statistically significant AMOC weakening of -2.39 Sv per century, whereas CoRea1860+ shows no statistically significant trend (-0.81 Sv per century), and the reconstruction of Dalaiden et al. (2025) indicates a statistically significant strengthening (1.32 Sv per century; Figure 7). As an important note, when extending the period for calculating the trend to 2020, both CoRea1860+ and the reconstruction of Dalaiden et al. (2025) show AMOC weakening. Despite these differences in the 20th century, paleo-NorCPM1 wind nudging indicates strong multi-decadal variability. In particular, both paleo-NorCPM1 wind nudging and CoRea1860+ show a strengthening of the AMOC from the late 19th century to about 1930. Thereafter, CoRea1860+ exhibits a pronounced weakening until the 1970s, while paleo-NorCPM1 wind nudging simulates a more moderate decline superimposed on pronounced multi-decadal variability.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

From the 1990s, paleo-NorCPM1 wind nudging indicates a decline in AMOC strength that is not present in CoRea1860+ and in the reconstruction of Dalaiden et al. (2025). Such a decline is consistent with the weakening of the AMOC simulated by climate models in response to anthropogenic forcing (Weijer et al., 2020). Overall, paleo-NorCPM1 wind nudging suggests a weakening of the AMOC over the 20th century, consistent with proxy-based reconstructions (e.g., Caesar et al., 2018). However, large uncertainties remain (Fox-Kemper et al., 2021), and a detailed investigation of the underlying mechanisms is required to assess the robustness and causes of this long-term decline. It is worth noting that the multi-decadal AMOC variability simulated in paleo-NorCPM1 wind nudging closely corresponds to variability in North Atlantic SST, supporting the central role of the AMOC in driving multi-decadal SST variability in the North Atlantic (Zhang et al., 2019).

4.2.5 Sea-ice extent

Sea ice plays a key role in the global climate system. It contributes to the global overturning circulation through brine rejection during sea-ice formation, which influences ocean density and deep water formation (Fox-Kemper et al., 2021). In addition, sea ice strongly affects the Earth energy balance by reflecting a large fraction of incoming solar radiation due to its high albedo (e.g., Goosse et al., 2018). Figure 8 shows the temporal evolution of Arctic and Antarctic sea-ice extent over the past centuries in paleo-NorCPM1 offline and wind nudging, together with other reconstructions.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

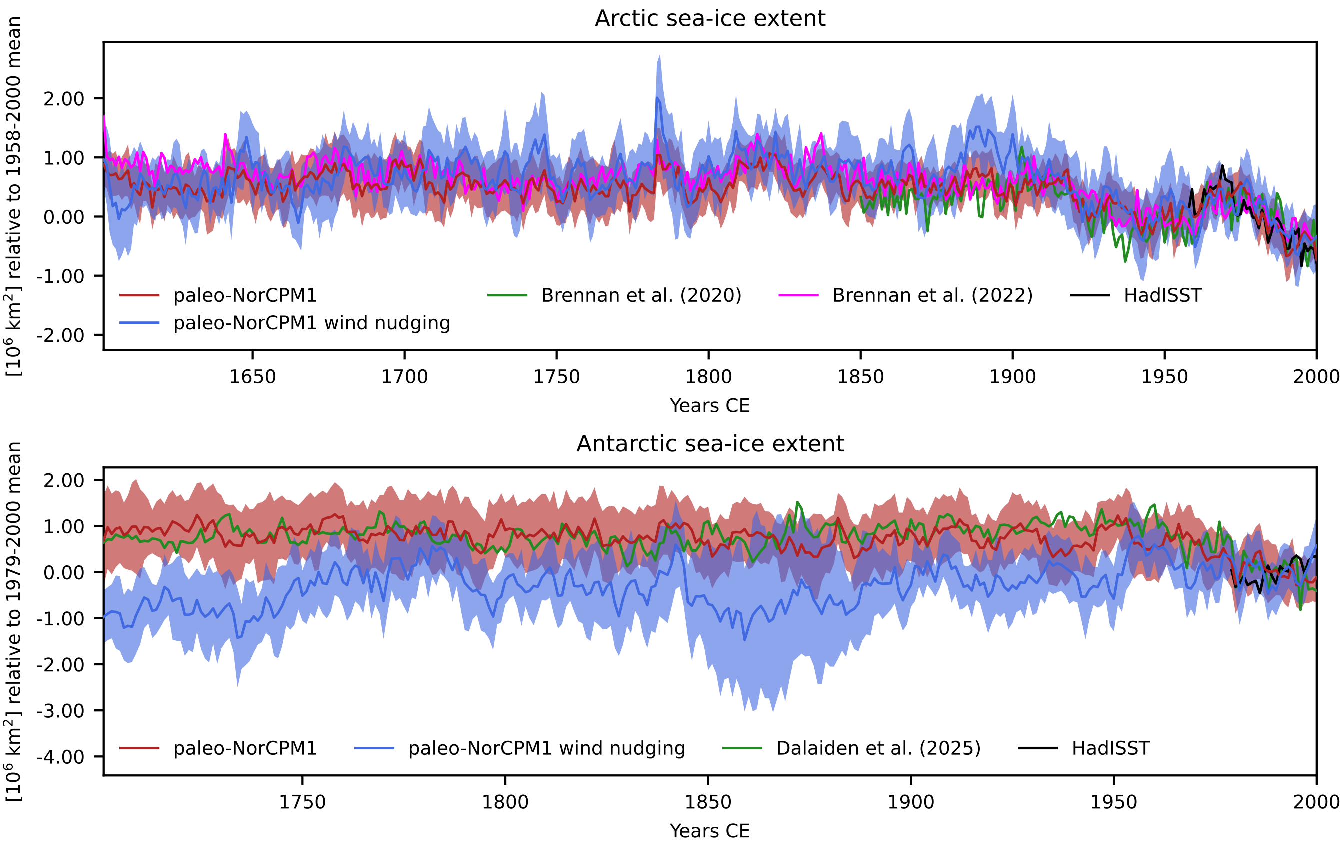

Figure 8. Time series of annual sea-ice extent anomalies in the Arctic and Antarctic from paleo-NorCPM1 offline and wind nudging. For the Arctic, the instrumental reconstruction of Brennan et al. (2020) and the paleo-based reconstruction of Brennan and Hakim (2022) are shown. For the Antarctic, the oceanic reconstruction of Dalaiden et al. (2025) is included. HadISST is also shown from 1958 onward for the Arctic and from 1979 onward for the Antarctic. Anomalies are computed relative to the 1958–2000 period for the Arctic and the 1979–2000 period for the Antarctic.

