

Coupled Seasonal Data Assimilation of Sea Ice, Ocean, and Atmospheric Dynamics over the Last Millennium

ZILU MENG¹,^a GREGORY J. HAKIM,^a AND ERIC J. STEIG^{b,a}

^a *Department of Atmospheric and Climate Science, University of Washington, Seattle, Washington*

^b *Department of Earth and Space Sciences, University of Washington, Seattle, Washington*

(Manuscript received 23 January 2025, in final form 1 July 2025, accepted 28 August 2025)

ABSTRACT: “Online” data assimilation (DA) is used to generate a seasonal-resolution reanalysis dataset over the last millennium by combining forecasts from an ocean–atmosphere–sea ice coupled linear inverse model with climate proxy records. Instrumental verification reveals that this reconstruction achieves the highest correlation skill while using fewer proxies in surface temperature reconstructions compared to other paleo-DA products, particularly during boreal winter when proxy data are scarce. Reconstructed ocean and sea ice variables also have high correlation with instrumental and satellite datasets. Verification against independent proxy records shows that reconstruction skill is robust throughout the last millennium. Analysis of the results reveals that the method effectively captures the seasonal evolution and amplitude of El Niño events, seasonal temperature trends that are consistent with orbital forcing over the last millennium, and polar-amplified cooling in the transition from the medieval climate anomaly to the little ice age.

SIGNIFICANCE STATEMENT: This paper introduces a new seasonal-resolution reanalysis of the last millennium, based on an “online” data assimilation method using a linear inverse model to assimilate paleoclimate proxies. We find good agreement when verifying the reconstruction against modern instrumental reanalyses and out-of-sample proxies. Results show that seasonal temperature trends are similar to predictions from orbital-insolation trends, and seasonal variability of modern El Niño events is similar to instrumental reanalyses. This framework offers a dynamically consistent, seasonally resolved view of past climate variability that supports broader applications in paleoclimate research.

KEYWORDS: ENSO; Climate; Climate variability; Paleoclimate; Data assimilation; Multidecadal variability

1. Introduction

Reconstructions of past climate are essential for understanding the dynamics of the long-term climate system. Such reconstructions are particularly important in the context of global warming (Pachauri et al. 2007), as they place contemporary climate variability within a larger sample of past climate. This historical and long-term perspective also enhances our ability to improve projections of future climate change by providing a reference against which model simulations can be compared. Before the instrumental era, when humans began using scientific tools to record weather and climate information like temperature and precipitation, climate information is primarily derived from natural proxies such as tree rings, corals, and ice cores. For example, the width of some tree rings reflects local moisture and temperature stress (e.g., Briffa et al. 2004). This information can be used to reconstruct past climate conditions. The challenges of using proxies to reconstruct past climate include their uneven spatial distribution, inconsistent temporal coverage, varying temporal resolution, complex relationships with physical drivers, and complex

noise characteristics (e.g., Christiansen and Ljungqvist 2017), complicating multiproxy interpretations of climate variability. Recently, data assimilation (DA) methods (e.g., Bouttier and Courtier 2002) have been increasingly used to reconstruct past climates (e.g., Dirren and Hakim 2005; Goose et al. 2010; Widmann et al. 2010; Franke et al. 2017; Perkins and Hakim 2021; Tardif et al. 2019; Steiger et al. 2018; Valler et al. 2024). This approach combines climate model physical constraints with proxy data to reconstruct climate variables. One of the most significant advantages of DA is that it allows for the reconstruction of variables not directly represented by the proxies (Hakim et al. 2016). For example, we can use temperature data from proxies to infer sea ice conditions and geopotential height, as there are strong and well-understood correlations among these variables (e.g., Hakim et al. 2016; Steiger et al. 2018; Brennan 2022; Brennan and Hakim 2022; Meng and Hakim 2024).

A general and flexible paleoclimate DA (PDA) framework, the Last Millennium Reanalysis (LMR), was proposed by Hakim et al. (2016) for reconstructing climate variables over the Common Era. The success of this framework has been followed by extensive research on PDA (e.g., Steiger et al. 2017; Dee et al. 2020; Sun et al. 2022; Luo et al. 2022; Meng and Hakim 2024; Zhu et al. 2023; Okazaki et al. 2021; Hu et al. 2024). One practical limitation of PDA compared to weather DA is the high cost of forecasts that generate the prior (“first guess”) because of the need for long integrations of climate models (e.g., Taylor et al. 2012). Consequently, the initial

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-25-0048.s1>.

Corresponding author: Zilu Meng, zilumeng@uw.edu

LMR framework used an “offline” DA method, where the prior is sampled from a static source, such as existing climate model simulations. This approach is particularly relevant in paleoclimate applications, where the forecast skill of dynamical climate models may be limited and the computational cost of running those models is high.

There are, however, patterns of variability, such as the Pacific decadal oscillation (PDO) (e.g., Mantua and Hare 2002) and El Niño–Southern Oscillation (ENSO) (e.g., McPhaden et al. 2006; Meng and Li 2024), that persist on seasonal to interannual time scales. Developing “online” PDA methods that exploit this persistence can lead to more accurate reconstructions since the memory of past proxies is transmitted into the future by the forecast model. Using DA with a skillful coupled atmosphere–ocean model, information from terrestrial proxies such as tree-ring width can be used to inform ocean state estimates, which then carry memory in the model forecast through ocean persistence. An example of this online PDA approach is shown by Perkins and Hakim (2021), who used a linear inverse model (LIM) to reconstruct climate fields over the last millennium and found improved representations of decadal variability.

A significant challenge with PDA reconstructions is resolving the seasonal cycle. For example, proxies from Northern Hemisphere trees, including tree-ring width (TRW) and late-wood density, primarily reflect warm-season temperature (PAGES2k Consortium 2013, 2017). Previous PDA studies have used this information to reconstruct annual-mean climate variability, leading to biases in the reconstructions and inconsistent results for significant climate periods such as the medieval climate anomaly and the little ice age (PAGES2k Consortium 2017; Hakim et al. 2016; Tardif et al. 2019; Steiger et al. 2018). Here, we present results for a seasonal reconstruction of the last millennium using online PDA. We use a LIM to forecast one season to the next: from March–May (MAM) to June–August (JJA), from JJA to September–November (SON), from SON to December–February (DJF), and from DJF to the next year.

The LIM incorporates sea ice variables (concentration and thickness), recognizing the long-lead memory of sea ice (Blanchard-Wrigglesworth et al. 2011) and therefore predictive skill, especially near the Arctic where seasonal variability is large. We assimilate proxies from the PAGES2k V2 (PAGES2k Consortium 2017) database at the season specific to each proxy. Proxies that represent annual-mean conditions are assimilated subsequently once an annual-mean forecast is available from the initial LIM forecast from seasonal proxies only.

The organization of the remainder of the paper is as follows. Section 2 details the PDA methods and data used in this study, and section 3 presents instrumental and proxy verification to measure the accuracy of the reconstruction. Section 4 applies the reconstruction to analyze seasonal climate variability and trends over the last millennium. Section 5 provides a concluding discussion.

2. LMR-seasonal framework data and methods

The LMR-seasonal approach utilizes an online “cycling” DA framework, consisting of three components, with the goal

TABLE 1. Variables reconstructed in LMR seasonal and their spatial coverage.

Variable	Description	Spatial coverage
TAS	2-m air temperature	Global ($2^\circ \times 2^\circ$)
TOS	Sea surface temperature	Global ($2^\circ \times 2^\circ$)
OHC300	Ocean heat content (0–300 m)	Global ($2^\circ \times 2^\circ$)
SIT	Sea ice thickness (NH)	NH ($2^\circ \times 2^\circ$)
SIC	Sea ice concentration (NH)	NH ($2^\circ \times 2^\circ$)

of reconstructing a subset of atmosphere–ocean–sea ice fields (Table 1). The first component involves a simplified climate model, a LIM trained for seasonal forecasting as described in section 2a. The second component, proxy system models, which estimate the proxies from the prior, are trained as described in section 2b. The third component is an ensemble Kalman filter (EnKF), which is used to combine the proxy and prior as described in section 2c.

a. LIM

Linear inverse models are a computationally efficient, widely applied, and skillful method for predicting climate fields (e.g., Penland and Magorian 1993; Penland and Sardeshmukh 1995; Newman 2013; Perkins and Hakim 2020; Meng and Hakim 2024). A LIM captures linear dynamics of anomalies about a mean state:

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \boldsymbol{\xi}, \quad (1)$$

where $\mathbf{x}$ is the state vector. The $\mathbf{L}$ is a matrix representing deterministic dynamics, and $\boldsymbol{\xi}$ is a random noise vector, which is temporally uncorrelated but may have correlations in the state space $\mathbf{x}$. For a stable linear system, the real part of the eigenvalues of $\mathbf{L}$ is negative (Penland and Magorian 1993).

In this study, $\mathbf{x}$ represents low-dimensional principal components (PCs) derived from a truncated set of empirical orthogonal functions (EOFs) (e.g., Meng et al. 2023). EOF truncation is applied to individual variables of interest, including 2-m air temperature (TAS), sea surface temperature (TOS), ocean heat content from 300 m to the surface (OHC300), Northern Hemisphere (NH) sea ice thickness (SIT), and NH sea ice concentration (SIC) using SACPY (Meng et al. 2023; Meng and Li 2024). We do not include Southern Hemisphere (SH) sea ice due to the sparseness of SH paleoclimate proxies and known challenges in reconciling climate model simulations of SH sea ice with observations (Roach et al. 2020). This selection of variables is guided by two primary considerations: 1) We limit the number of PCs within the LIM to prevent overfitting, which could degrade the quality of reconstructions; and 2) we exclude high-frequency atmospheric variables such as sea level pressure (SLP) to avoid reducing the forecast skill of the primary variables of interest. We include both TAS and TOS in the LIM state vector to ensure that important variability over both land (TAS) and ocean (TOS) regions is captured, as demonstrated in Fig. S1 in the online supplemental material. This choice is necessary because we assimilate temperature-sensitive proxies located over both land and ocean. For all variables except

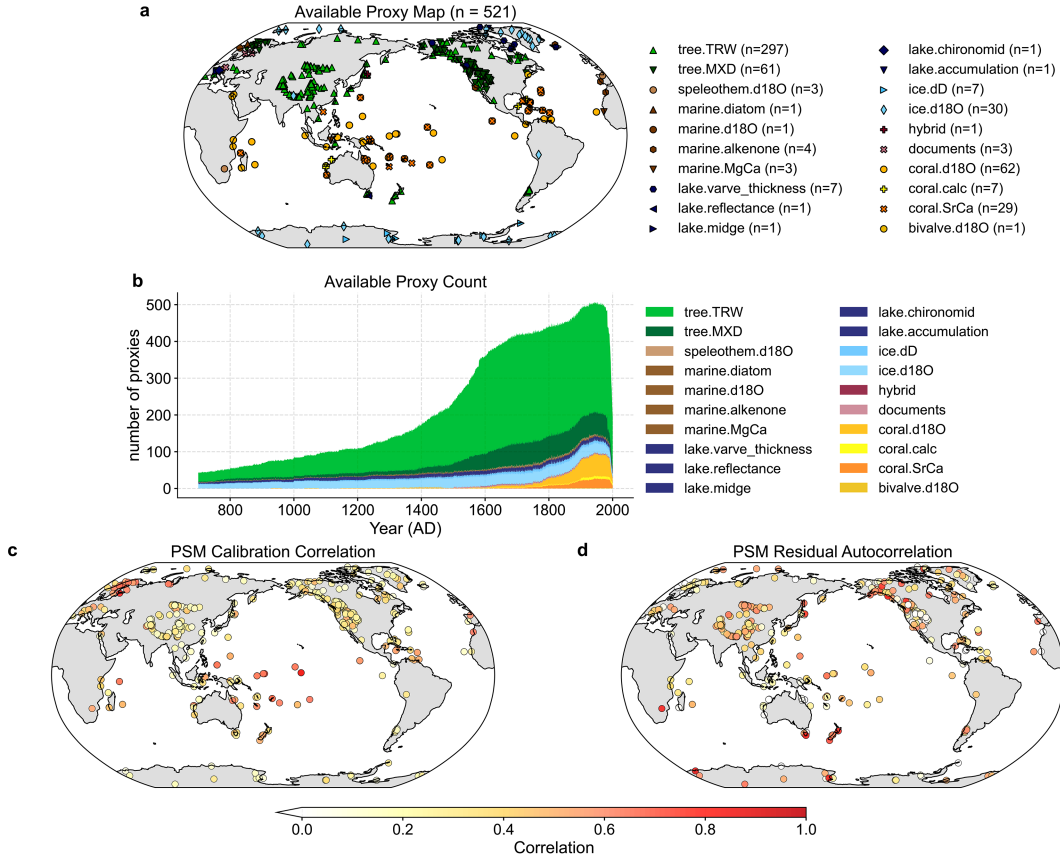


FIG. 1. Proxies from PAGES2k V2 (PAGES2k Consortium 2017). (a) Locations and counts of proxy types after filtering by the specified standards indicated by section 2b. (b) Evolution of the number of proxies over time. (c),(d) Spatial distribution of PSM calibration correlations and 1-yr lag residual (error) autocorrelations.

OHC300, we select the first 15 PCs, which account for around 80% of the total variance of each variable. Following Perkins and Hakim (2020), we select 30 PCs for OHC300 to better capture the extended memory of this variable within the LIM. Thus, the state vector is defined as

$$\mathbf{x} = [\mathbf{PC}_{\text{TAS}}^T, \mathbf{PC}_{\text{TOS}}^T, \mathbf{PC}_{\text{OHC300}}^T, \mathbf{PC}_{\text{SIT}}^T, \mathbf{PC}_{\text{SIC}}^T]^T. \quad (2)$$

LINEAR INVERSE MODEL TRAINING PROCESS

We utilize output from two models in the Coupled Model Intercomparison Project phase 5 (CMIP5) Last Millennium experiments, specifically CCSM4 and MPI-ESM-R (Taylor et al. 2012). We choose these models primarily to maintain consistency with LMR v1 (Hakim et al. 2016), LMR v2 (Tardif et al. 2019), and LMR online (Perkins and Hakim 2021). Furthermore, seasonal climate variability statistics have not changed significantly from CMIP5 to CMIP6 (Brown et al. 2020). The LIM is trained separately on data from CCSM4 and MPI-ESM-R to capture the internal covariability and physical relationships specific to each parent model. We define four seasons by 3-month averages: MAM, JJA, SON, and DJF. Prior to taking the seasonal average, model output data are placed on a $2^\circ \times 2^\circ$ latitude–longitude grid using linear interpolation in the Climate Data Operators package

(Schulzweida et al. 2019), and the last millennium trend for each month is removed by simple linear regression.