For the Arctic, we compare paleo-NorCPM1 with the instrumental reconstruction of Brennan et al. (2020) and the paleo-based reconstruction of Brennan and Hakim (2022), both derived using the Last Millennium Reanalysis (LMR) framework. The reconstruction of Brennan et al. (2020) uses Arctic temperature observations from HadCRUT to constrain sea-ice variability, whereas Brennan and Hakim (2022) relies on the PAGES2k database (Emile-Geay et al., 2017). Although two reconstructions are available for each (using two different model priors), we selected the version based on the Community Climate System Model version 4 (CCSM4), as differences between prior choices are minimal. Because climate models tend to underestimate the sensitivity of Arctic sea ice to temperature changes compared with observations (Rosenblum and Eisenman, 2017), Brennan et al. (2020) and (Brennan and Hakim, 2022) applied a covariance inflation factor of 2.6 between sea-ice concentration and temperature. This correction effectively increases the amplitude of reconstructed sea-ice variability to better match observed sensitivity. A similar approach has been employed here. In practice, we scale the variance of the Arctic sea-ice reconstruction

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

from paleo-NorCPM1 to match the variance of the observed Arctic sea-ice extent over the 1958–2000 period. HadISST sea-ice extent is shown from 1958 onward to provide an instrumental observational reference.

From 1958 onward, all reconstructions agree with HadISST (Figure 8), showing a modest sea-ice expansion until the early 1970s followed by a pronounced decline. During the early 20th century warming period (1920–1950), both paleo-NorCPM1 offline and wind nudging simulate a significant decline in sea-ice extent (-0.11 and -0.12 x 10^6 km^2 per decade, respectively; both statistically significant at the 95% level), consistent with the reconstructions of Brennan et al. (2020) (-0.13 x 10^6 km^2 per decade; statistically significant) and Brennan and Hakim (2022) (-0.15 x 10^6 km^2 per decade; statistically significant). This decline coincides with the early 20th century warming (Hegerl et al., 2018). Prior to the 20th century, both paleo-NorCPM1 and Brennan and Hakim (2022) indicate relatively stable sea-ice conditions, but with pronounced multi-decadal variability. Although differences between paleo-NorCPM1 offline and wind nudging remain small, paleo-NorCPM1 wind nudging exhibits stronger temporal variability, with a standard deviation of 0.16 10^6 km^2 compared with 0.12 10^6 km^2 for paleo-NorCPM1 offline. This enhanced variability reflects the dynamical adjustment, including the coupled ocean–atmosphere feedback, to wind variability. For example, paleo-NorCPM1 wind nudging simulates a higher sea-ice extent at the end of the 19th century compared with the offline reconstruction. Overall, the agreement with independent reconstructions and the physically consistent response to atmospheric variability support the credibility of paleo-NorCPM1 in reproducing Arctic sea-ice variability.

As for the Antarctic sea-ice extent, substantially larger differences are noticed between paleo-NorCPM1 offline and wind nudging (Figure 8). The offline reconstruction closely resembles the oceanic reconstruction of Dalaiden et al. (2025), with relatively stable sea-ice extent from 1700 to approximately 1960, followed by a marked decline afterwards (-0.69 × 10^6 km^2 per decade and -0.33 × 10^6 km^2 per decade for paleo-NorCPM1 offline and Dalaiden et al. (2025), respectively; both statistically significant at the 95% level). A similar, though weaker and statistically insignificant, decline is also simulated in paleo-NorCPM1 wind nudging (-0.17 × 10^6 km^2 per decade). Because continuous satellite observations of Antarctic sea ice are only available from 1979 onward (e.g., Hobbs et al., 2016), and paleo-NorCPM1 ends in 2000, the overlap with observational estimates remains too short to draw robust conclusions from direct comparison with HadISST.

In contrast to the offline reconstruction, paleo-NorCPM1 wind nudging exhibits pronounced low-frequency variability prior to 1950, accompanied by substantially larger ensemble spread for some periods, in particular from around 1840 to 1900. This enhanced variability reflects the impacts of the coupled ocean-atmosphere feedback. Regional analysis shows that these variations are primarily driven by changes in the Weddell Sea (70°W-15°W) and King Håkon sector (15°W-70°E; Figure S11). These regions are known to be associated with intense deep convection and strong ocean-atmosphere heat exchange (e.g., Zhou et al., 2026), which has been responsible for the opening of large-scale, open-ocean, multi-year polynyas, as observed in the 1970s (Comiso and Gordon, 1987). Consistent with this mechanism, pronounced positive mixed layer depth anomalies are found in the Atlantic sector (30°W-30°E, south of 60°S) between approximately 1850 and 1900 (Figure S12), indicating enhanced vertical mixing during this period. However, no well-defined open-ocean polynyas (i.e., fully surrounded by sea ice) are observed in paleo-NorCPM1 wind nudging. The enhanced vertical mixing rather results in a homogeneous sea cover decrease. Importantly, this behavior is not uniform across ensemble members since only two out of ten members exhibit particularly strong convection events (Figure S13), while others show variability more consistent with the reconstruction of

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Dalaiden et al. (2025), albeit with larger amplitude. This large ensemble spread highlights the strong sensitivity of Antarctic sea ice to wind-driven ocean dynamics and internal variability. Enhanced deep convection in the Southern Ocean is a known feature of climate models (e.g., Zhou et al., 2026), and has also been reported in forced ocean simulations, including experiments with surface salinity restoring (e.g., Dalaiden et al., 2025).