EOF analysis on area-weighted variables yields the first 15 PCs for each variable (30 PCs for OHC300). Then, $\mathbf{L}$ is calculated by

$$\mathbf{L} = \tau^{-1} \ln \mathbf{C}(\tau) \mathbf{C}(0)^{-1}. \quad (3)$$

Here, $\mathbf{C}(\tau)$ is the τ -lag covariance matrix of $\mathbf{x}$, $\mathbf{C}(\tau) = \langle \mathbf{x}(\tau) \mathbf{x}^T(0) \rangle$, where “ $\langle \rangle$ ” represents a sample average. Here, τ is 3 months for our seasonal LIM. The stochastic part of the dynamics ξ has covariance matrix $\mathbf{Q}$, such that $\langle \xi \xi^T \rangle = \mathbf{Q}$. The matrix $\mathbf{Q}$ is calculated based on stationary statistics:

$$\frac{d\mathbf{C}(0)}{dt} = \mathbf{L}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{L}^T + \mathbf{Q} = 0. \quad (4)$$

Using $\mathbf{Q}$ and $\mathbf{L}$, stochastic integration of (1) yields a sample trajectory using the two-step integration process of Penland and Matrosova (1994):

$$\mathbf{x}_{t+\delta t} = (\mathbf{L}\delta t + \mathbf{I})\mathbf{x}_t + \hat{\mathbf{Q}}\sqrt{\Delta t}\alpha, \quad (5)$$

$$\mathbf{x}_{t+\delta t/2} = \frac{1}{2}(\mathbf{x}_{t+\delta t} + \mathbf{x}_t), \quad (6)$$

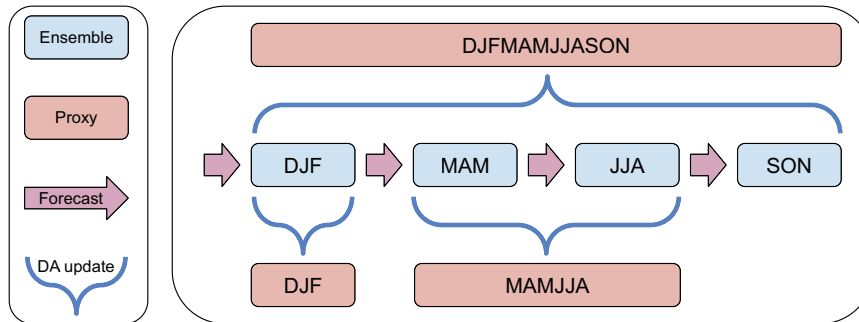


FIG. 2. LMR seasonal update strategy. The light-blue box represents the ensemble, the rose box represents the proxy, and the pink arrow represents the forecast step from the LIM. Curly brackets denote the update from the EnKF to integrate the proxy data into updating the prior ensemble. The text within the box indicates the seasonality of either the ensemble or the proxies.

where δt is the integration time step, set at 6 h for this study; $\mathbf{I}$ is the identity matrix, $\hat{\mathbf{Q}}$ denotes the matrix where columns are eigenvectors of $\mathbf{Q}$, and $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues of $\mathbf{Q}$; $\boldsymbol{\alpha}$ is a vector of independent standard normal random variables. We exclude eigenvectors associated with negative eigenvalues in $\mathbf{Q}$ and normalize the remaining eigenvalues to preserve total variance, following Penland and Matrosova (1994) and Perkins and Hakim (2020). A complete derivation of the LIM is provided in the supplemental appendix. After training the LIM on output from CCSM4 and MPI-ESM-R Last Millennium simulations (850–1850 CE) (Taylor et al. 2012), the LIM demonstrates predictive skill to at least 12 months as evidenced by out-of-sample tests shown in Figs. S2 and S3.

b. PSMs

We use the temperature-sensitive PAGES2k V2 dataset (PAGES2k Consortium 2017) as the observational inputs for our data assimilation process. The PAGES2k V2 dataset comprises approximately 700 proxy records, mainly from tree rings, corals, and ice cores. We calibrate proxy system models (PSMs) for each proxy record, using only surface temperatures from GISTEMP v4 (Lenssen et al. 2019) for terrestrial proxies and sea surface temperature (SST) from ERSST v5 (Huang et al. 2017) for marine records. Precipitation and other moisture variables are not included in the PSM calibration. Comparisons with calibration on other instrumental observations,

such as Berkeley Earth (Rohde and Hausfather 2020) and MLOST (Smith et al. 2008), yield similar results (not shown).

The truncated EOF basis of the LIM does not fully resolve local details of climate fields such as surface temperature and SST, which is a “representativeness” error we account for by calibrating the PSMs in the EOF-truncated space, $\hat{\mathbf{x}} = \mathbf{U}\mathbf{U}^T\mathbf{x}$, where $\mathbf{U}^T$ is the matrix with the first 15 PCs derived from EOF analysis:

$$\mathbf{y} = \mathbf{H}\hat{\mathbf{x}} + \boldsymbol{\epsilon}. \quad (7)$$

Here, $\mathbf{H}$ is the matrix that maps the climate variables to the proxy data, $\mathbf{y}$, and $\boldsymbol{\epsilon}$ is the error term.

As will be discussed in section 2c, an important factor for data assimilation is the observation error covariance matrix $\mathbf{R} = \langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle$. Hakim et al. (2022) employ linear regression to estimate $\mathbf{H}$ and a full covariance matrix $\mathbf{R}$. In our study, we adopt a similar method but with a diagonal $\mathbf{R}$, implying zero error covariance among errors for the proxy PSMs. The use of a full $\mathbf{R}$ matrix is impractical because the calibration period does not provide a sufficiently long overlap between each of the proxies to estimate the off-diagonal terms, and tests suggest that the diagonal values of $\mathbf{R}$ are several orders of magnitude larger than the off-diagonal elements (not shown).

SEASONALITY

Seasonality refers to the specific season that a proxy’s temperature represents. Tardif et al. (2019) assessed both expert-

TABLE 2. Summary of DA products over the Last Millennium.

Name	PDA method	N_{proxy}^a	Time resolution	Reference
PHYDA	Offline	2978	(Sub)Annual ^b	Steiger et al. (2018)
LMR v2	Offline	2250	Annual	Tardif et al. (2019)
LMR online	Online	545	Annual	Perkins and Hakim (2021)
LMR seasonal	Online	521	Seasonal ^c	This study

^a N_{proxy} denotes the total number of proxies for data assimilation.

^b“(Sub)Annual” refers to PHYDA’s time resolutions finer than a year because it has the annual mean, DJF, and JJA reconstructions, and monthly Niño-3.4 index reconstructions.

^c“Seasonal” refers to 4 time steps (MAM, JJA, SON, and DJF) every year.

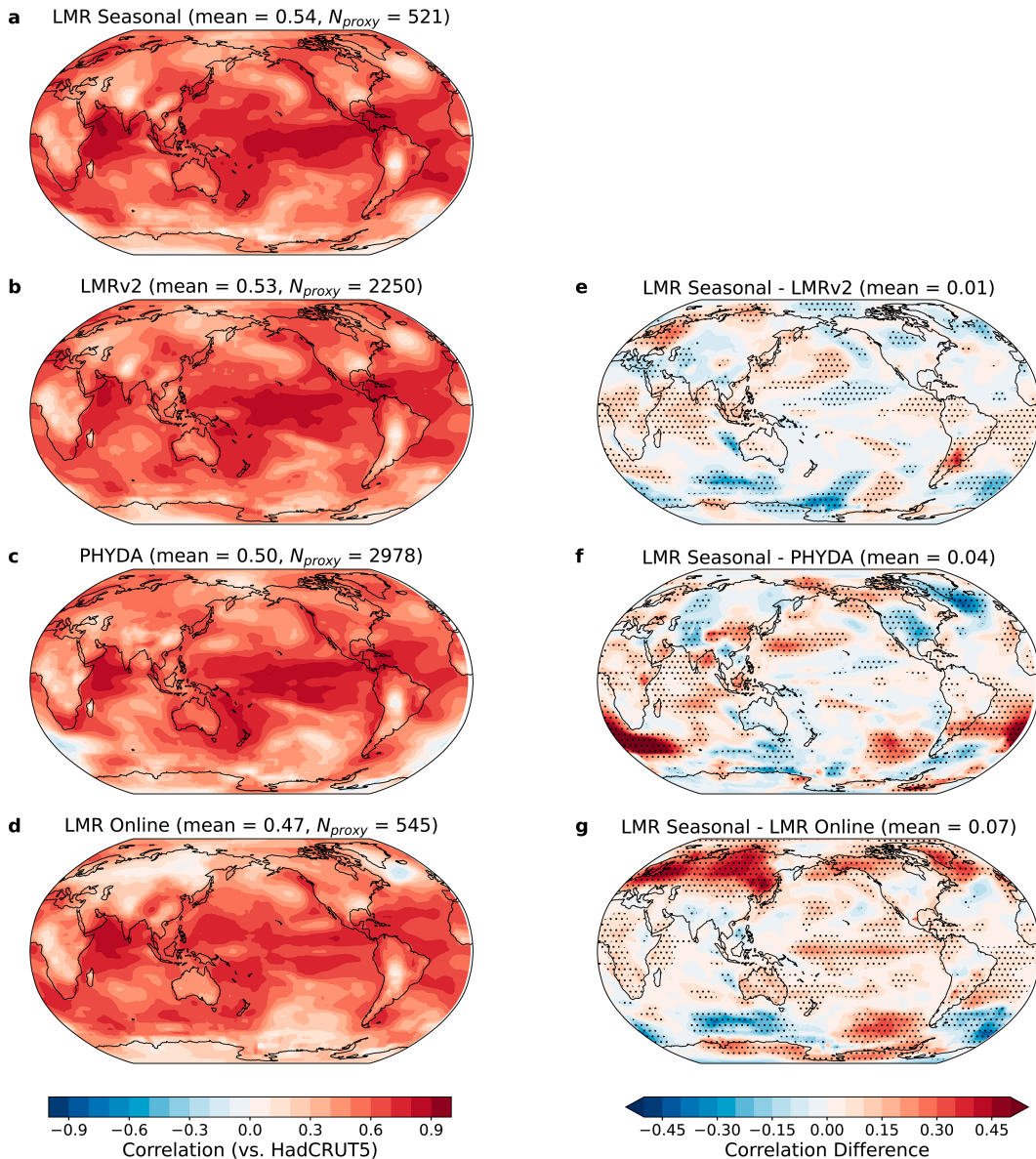


FIG. 3. Annual-mean surface temperature instrumental verification. (a)–(d) Correlation between various DA reconstruction's ensemble mean and HadCRUT5 (Morice et al. 2021) 2-m air temperature during 1880–2000. Results are shown for (a) LMR seasonal, (b) LMRv2, (c) PHYDA, and (d) LMR online, with the global-mean correlation and the number of used proxies given in the title for each subpanel. Correlation differences between (a) LMR seasonal and other reconstructions are shown for (e) LMRv2, (f) PHYDA, and (g) LMR online, with global-mean correlation differences indicated in the titles. Black dots show the correlation difference values passing the confidence level of 95% using Williams test (Williams 1959).

based seasonality, derived from PAGES2k metadata, and objectively determined seasonality, given by the best correlation with instrumental data during PSM calibration. We evaluate both definitions of seasonality and find no significant differences in calibration results (cf. Fig. 1 with Fig. S4), or in the PDA results as measured by instrumental verification (cf. Fig. 3 with Fig. S5; and Fig. 4 with Fig. S6) and independent proxies (cf. Fig. 12 with Fig. S15; see section 3). We therefore use the objectively determined seasonality to maintain consistency with LMR

v2 (Tardif et al. 2019) and LMR online (Perkins and Hakim 2021). We note that subseasonal coral records are averaged to seasonal resolution (DJF, MAM, JJA, and SON) for PSM calibration and assimilation.

We remove proxies from assimilation if they have an insignificant correlation with local temperature or high temporal error autocorrelation. Specifically, we remove proxies that have a PSM calibration correlation below 0.05 or a 1-yr lag autocorrelation in PSM calibration residuals exceeding 0.90.