5 Summary and future perspectives

In this paper, we present and evaluate methodological developments that extend NorCPM to enable the assimilation of paleoclimate records and thereby extend climate reanalyses beyond the instrumental period. NorCPM was augmented with a comprehensive suite of PSMs that map climate variables into proxy space, allowing the integration of paleoclimate records from a wide range of proxy archives within a consistent data assimilation framework. To provide a dynamically consistent prior, we first generated a 30-member transient ensemble with NorESM1 covering 1500–1850 to complement the existing historical ensemble (Bethke et al., 2021). Using this fully coupled transient ensemble, we produced a new offline PDA reconstruction spanning 1500–2000 (paleo-NorCPM1 offline) by assimilating an extensive proxy database from multiple archives (Emile-Geay et al., 2017; Konecky et al., 2020; Breitenmoser et al., 2014; Thomas et al., 2017, 2022). The hybrid covariance approach employed in the DA system mitigates sampling error but still ensures consistency with external forcing and provides a physically informed prior, which is particularly relevant during periods and in regions characterized by sparse observational coverage. Compared with frameworks that draw priors from randomly sampled years of long control simulations, the use of a transient, year-specific forced ensemble helps retain externally forced signals even during periods with a weak observational network, while reducing sensitivity to an equilibrium climatology. Evaluation of paleo-NorCPM1 offline against independent instrumental datasets over the 20$^{\text{th}}$ century (20CRv3 (Slivinski et al., 2021), HadCRUT5 (Morice et al., 2021), ERSSTv5 (Huang et al., 2017), GSWP3 (Lange et al., 2021), HadISST, (Kennedy et al., 2019)) demonstrates robust skill in reproducing large-scale climate variability, major modes of climate variability, and key regional hydroclimate signals, including over Europe, the two polar ice sheets, the Pacific and Atlantic Oceans, as well as sea-ice extent. Comparisons with existing paleo-based reconstructions (ModE-RA (Valler et al., 2024), LMR Offline (Tardif et al., 2019), LMR Online (Perkins and Hakim, 2017)) show broadly comparable performance, providing confidence in both the methodological framework and the realism of the reconstructed variability.

As a step toward a fully coupled ocean-atmosphere paleo reanalysis, we conducted a 10-member NorESM1 ensemble over 1600–2000 in which the wind fields are nudged toward the atmospheric circulation reconstructed by paleo-NorCPM1 offline. This wind nudging ensemble (hereafter paleo-NorCPM1 wind nudging) allows the ocean to dynamically adjust to the reconstructed atmospheric variability and external forcing. In contrast to the offline reconstruction, where the ocean state is not dynamically updated between assimilation steps and is therefore not suited for representing ocean dynamics, the wind nudging framework provides a continuous, dynamically consistent ocean evolution driven by the reconstructed atmospheric circulation. Paleo-NorCPM1 wind nudging therefore enables reconstruction of ocean variability, including diagnostics such as the OHC and AMOC. The simulated OHC and AMOC variability compare broadly with a recent paleo-based ocean reconstruction

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

(Dalaiden et al., 2025) and with a long climate reanalysis constrained by instrumental SST observations (Wang et al., 2025), although differences remain. For most surface variables evaluated, paleo-NorCPM1 wind nudging shows similar performance to paleo-NorCPM1 offline, indicating that reconstructed atmospheric circulation and external forcing provide the dominant large-scale constraints on climate variability. By explicitly resolving ocean dynamics and atmosphere–ocean feedback, paleo-NorCPM1 wind nudging exhibits enhanced multi-decadal variability in near-surface air temperature and SST, reflecting the thermodynamic contribution of ocean adjustment. Notable differences between the two products emerge in the Southern Ocean and surrounding regions, particularly near the Antarctic Ice Sheet. In paleo-NorCPM1 wind nudging, strong multi-decadal subsurface variability associated with episodic deep convection leads to amplified surface variability. While such behavior reflects the dynamical sensitivity of the Southern Ocean to wind forcing, it also highlights known model sensitivities in representing deep convection (e.g., Zhou et al., 2026). Paleo-NorCPM1 wind nudging thus provides a dynamically consistent reconstruction of coupled variability, while emphasizing that ocean responses, especially in the Southern Ocean, remain sensitive to atmospheric circulation changes and model characteristics.

Taken together, the free transient ensemble, paleo-NorCPM1 offline, and paleo-NorCPM1 wind nudging provide a coherent framework (all products are based on the same NorESM1 configuration) to investigate the mechanisms underlying the long-term changes and low-frequency variability of the climate over the past centuries. By contrasting the ensemble mean of the free transient ensemble (representing the externally forced response) with paleo-NorCPM1, it becomes possible to assess the relative contributions of externally forced and internally generated variability to the reconstructed climate evolution. Additionally, paleo-NorCPM1 wind nudging further enables examination of how reconstructed atmospheric circulation drives ocean adjustment and coupled feedback processes, with the caveat that Southern Ocean variability exhibits strong sensitivity to wind forcing.

Data availability. The two reanalyses, paleo-NorCPM1 offline and the wind-nudged ensemble, are available on Zenodo at https://doi.org/10.5281/zenodo.19660743 (Dalaiden, 2026). Due to the size of the files, only the ensemble mean and variance from the two reanalyses are made available. All the members from the two reanalyses are available following this link: https://ns9039k.web.sigma2.no/datalake/dalaiden/paleo-NorCPM1/all_members.

Author contributions. QD and FC conceived the study, with input from LS and NK. QD carried out the NorESM1 experiments, with assistance from IB, and adapted NorCPM for the assimilation of paleoclimate records, with support from FC. All authors contributed to the discussion and interpretation of the results. QD prepared the original draft of the manuscript, and all authors contributed to reviewing and editing the final version.

Competing interests. The authors declare that they have no conflict of interest.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Acknowledgements. We gratefully acknowledge the paleoclimate community for making the proxy records used in this study publicly available. We also thank Sigma2 for providing computing and storage resources (projects ns11071k, nn11071k, ns9039k and nn9039k). QD acknowledges support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101149188.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Open Access Earth System Science Data Discussions

References

Abram, N. J., Henley, B. J., Sen Gupta, A., Lippmann, T. J. R., Clarke, H., Dowdy, A. J., Sharples, J. J., Nolan, R. H., Zhang, T., Wooster, M. J., Wurtzel, J. B., Meissner, K. J., Pitman, A. J., Ukkola, A. M., Murphy, B. P., Tapper, N. J., and Boer, M. M. (2021). Connections of climate change and variability to large and extreme forest fires in southeast Australia. *Communications Earth & Environment*, 2(1).

Allan, R. J. (2000). *ENSO and Climatic Variability in the Past 150 Years*, page 3–56. Cambridge University Press.

Anderson, J. L. (2001). An ensemble adjustment kalman filter for data assimilation. *Monthly Weather Review*, 129(12):2884 – 2903.

Anderson, J. L. and Collins, N. (2007). Scalable Implementations of Ensemble Filter Algorithms for Data Assimilation. *Journal of Atmospheric and Oceanic Technology*, 24(8):1452–1463.

Armour, K. C., Marshall, J., Scott, J. R., Donohoe, A., and Newsom, E. R. (2016). Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geoscience*, 9(7):549–554.

Badgeley, J. A., Steig, E. J., Hakim, G. J., and Fudge, T. J. (2020). Greenland temperature and precipitation over the last 20,000 years using data assimilation. *Climate of the Past*, 16:1335–1346.

Barthélémy, S., Counillon, F., and Wang, Y. (2024). Adaptive covariance hybridization for the assimilation of sst observations within a coupled earth system reanalysis. *Journal of Advances in Modeling Earth Systems*, 16(6):e2023MS003888. e2023MS003888 2023MS003888.

Beck, J. W., Edwards, R. L., Ito, E., Taylor, F. W., Recy, J., Rougerie, F., Joannot, P., and Henin, C. (1992). Sea-surface temperature from coral skeletal strontium/calcium ratios. *Science (New York, N.Y.)*, 257(5070):644–647.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E. (2013). The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. *Geoscientific Model Development*, 6(3):687–720.

Berger, A. (1978). Long-term variations of daily insolation and quaternary climatic changes. *Journal of Atmospheric Sciences*, 35(12):2362 – 2367.

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H., Svendsen, L., Chiu, P.-G., Passos, L., Bentsen, M., Guo, C., Gupta, A., Tjiputra, J., Kirkevåg, A., Olivié, D., Seland, Ø., Solsvik Vågane, J., Fan, Y., and Eldevik, T. (2021). NorCPM1 and its contribution to CMIP6 DCPP. *Geoscientific Model Development*, 14(11):7073–7116.

Box, J. E., Cressie, N., Bromwich, D. H., Jung, J. H., Van Den Broeke, M., Van Angelen, J. H., Forster, R. R., Miège, C., Mosley-Thompson, E., Vinther, B., and Mcconnell, J. R. (2013). Greenland ice sheet mass balance reconstruction. Part I: Net snow accumulation (1600-2009). *Journal of Climate*, 26(11):3919–3934.

Breitenmoser, P., Brönnimann, S., and Frank, D. (2014). Forward modelling of tree-ring width and comparison with a global network of tree-ring chronologies. *Climate of the Past*, pages 437–449.

Brennan, M. K. and Hakim, G. J. (2022). Reconstructing Arctic Sea Ice over the Common Era Using Data Assimilation. *Journal of Climate*, 35(4):1231–1247.

Brennan, M. K., Hakim, G. J., and Blanchard-Wrigglesworth, E. (2020). Arctic Sea-Ice Variability During the Instrumental Era. *Geophysical Research Letters*, 47(7):1–9.

Bromwich, D. H., Ensign, A., Wang, S. H., and Zou, X. (2024). Major Artifacts in ERA5 2-m Air Temperature Trends Over Antarctica Prior to and During the Modern Satellite Era. *Geophysical Research Letters*, 51(21).

Brönnimann, S., Luterbacher, J., Staehelin, J., Svendby, T. M., Hansen, G., and Svenoe, T. (2004). Extreme climate of the global troposphere and stratosphere in 1940-42 related to El Niño. *Nature*, 431(7011):971–974.

Buckley, M. W. and Marshall, J. (2016). Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review. *Reviews of Geophysics*, 54(1):5–63.

Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., and Saba, V. (2018). Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature*, 556(7700):191–196.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):1–50.

Comiso, J. C. and Gordon, A. L. (1987). Recurring polynyas over the cosmonaut sea and the maud rise. *Journal of Geophysical Research: Oceans*, 92(C3):2819–2833.

Conroy, J. L., Thompson, D. M., Cobb, K. M., Noone, D., Rea, S., and Legrande, A. N. (2017). Spatiotemporal variability in the δ^{18}o-salinity relationship of seawater across the tropical pacific ocean. *Paleoceanography*, 32(5):484–497.

Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, L., and Zheng, F. (2014). Seasonal-to-decadal predictions with the ensemble Kalman filter and the Norwegian Earth System Model: a twin experiment. *Tellus B: Chemical and Physical Meteorology*.

Counillon, F., Keenlyside, N., Bethke, I., Wang, Y., Billeau, S., Shen, M. L., and Bentsen, M. (2016). Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the norwegian climate prediction model. *Tellus B: Chemical and Physical Meteorology*.

Counillon, F., Sakov, P., and Bertino, L. (2009). Application of a hybrid EnKF-OI to ocean forecasting. *Ocean Science*, 5(4):389–401.

Dai, A., Fyfe, J. C., Xie, S.-P., and Dai, X. (2015). Decadal modulation of global surface temperature by internal climate variability. *Nature Climate Change*, 5(6):555–559.

Dalaiden, Q. (2026). Paleo-based climate reanalysis: paleo-NorCPM1. doi: https://doi.org/10.5281/zenodo.19660743.

Dalaiden, Q. and Bethke, I. (2026). Brief communication: Uncertainties in southern ocean sea surface conditions and their impact on antarctic climate over 1958–1978. *The Cryosphere*, 20(4):2089–2098.

Dalaiden, Q., Goosse, H., Holland, P. R., and Barthelemy, A. (2025). Dynamical reconstruction of Southern Ocean and Antarctic climate variability since 1700. *Scientific data*, 12(1):1574.

Dalaiden, Q., Rezsöhazy, J., Goosse, H., Thomas, E. R., Vladimirova, D. O., and Tetzner, D. (2023). An unprecedented sea ice retreat in the weddell sea driving an overall decrease of the antarctic sea-ice extent over the 20th century. *Geophysical Research Letters*, 50(21):e2023GL104666. e2023GL104666 2023GL104666.

Dalaiden, Q., Schurer, A. P., Kirchmeier-young, M. C., Goosse, H., and Hegerl, G. C. (2022). West Antarctic Surface Climate Changes Since the Mid-20th Century Driven by Anthropogenic Forcing Geophysical Research Letters. *Geophysical Research Letters*, 49(e2022GL099543).

Dee, S., Emile-Geay, J., Evans, M. N., Allam, A., Steig, E. J., and Thompson, D. (2015). Prysm: An open-source framework for proxy system modeling, with applications to oxygen-isotope systems. *Journal of Advances in Modeling Earth Systems*, 7(3):1220–1247.