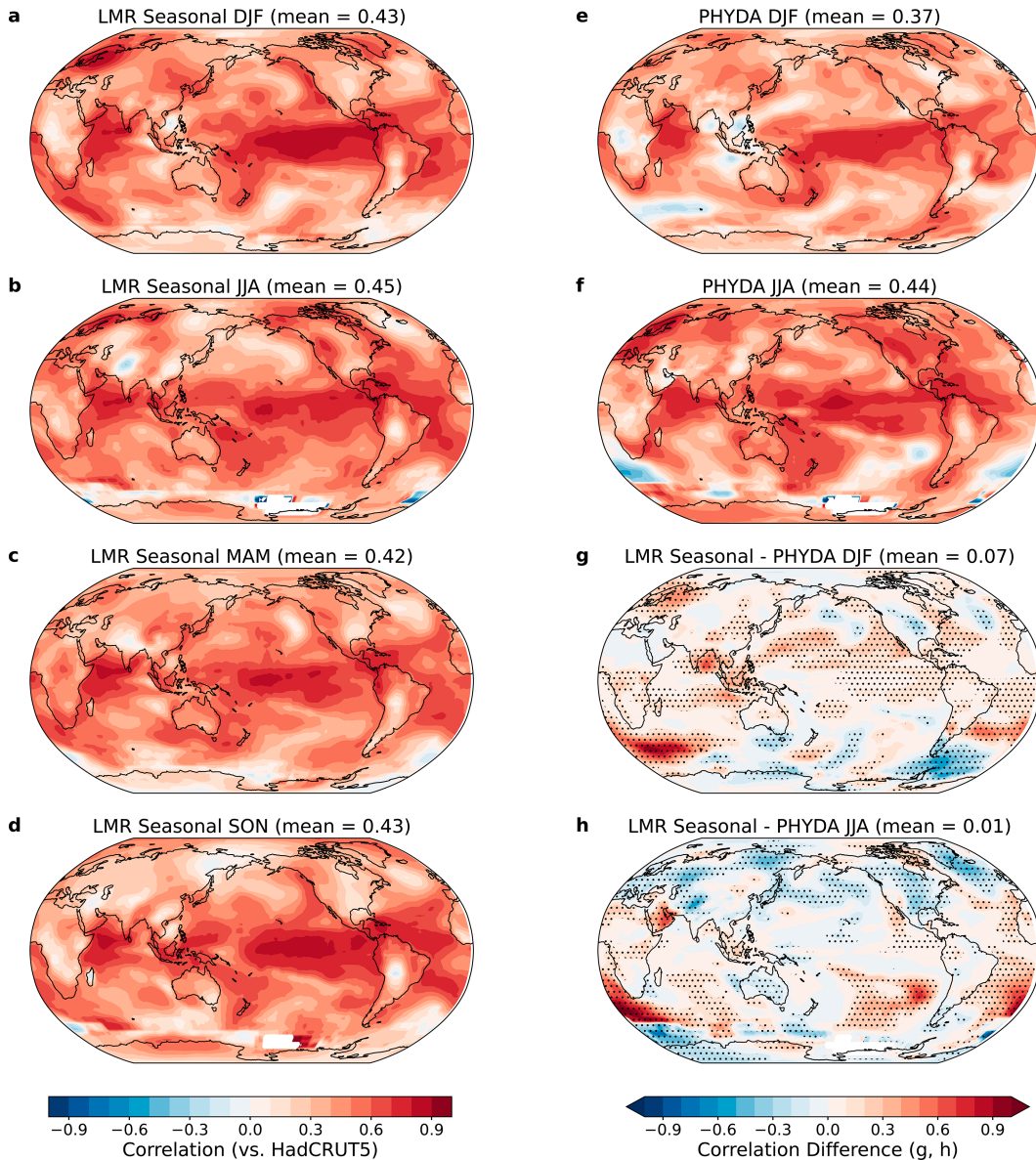


FIG. 4. (a)–(d) Surface temperature seasonal instrumental verification. The correlations between the LMR seasonal ensemble mean and HadCRUT5 (Morice et al. 2021). For (a) DJF, (c) MAM, (b) JJA, and (d) SON during 1880–2000. (e),(f). Correlations between PHYDA and HadCRUT5 are shown for (e) DJF and (f) JJA. (g),(h) The correlation differences between LMR seasonal and PHYDA in (g) DJF and (h) JJA. Black dots show the correlation difference values passing the confidence level of 95% using Williams test (Williams 1959).

As shown in Fig. 1, some Pacific corals have high PSM calibration correlation but also a 1-yr-lag error autocorrelation. High error autocorrelation is problematic for Kalman filters, which assume that observation errors are uncorrelated in time.

c. Ensemble Kalman filter and update strategy

Here, we introduce the update strategy for our seasonal reconstruction, which is based on an ensemble implementation of the Kalman filter (e.g., Evensen 2003).

1) ENSEMBLE KALMAN FILTER

The EnKF is extensively applied in a variety of paleo-DA tasks and has consistently shown strong performance (Hakim et al. 2016; Franke et al. 2017; Perkins and Hakim 2021; Steiger et al. 2018; Zhu et al. 2022; Valler et al. 2024). The primary step in the EnKF process is the “update”:

$$\mathbf{x}_a = \mathbf{x}_p + \mathbf{K}[\mathbf{y} - \mathcal{H}(\mathbf{x}_p)], \quad (8)$$

where $\mathbf{x}_a$ represents the posterior (“analysis”) state vector, $\mathbf{x}_p$ denotes the prior state vector, and $\mathcal{H}$ is the observation

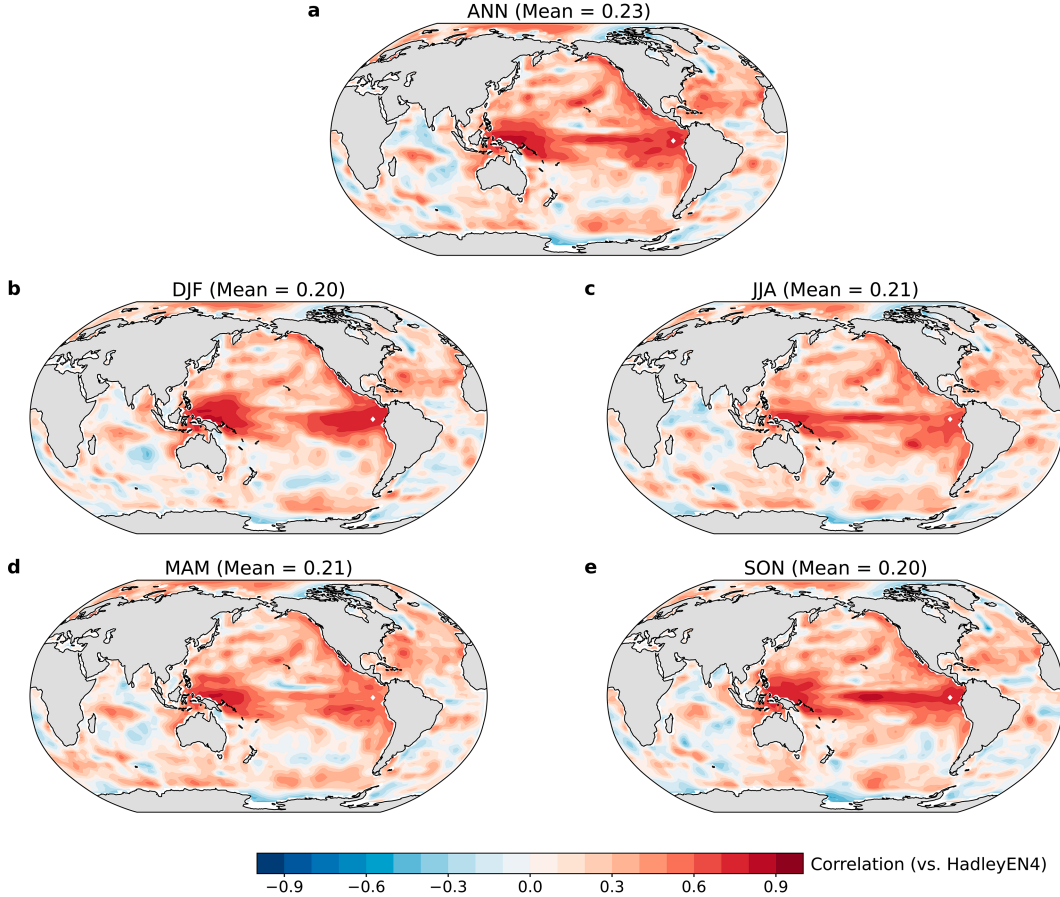


FIG. 5. Ocean heat content from 300 m to the surface (OHC300) instrumental verification. Correlation between the LMR seasonal OHC300 ensemble mean and HadleyEN4 OHC300 (Good et al. 2013) over the period 1940–2000 for (a) the annual mean, (b) DJF, (c) JJA, (d) MAM, and (e) SON. Global-mean correlations are indicated in the titles.

operator that maps to the corresponding observation vector (i.e., the PSMs). Matrix $\mathbf{K}$, the Kalman gain, is defined by