Delhasse, A., Kittel, C., Amory, C., Hofer, S., van As, D., S. Fausto, R., and Fettweis, X. (2020). Brief communication: Evaluation of the near-surface climate in ERA5 over the Greenland Ice Sheet. *The Cryosphere*, 14(3):957–965.

DelSole, T., Tippett, M. K., and Shukla, J. (2011). A Significant Component of Unforced Multidecadal Variability in the Recent Acceleration of Global Warming. *Journal of Climate*, 24(3):909–926.

Deser, C., Alexander, M. A., Xie, S. P., and Phillips, A. S. (2010). *Sea surface temperature variability: Patterns and mechanisms*, volume 2.

Deser, C. and Phillips, A. S. (2021). Defining the internal component of atlantic multidecadal variability in a changing climate. *Geophysical Research Letters*, 48(22):e2021GL095023. e2021GL095023 2021GL095023.

Deser, C., Phillips, A. S., Simpson, I. R., Rosenbloom, N., Coleman, D., Lehner, F., Pendergrass, A. G., Dinezio, P., and Stevenson, S. (2020). Isolating the evolving contributions of anthropogenic aerosols and greenhouse gases: A new CESM1 large ensemble community resource. *Journal of Climate*, 33(18):7835–7858.

Emile-Geay, J., McKay, N. P., Kaufman, D. S., von Gunten, L., Wang, J., Anchukaitis, K. J., Abram, N. J., Addison, J. A., Curran, M. A. J., Evans, M. N., Henley, B. J., Hao, Z., Martrat, B., McGregor, H. V., Neukom, R., Pederson, G. T., Stenni, B., Thirumalai, K., Werner, J. P., Xu, C., Divine, D. V., Dixon, B. C., Gergis, J., Mundo, I. A., Nakatsuka, T., Phipps, S. J., Routson, C. C., Steig, E. J., Tierney, J. E., Tyler, J. J., Allen, K. J., Bertler, N. A. N., Björklund, J., Chase, B. M., Chen, M.-T., Cook, E., de Jong, R., DeLong, K. L., Dixon, D. A., Ekaykin, A. A., Ersek, V., Filipsson, H. L., Francus, P., Freund, M. B., Frezzotti, M., Gaire, N. P., Gajewski, K., Ge, Q., Goosse, H., Gornostaeva, A., Grosjean, M., Horiuchi, K., Hormes, A., Husum, K., Isaksson, E., Kandasamy, S., Kawamura, K., Kilbourne, K. H., Koç, N., Leduc, G., Linderholm, H. W., Lorrey, A. M., Mikhalenko, V., Mortyn, P. G., Motoyama, H., Moy, A. D., Mulvaney, R., Munz, P. M., Nash, D. J., Oerter, H., Opel, T., Orsi, A. J., Ovchinnikov, D. V., Porter, T. J., Roop, H. A., Saenger, C., Sano, M., Sauchyn, D., Saunders, K. M., Seidenkrantz, M.-S., Severi, M., Shao, X., Sicre, M.-A., Sigl, M., Sinclair, K., St. George, S., St. Jacques, J.-M., Thamban, M., Kuwar Thapa, U., Thomas, E. R., Turney, C., Uemura, R., Viau, A. E., Vladimirova, D. O., Wahl, E. R., White, J. W. C., Yu, Z., and Zinke, J. (2017). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific data*, 4(July):170088.

England, M. H., McGregor, S., Spence, P., Meehl, G. A., Timmermann, A., Cai, W., Gupta, A. S., McPhaden, M. J., Purich, A., and Santoso, A. (2014). Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change*, 4(3):222–227.

Epstein, S., , and Mayeda, T. (1953). Variation of o18 content of waters from natural sources. *Geochimica et Cosmochimica Acta*, 4(5):213–224.

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J. (2013). Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary Science Reviews*, 76:16–28.

Evensen, G. (2003). The Ensemble Kalman Filter : theoretical formulation and practical implementation. pages 343–367.

Fischer, H., Meissner, K. J., Mix, A. C., Abram, N. J., Austermann, J., Brovkin, V., Capron, E., Colombaroli, D., Daniau, A. L., Dyez, K. A., Felis, T., Finkelstein, S. A., Jaccard, S. L., McClymont, E. L., Rovere, A., Sutter, J., Wolff, E. W., Affolter, S., Bakker, P., Ballesteros-Cánovas, J. A., Barbante, C., Caley, T., Carlson, A. E., Churakova, O., Cortese, G., Cumming, B. F., Davis, B. A., De Vernal, A., Emile-Geay, J., Fritz, S. C., Gierz, P., Gottschalk, J., Holloway, M. D., Joos, F., Kucera, M., Loutre, M. F., Lunt, D. J., Marcisz, K., Marlon, J. R., Martinez, P., Masson-Delmotte, V., Nehrbass-Ahles, C., Otto-Bliesner, B. L., Raible, C. C., Risebrobakken, B., Sánchez Goñi, M. F., Arrigo, J. S., Sarnthein, M., Sjolte, J., Stocker, T. F., Velasquez Alvárez, P. A., Tinner, W., Valdes, P. J., Vogel, H., Wanner, H., Yan, Q., Yu, Z., Ziegler, M., and Zhou, L. (2018). Palaeoclimate constraints on the impact of 2 °c anthropogenic warming and beyond. *Nature Geoscience*, 11(7):474–485.

Fox-Kemper, B., Hewitt, H., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slangen, A. B. A., and Yu, Y. (2021). Ocean, cryosphere and sea level change. Ocean, Cryosphere and Sea Level Change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Franke, J., Valler, V., Brönnimann, S., Neukom, R., and Jaume-Santero, F. (2020). The importance of input data quality and quantity in climate field reconstructions – results from the assimilation of various tree-ring collections. *Climate of the Past*, 16(3):1061–1074.

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., and Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, 28(2):862–886.

Gao, C., Robock, A., and Ammann, C. (2008). Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models. *Journal of Geophysical Research: Atmospheres*, 113(D23).

Garcia-Oliva, L., Counillon, F., Bethke, I., and Keenlyside, N. (2024). Intercomparison of initialization methods for seasonal-to-decadal climate predictions with the NorCPM. *Climate Dynamics*, 62(6):5425–5444.

Gaspari, G. and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757.

Gilbert, A. L., Kay, J. E., Blanchard-Wrigglesworth, E., Bailey, D. A., Holland, M. M., Jahn, A., and Schneider, D. P. (2025). Observed winds alone cannot explain recent arctic warming and sea ice loss. *Environmental Research: Climate*, 4(4):045009.

Good, S. A., Martin, M. J., and Rayner, N. A. (2013). En4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, 118(12):6704–6716.

Goosse, H., Kay, J. E., Armour, K. C., Bodas-Salcedo, A., Chepfer, H., Docquier, D., Jonko, A., Kushner, P. J., Lecomte, O., Massonnet, F., Park, H. S., Pithan, F., Svensson, G., and Vancoppenolle, M. (2018). Quantifying climate feedbacks in polar regions. *Nature Communications*, 9(1).

Goosse, H., Lefebvre, W., de Montety, A., Crespin, E., and Orsi, A. H. (2009). Consistent past half-century trends in the atmosphere, the sea ice and the ocean at high southern latitudes. *Climate Dynamics*, 33(7-8):999–1016.

Gossart, A., Helsen, S., Lenaerts, J. T. M., Broucke, S. V., van Lipzig, N. P. M., and Souverijns, N. (2019). An Evaluation of Surface Climatology in State-of-the-Art Reanalyses over the Antarctic Ice Sheet. *Journal of Climate*, 32(20):6899–6915.

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N., and Perkins, W. A. (2016). The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research*, 121(12):6745–6764.

Hegerl, G. and Zwiers, F. (2011). Use of models in detection and attribution of climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 2(4):570–591.

Hegerl, G. C., Brönnimann, S., Schurer, A., and Cowan, T. (2018). The early 20th century warming: Anomalies, causes, and consequences. *Wiley interdisciplinary reviews. Climate change*, 9(4):e522.

Hegglin, M., Kinnison, D., Lamarque, J.-F., and Plummer, D. (2016). CCMI ozone in support of CMIP6 - version 1.0.

Henley, B. J. (2017). Pacific decadal climate variability: Indices, patterns and tropical-extratropical interactions. *Global and Planetary Change*, 155(May):42–55.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., and Folland, C. K. (2015). A Tripole Index for the Interdecadal Pacific Oscillation. *Climate Dynamics*, 45(11):3077–3090.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Hobbs, W. R., Massom, R., Stammerjohn, S., Reid, P., Williams, G., and Meier, W. (2016). A review of recent changes in Southern Ocean sea ice, their drivers and forcings. *Global and Planetary Change*, 143:228–250.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Holland, M. M., Bailey, D. A., Briegleb, B. P., Light, B., and Hunke, E. (2012). Improved sea ice shortwave radiation physics in ccsm4: The impact of melt ponds and aerosols on arctic sea ice. *Journal of Climate*, 25(5):1413 – 1430.

Hu, W., Yan, M., Xing, F., and Lei, L. (2025). A new last two millennium reanalysis based on hybrid gain analog of fl ine EnKF and an expanded proxy database. *npj Climate and Atmospheric Science*.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M. (2017). Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate*, 30(20):8179–8205.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J. F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., and Marshall, S. (2013). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, 94(9):1339–1360.

Hébert, R. and Laepple, T. (2025). Lacking oceanic-driven internal multidecadal climate variability is compensated by forced variability in model simulations.

Jouzel, J., Alley, R. B., Cuffey, K. M., Dansgaard, W., Grootes, P., Hoffmann, G., Johnsen, S. J., Koster, R. D., Peel, D., Shuman, C. A., Stievenard, M., Stuiver, M., and White, J. (1997). Validity of the temperature reconstruction from ice cores. *Journal of Geophysical Research*, 102(C12):26471–26487.

Kalnay, E. (2003). *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press.

Kennedy, J. J., Rayner, N. A., Atkinson, C. P., and Killick, R. E. (2019). An ensemble data set of sea surface temperature change from 1850: The met office hadley centre hadsst.4.0.0.0 data set. *Journal of Geophysical Research: Atmospheres*, 124(14):7719–7763.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E., Struthers, H., Ekman, A. M. L., Ghan, S., Griesfeller, J., Nilsson, E. D., and Schulz, M. (2013). Aerosol–climate interactions in the norwegian earth system model – noresm1-m. *Geoscientific Model Development*, 6(1):207–244.

Konecky, B. L., Mckay, N. P., Sidorova, O. V. C., Comas-bru, L., Dassié, E. P., Delong, K. L., Falster, G. M., Fischer, M. J., and Jones, M. D. (2020). The Iso2k database : a global compilation of paleo- δ 18 O and δ 2 H records to aid understanding of Common Era climate. *Earth System Science Data*, pages 2261–2288.

Kosaka, Y. and Xie, S.-P. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, 501(7467):403–407.

Kruse, C. G., Bacmeister, J. T., Zarzycki, C. M., Larson, V. E., and Thayer-Calder, K. (2022). Do nudging tendencies depend on the nudging timescale chosen in atmospheric models? *Journal of Advances in Modeling Earth Systems*, 14(10):e2022MS003024. e2022MS003024 2022MS003024.

Laepple, T., Ziegler, E., Weitzel, N., Hébert, R., Ellerhoff, B., Schoch, P., Martrat, B., Bothe, O., Moreno-Chamarro, E., Chevalier, M., Herbert, A., and Rehfeld, K. (2023). Regional but not global temperature variability underestimated by climate models at supradecadal timescales. *Nature Geoscience*, 16(11):958–966.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Schmied, H. M., Hersbach, H., Buontempo, C., and Cagnazzo, C. (2021). Wfde5 over land merged with era5 over the ocean (w5e5 v2.0).

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G. (2011). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, 3(1).

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Lehner, F. and Deser, C. (2023). Origin, importance, and predictive limits of internal climate variability. *Environmental Research: Climate*, 2(2):023001.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences*, 105(36):13252–13257.

Marshall, G. J. (2003). Trends in the Southern Annular Mode from Observations and Reanalyses. *Journal of Climate*, 16(1999):4134–4143.

Matsikaris, A., Widmann, M., and Jungclaus, J. (2015). On-line and off-line data assimilation in palaeoclimatology: A case study. *Climate of the Past*, 11(1):81–93.

Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok de Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretzschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., van de Kamp, M., Verronen, P. T., and Versick, S. (2017). Solar forcing for cmip6 (v3.2). *Geoscientific Model Development*, 10(6):2247–2302.

McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., Baringer, M. O., Meinen, C. S., Collins, J., and Bryden, H. L. (2015). Measuring the Atlantic Meridional Overturning Circulation at 26°N. *Progress in Oceanography*, 130:91–111.