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T[\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}]^{-1}, \quad (9)$$

where $\mathbf{B}$ is the prior covariance matrix and $\mathbf{H}$ is the linearization of $\mathcal{H}$. The matrix $\mathbf{R}$ is the observation error covariance matrix derived from Eq. (7). Given that all PSMs in this study are linear, $\mathcal{H} = \mathbf{H}$. To solve (8) and (9) using ensemble sampling, we employ the ensemble square root filter (EnSRF) method (Whitaker and Hamill 2002) incorporating a serial observation update strategy. For the k th proxy, whose value is y_k , the update proceeds by separating the ensemble into the ensemble mean ($\bar{\mathbf{x}}$) and perturbations ($\mathbf{x}'_i$):

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{x}'_i. \quad (10)$$

For the ensemble mean $\bar{\mathbf{x}}$, the update equation is

$$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_p + \frac{\text{cov}(\mathbf{x}_p, y_{e,k})}{\text{var}(y_{e,k}) + R_k}(y_k - \bar{y}_{e,k}), \quad (11)$$

where $y_{e,k}$ denotes the k th proxy estimate from the ensemble, represented as $y_{e,k} = \bar{y}_{e,k} + y'_{e,k}$, and R_k is the k th proxy error

variance. The “var” and “cov” operators denote the variance and covariance, respectively. Ensemble perturbations $\mathbf{x}'_i$ are update by

$$\mathbf{x}'_a = \mathbf{x}'_p - \left[1 + \sqrt{\frac{R_k}{\text{var}(y_{e,k}) + R_k}} \right]^{-1} \frac{\text{cov}(\mathbf{x}_p, y_{e,k})}{\text{var}(y_{e,k}) + R_k}(y'_{e,k}). \quad (12)$$

This ensemble update is completed for the i th member using (10) to obtain the full analysis state. In this study, the ensemble size is 800, which allows us to avoid ensemble inflation and localization methods (e.g., Anderson 2012). Localization techniques are complicated by the EOF state space of the LIM, so we use an ensemble large enough to minimize the need for such localization.

2) SEASONAL UPDATE STRATEGY

Unlike previous PDA reconstructions (e.g., Hakim et al. 2016; Steiger et al. 2018; Tardif et al. 2019; Perkins and Hakim 2021) that use seasonal proxies to update the annual mean, our approach updates specific seasons corresponding to the proxy seasonality. An example illustration of this update strategy is shown

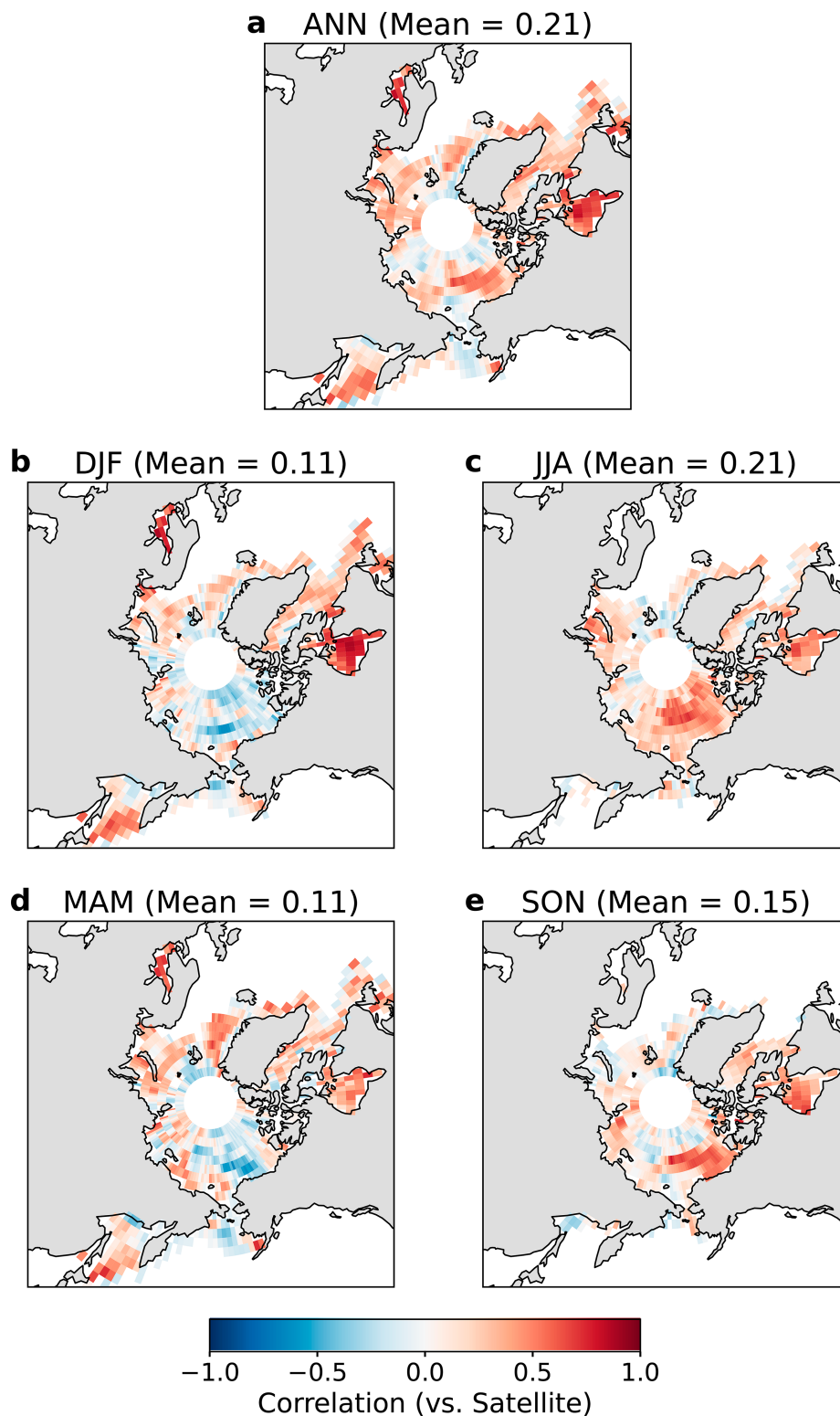


FIG. 6. NH SIC instrumental verification. Correlation between the LMR seasonal SIC ensemble mean and satellite SIC data (Fetterer et al. 2017) during 1980–2000 is presented for the (a) annual mean, (b) DJF, (c) JJA, (d) MAM, and (e) SON. Global-mean correlations are indicated in the titles.

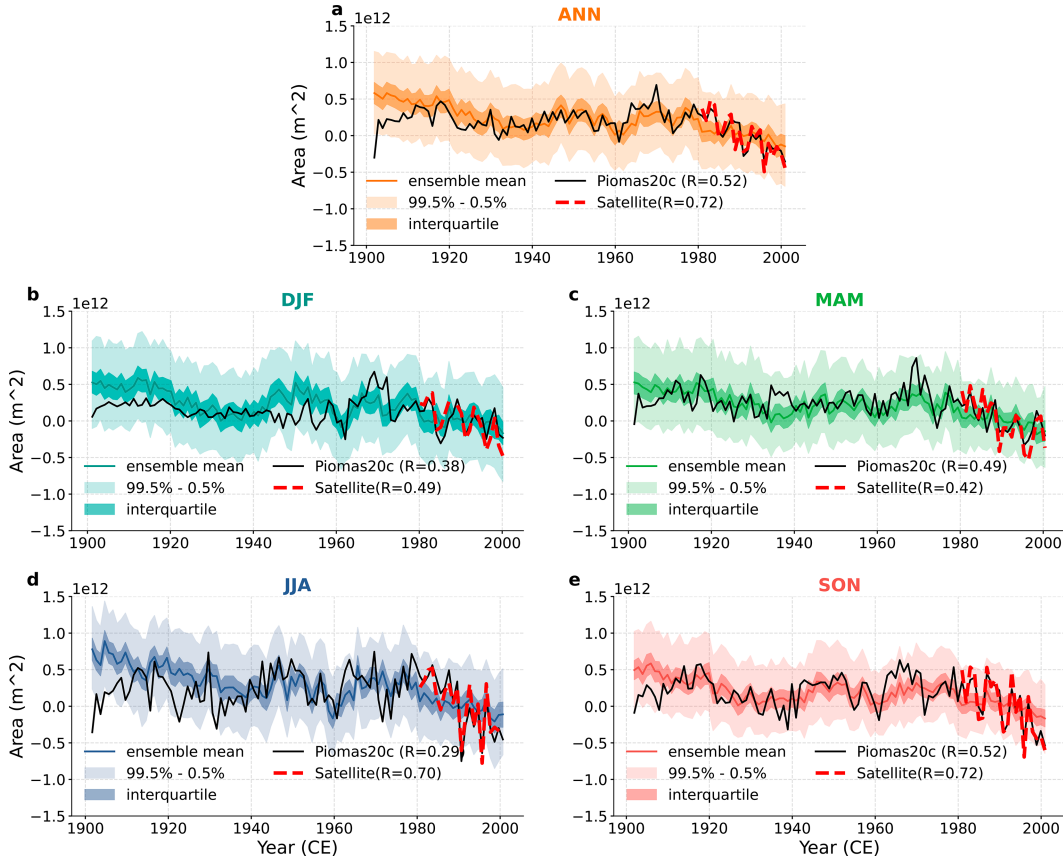


FIG. 7. NH SIA verification. Temporal comparison of the LMR seasonal reconstructed NH SIA series (colored curves) against Piomas20c (Schweiger et al. 2019) SIA (black solid curve), and satellite SIA (Fetterer et al. 2017) (red dashed curve) in the (a) annual mean, (b) DJF, (c) MAM, (d) JJA, and (e) SON. The reference time period for anomalies is 1980–2000. Dark shading denotes the ensemble interquartile range, and light shading denotes the 0.5%–99.5% interval.

in Fig. 2 for three proxies having different seasonality: DJF, MAMJJA, and DJFMAMJJASON. When the LIM forecast completes the DJF season, the DJF proxy is used to update the DJF prior ensemble. Subsequently, the LIM advances by updating the state to the MAM and JJA seasons. Upon reaching JJA, the MAMJJA proxy is used to update the MAM and JJA ensembles. Finally, when the LIM ensemble progresses to the SON season, the DJFMAMJJASON proxy is used to update the DJF, MAM, JJA, and SON ensembles. In summary, our methodology emphasizes a season-to-season update mechanism. This novel update strategy has a significant impact on reconstructions of the differences between the medieval climate anomaly (MCA) and little ice age (LIA), as discussed in section 4.

d. Verification metrics and significance test

We validate the LMR seasonal reconstruction against both calibration and reanalysis datasets (discussed below) using two primary verification metrics: correlation,

$$\text{corr} = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})(v_i - \bar{v})}{\sigma_x \sigma_v}, \quad (13)$$

and the coefficient of efficiency (CE) (Nash and Sutcliffe 1970),

$$\text{CE} = 1 - \frac{\sum_{i=1}^n (v_i - x_i)^2}{\sum_{i=1}^n (v_i - \bar{v})^2}. \quad (14)$$

Here, an overbar ($\bar{x}$) represents a mean value, σ represents the standard deviation, n is the number of temporal samples, and x and v represent the reconstructed and verification values, respectively. Correlation measures errors in signal timing, whereas CE measures errors in signal timing and amplitude.

To assess whether the correlation skill of LMR seasonal is significantly different from that of other datasets, we employ the Williams test (Williams 1959). This test is specifically designed to compare two correlation coefficients that are statistically dependent because they share a common variable, and it accounts for this dependency using the correlation between the two predictors. Specifically, let X represents our reconstruction, Y represents another PDA product, and Z represents the instrumental verification dataset. The goal is to test whether the correlation between X and Z (r_{xz}) is significantly different from that between

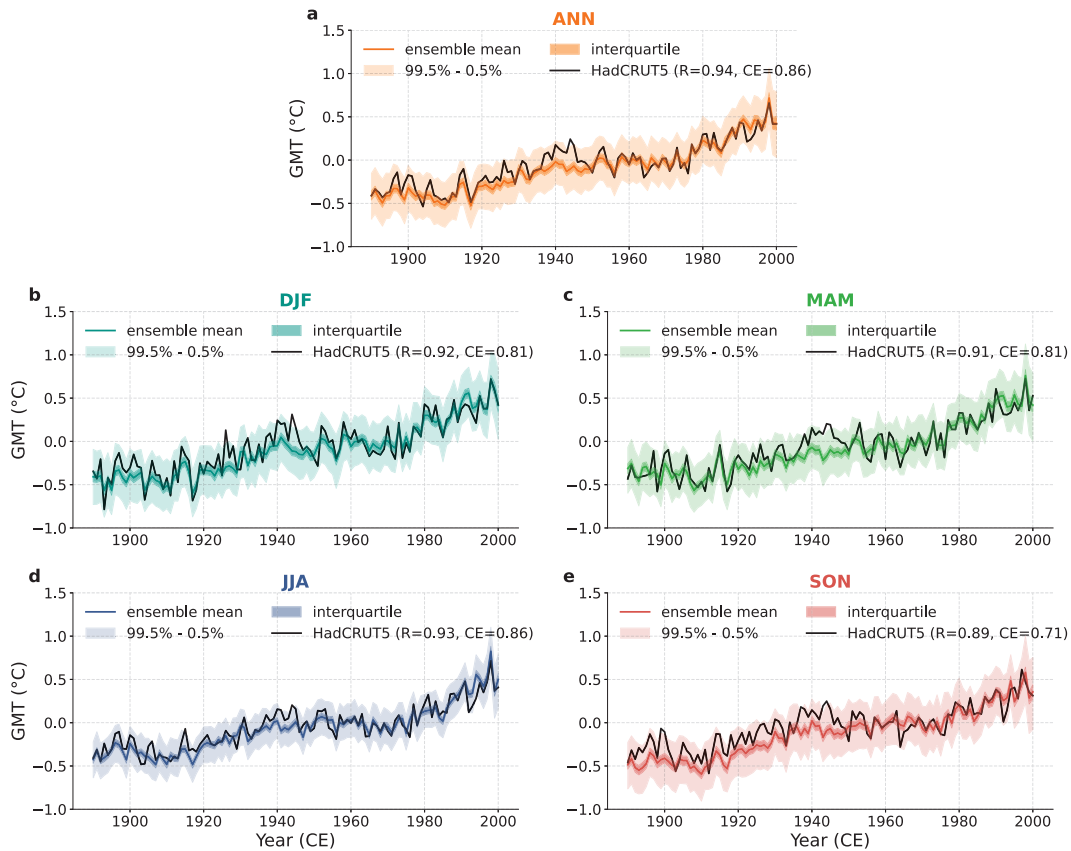


FIG. 8. GMT instrumental verification. Temporal verification of the ensemble mean LMR seasonal reconstructed GMT series (colored curves) against HadCRUT5 (Morice et al. 2021) GMT (black solid curve) in (a) annual mean, (b) DJF, (c) MAM, (d) JJA, and (e) SON. The reference time period for anomalies is 1950–80. For each reconstruction, dark shading denotes the ensemble interquartile range, and light shading denotes the 0.5%–99.5% interval. R = correlation.