Medley, B., McConnell, J. R., Neumann, T. A., Reijmer, C. H., Chellman, N., Sigl, M., and Kipfstuhl, S. (2018). Temperature and Snowfall in Western Queen Maud Land Increasing Faster Than Climate Model Projections. *Geophysical Research Letters*, 45(3):1472–1480.

Medley, B. and Thomas, E. R. (2018). Increased snowfall over the Antarctic Ice Sheet mitigated 20th century sea level rise. *Nature Clim. Change*, in press(2).

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (cmip6). *Geoscientific Model Development*, 10(5):2057–2116.

Miyoshi, T., Kondo, K., and Imamura, T. (2014). The 10,240-member ensemble Kalman filtering with an intermediate AGCM. pages 5264–5271.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R. (2021). An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. *Journal of Geophysical Research: Atmospheres*, 126(3):e2019JD032361.

Nash, J. and Sutcliffe, J. (1970). River flow forecasting through conceptual models part i — a discussion of principles. *Journal of Hydrology*, 10(3):282 – 290.

Naughten, K. A., Meissner, K. J., Galton-Fenzi, B. K., England, M. H., Timmermann, R., Hellmer, H. H., Hattermann, T., and Debernard, J. B. (2018). Intercomparison of Antarctic ice-shelf, ocean, and sea-ice interactions simulated by MetROMS-iceshelf and FESOM 1.4. *Geoscientific Model Development*, 11(4):1257–1292.

Neale, R., Richter, J., Conley, A., Park, S., Lauritzen, P., Gettelman, A., Williamson, D., Rasch, P., Vavrus, S., Taylor, M., Collins, W., Zhang, M., and Lin, S.-J. (2010). Description of the community atmosphere model (cam 4.0). *NCAR Technical Note*, TN-485+STR.

Okazaki, A., Miyoshi, T., Yoshimura, K., Greybush, S. J., and Zhang, F. (2021). Revisiting online and offline data assimilation comparison for paleoclimate reconstruction: An idealized osse study. *Journal of Geophysical Research: Atmospheres*, 126(16):e2020JD034214. e2020JD034214 2020JD034214.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Osman, M. B., Tierney, J. E., Zhu, J., Tardif, R., Hakim, G. J., King, J., and Poulsen, C. J. (2021). Globally resolved surface temperatures since the Last Glacial Maximum. *Nature*, 599(7884):239–244.

Perkins, W. A. and Hakim, G. J. (2017). Reconstructing paleoclimate fields using online data assimilation with a linear inverse model. *Climate of the Past*, 13(5):421–436.

Rayner, N. A., Parker, E. B., Horton, C. K., Holland, 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. *Journal of Geophysical Research*, 108(D14):4407.

Revell, L. E., Stenke, A., Luo, B., Kremser, S., Rozanov, E., Sukhodolov, T., and Peter, T. (2017). Impacts of mt pinatubo volcanic aerosol on the tropical stratosphere in chemistry–climate model simulations using ccmi and cmip6 stratospheric aerosol data. *Atmospheric Chemistry and Physics*, 17(21):13139–13150.

Rezsöhazy, J., Dalaiden, Q., Klein, F., Goosse, H., and Guiot, J. (2022). Using a process-based dendroclimatic proxy system model in a data assimilation framework: a test case in the southern hemisphere over the past centuries. *Climate of the Past*, 18(9):2093–2115.

Riser, S. C., Freeland, H. J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., Gilbert, D., Xu, J., Pouliquen, S., Thresher, A., Le Traon, P.-Y., Maze, G., Klein, B., Ravichandran, M., Grant, F., Poulain, P.-M., Suga, T., Lim, B., Sterl, A., Sutton, P., Mork, K.-A., Vélez-Belchí, P. J., Ansorge, I., King, B., Turton, J., Baringer, M., and Jayne, S. R. (2016). Fifteen years of ocean observations with the global Argo array. *Nature Climate Change*, 6(2):145–153.

Roach, L. A., Blanchard-Wrigglesworth, E., Ragen, S., Cheng, W., Armour, K. C., and Bitz, C. M. (2022). The Impact of Winds on AMOC in a Fully-Coupled Climate Model. *Geophysical Research Letters*, 49(24):e2022GL101203.

Rosenblum, E. and Eisenman, I. (2017). Sea Ice Trends in Climate Models Only Accurate in Runs with Biased Global Warming. *Journal of Climate*, 30(16):6265–6278.

Sakov, P. and Oke, P. R. (2008). A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A*, 60(2):361–371.

Sakov, P., Oliver, D. S., and Bertino, L. (2012). An Iterative EnKF for Strongly Nonlinear Systems. *Monthly Weather Review*, 140(6):1988–2004.

Sallée, J.-B., Morrison, A. K., Naughten, K., and Thompson, A. F. (2025). Southern ocean circulation: Global drivers and ongoing changes. In Meredith, M. P., Melbourne-Thomas, J., Naveira Garabato, A. C., and Raphael, M., editors, *Antarctica and the Earth System*. Routledge.

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., Delaygue, G., Joos, F., Krivova, N. A., Muscheler, R., Otto-Bliesner, B. L., Pongratz, J., Shindell, D. T., Solanki, S. K., Steinhilber, F., and Vieira, L. E. A. (2012). Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geoscientific Model Development*, 5(1):185–191.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Statist.*, 6(2):461–464.

Sheffield, J., Goteti, G., and Wood, E. F. (2006). Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19(12):3088–3111.

Skiba, V., Ackermann, L., Hirsch, N., Sun, Y., Dolman, A. M., Lohmann, G., and Laepple, T. (2026). Enhancing climate variability in earth system models through paleo-informed nudging of sea surface temperatures. *Environmental Research: Climate*, 5(1):15025.

Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D. C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood,

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

K., and Wyszyński, P. (2021). An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3. *Journal of Climate*, 34(4):1417–1438.

Steiger, N. and Hakim, G. (2016). Multi-timescale data assimilation for atmosphere-ocean state estimates. *Climate of the Past*, 12(6):1375–1388.

Steiger, N. J., Smerdon, J. E., Cook, E. R., and Cook, B. I. (2018). A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Scientific Data*, 5:1–15.

Steiger, N. J., Steig, E. J., Dee, S. G., Roe, G. H., and Hakim, G. J. (2017). Climate reconstruction using data assimilation of water isotope ratios from ice cores. *Journal of Geophysical Research: Atmospheres*, 122(3):1545–1568.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1):79–107.

Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay, J., Anderson, D. M., Steig, E. J., and Noone, D. (2019). Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. *Climate of the Past*, 15(4):1251–1273.

Thomas, E. R., Melchior Van Wessem, J., Roberts, J., Isaksson, E., Schlosser, E., Fudge, T. J., Vallelonga, P., Medley, B., Lenaerts, J., Bertler, N., Van Den Broeke, M. R., Dixon, D. A., Frezzotti, M., Stenni, B., Curran, M., and Ekaykin, A. A. (2017). Regional Antarctic snow accumulation over the past 1000 years. *Climate of the Past*, 13(11):1491–1513.

Thomas, E. R., Vladimirova, D. O., Tetzner, D. R., Emanuelsson, B. D., Chellman, N., Dixon, D. A., Goosse, H., Grieman, M. M., King, A. C. F., Sigl, M., Udy, D., Vance, T. R., Winski, D. A., Winton, V. H. L., Bertler, N. A. N., Hori, A., Laluraj, C. M., McConnell, J. R., Motizuki, Y., Takahashi, K., Motoyama, H., Nakai, Y., Schwanck, F., Simões, J. C., Gaudie Ley Lindau, F., Severi, M., Traversi, R., Wauthy, S., Xiao, C., Yang, J., Mosely-Thompson, E., Khodzher, T., Golobokova, L., and Ekaykin, A. (2022). Ice core chemistry database: an antarctic compilation of sodium and sulphate records spanning the past 2000 years. *Earth System Science Data Discussions*, 2022:1–20.

Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T. (2018). A global space-based stratospheric aerosol climatology: 1979–2016. *Earth System Science Data*, 10(1):469–492.

Thompson, D. M., Ault, T. R., Evans, M. N., Cole, J. E., and Emile-Geay, J. (2011). Comparison of observed and simulated tropical climate trends using a forward model of coral δ^{18}o. *Geophysical Research Letters*, 38(14).

Thorne, P. W., Nicklas, J. M., Kennedy, J. J., Calvert, B., Fox-kemper, B., Richardson, M. T., Hawkins, E., Rohde, R., Cowtan, K., Abram, N. J., Andersson, A., and Noone, S. (2026). *How well can we quantify when 1 . 5 ° C of global warming has been exceeded ?* Number January.

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, Ø., and Heinze, C. (2013). Evaluation of the carbon cycle components in the norwegian earth system model (noresm). *Geoscientific Model Development*, 6(2):301–325.

Trenberth, K. E. and Solomon, A. (1994). The global heat balance: heat transports in the atmosphere and ocean. *Climate Dynamics*, 10(3):107–134.

Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., Lagun, V., Reid, P. A., and Iagovkina, S. (2004). The SCAR READER project: Toward a high-quality database of mean Antarctic meteorological observations. *Journal of Climate*, 17(14):2890–2898.

Valler, V., Franke, J., Brugnara, Y., Samakinwa, E., Hand, R., Lundstad, E., Burgdorf, A. M., Lipfert, L., Friedman, A. R., and Brönnimann, S. (2024). ModE-RA: a global monthly paleo-reanalysis of the modern era 1421 to 2008. *Scientific Data*, 11(1):1–19.

https://doi.org/10.5194/essd-2026-433
Preprint. Discussion started: 17 August 2026

Vieira, L. E. A., Solanki, S. K., Krivova, N. A., and Usoskin, I. (2011). Evolution of the solar irradiance during the Holocene. 6:1–20.

Walter, R. M., Sayani, H. R., Felis, T., Cobb, K. M., Abram, N. J., Arzey, A. K., Atwood, A. R., Brenner, L. D., Dassié, É. P., Delong, K. L., Ellis, B., Emile-Geay, J., Fischer, M. J., Goodkin, N. F., Hargreaves, J. A., Kilbourne, K. H., Krawczyk, H., Mckay, N. P., Moore, A. L., Murty, S. A., Ong, M. R., Ramos, R. D., Reed, E. V., Samanta, D., Sanchez, S. C., and Zinke, J. (2023). The CoralHydro2k database: a global, actively curated compilation of coral δ18O and Sr/Ca proxy records of tropical ocean hydrology and temperature for the Common Era. *Earth System Science Data*, 15(5):2081–2116.

Wang, Y., Counillon, F., Bethke, I., Keenlyside, N., Bocquet, M., and lin Shen, M. (2017). Optimising assimilation of hydrographic profiles into isopycnal ocean models with ensemble data assimilation. *Ocean Modelling*, 114:33–44.

Wang, Y., Counillon, F., Keenlyside, N., Svendsen, L., Gleixner, S., Kimmritz, M., Dai, P., and Gao, Y. (2019). Seasonal predictions initialised by assimilating sea surface temperature observations with the EnKF. *Climate Dynamics*, 53(9):5777–5797.

Wang, Y., Counillon, F., Svendsen, L., Chiu, P.-G., Keenlyside, N., Laloyaux, P., Koseki, M., and de Boisseson, E. (2025). An ensemble-based coupled reanalysis of the climate from 1860 to the present (corea1860+). *Earth System Science Data*, 17(8):4185–4211.

Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T. (2020). CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation. *Geophysical Research Letters*, 47(12):e2019GL086075.

Whitaker, J. S. and Hamill, T. M. (2002). Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, 130(7):1913–1924.

Yasunaka, S. and Hanawa, K. (2011). Intercomparison of historical sea surface temperature datasets. *International Journal of Climatology*, 31(7):1056–1073.

Yeager, S. and Danabasoglu, G. (2014). The Origins of Late-Twentieth-Century Variations in the Large-Scale North Atlantic Circulation. *Journal of Climate*, 27(9):3222–3247.

Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y. O., Marsh, R., Yeager, S. G., Amrhein, D. E., and Little, C. M. (2019). A Review of the Role of the Atlantic Meridional Overturning Circulation in Atlantic Multidecadal Variability and Associated Climate Impacts. *Reviews of Geophysics*, 57(2):316–375.

Zhou, L., Ayres, H., Gülk, B., Narayanan, A., de Lavergne, C., Ödalen, M., Silvano, A., Wang, X., Lindeman, M., and Steiger, N. (2026). Review article: Weddell Sea Polynya formation, cessation and climatic impacts. *The Cryosphere*, 20(1):285–308.