Y and Z (r_{yz}), accounting for the correlation between X and Y (r_{xy}). The Williams test statistic is calculated as

$$t = (r_{xz} - r_{yz}) / \sqrt{\frac{(1 - r_{xz}^2 - r_{yz}^2 - r_{xy}^2 + 2r_{xz}r_{yz}r_{xy})(1 + r_{xy})}{2(n - 3)}}, \quad (15)$$

where n is the sample size. Under the null hypothesis that $r_{xz} = r_{yz}$, the test statistic approximately follows a t distribution with $(n - 3)$ degrees of freedom.

e. Comparisons to other PDA last millennium reconstructions

To assess agreement with other PDA products, we compare our results to three DA products over the last millennium: PHYDA (Steiger et al. 2018), LMR v2 (Tardif et al. 2019), and LMR online (Perkins and Hakim 2021). Details on the DA methods, proxy number, and time resolution for each product are given in Table 2. Besides LMR seasonal, only PHYDA provides reconstructions for DJF and JJA (but not for SON or MAM), while LMR v2 and LMR

online are limited to annual means. A significant distinction between the online and offline DA methods involves whether the prior is derived from random (time independent) draws from an existing climate model simulation (offline) or from a forecast of the analysis at the previous assimilation time (online). An advantage of the online DA method is that the “memory” of past proxy information is carried to the next assimilation time. This feature is particularly vital for our seasonal DA approach, as NH trees are the dominant source of the climate signal and are primarily sensitive to summer growing conditions. With online DA, the JJA posterior, for example, serves as the initial condition for the SON prior, informing the SON and future season’s reconstructions. This results in more information persisting into seasons with fewer proxies (e.g., winter and spring).

3. Verification

We verify our reconstruction using instrumental observations and proxy data, as described below. For instrumental verification, we use 2-m air temperature from HadCRUT5

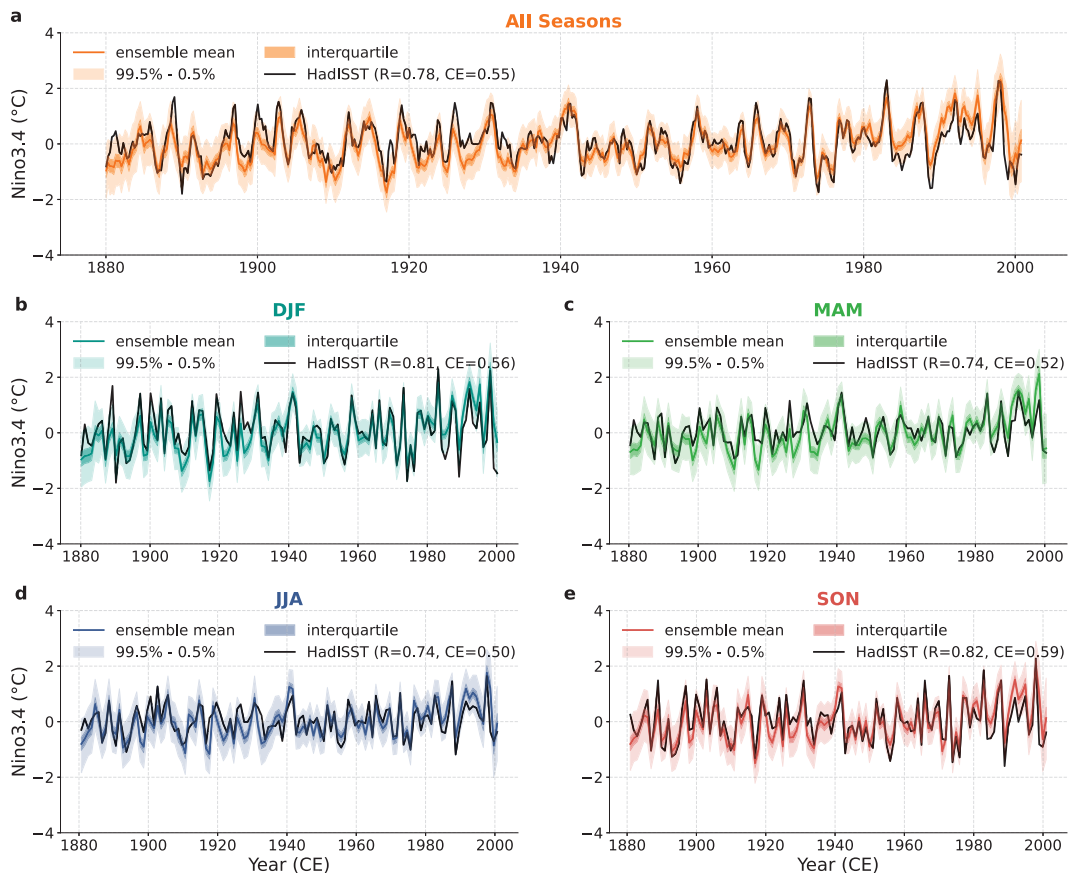


FIG. 9. Niño-3.4 index instrumental verification. Temporal verification of the ensemble mean LMR seasonal reconstructed Niño-3.4 index (colored curves) against HadISST (Rayner et al. 2003) (black solid curve) in (a) all seasons, (b) DJF, (c) MAM, (d) JJA, and (e) SON. Dark shading denotes the interquartile range, and light shading denotes the 0.5%–99.5% interval.

(Morice et al. 2021) and the ERA-20C reanalysis (Poli et al. 2016), ocean temperature data from Hadley EN4 (Good et al. 2013), SST from HadISST (Rayner et al. 2003), sea ice concentration from the satellite dataset of Fetterer et al. (2017), and NH sea ice concentration and sea ice thickness from PIOMAS20C (Schweiger et al. 2019). For proxy data verification, we use the PAGES2k V2 dataset (PAGES2k Consortium 2017) by withholding some proxies from assimilation using the bootstrap procedure described below.

a. Instrumental verification

Measured by correlation with the HadCRUT5 instrumental dataset during 1880–2000 CE, the LMR seasonal reconstruction skill in annual-mean 2-m air temperature is similar to or better than other reconstructions in the global mean (Fig. 3) despite assimilating fewer proxies. The spatial correlation pattern shows that LMR seasonal performs better than LMR v2 primarily in Europe, the North Pacific Ocean, the Indian Ocean, and the South Atlantic Ocean, but less well over the Southern Ocean and portions of Asia. Compared to PHYDA, the most pronounced differences are found in the Southern Ocean and the Indian Ocean. Similar results are found when

verifying against the ERA-20C reanalysis (Figs. S7 and S8). The limited number of proxies in the Southern Hemisphere means that the main signal for reconstructing the Southern Ocean relies heavily on long-distance teleconnections in the model prior. Since we do not use covariance localization, it is possible that bias in the Southern Ocean teleconnections from LIM forecasts degrades the performance of LMR seasonal in this location. We note that in Antarctica, where there are abundant ice core records (e.g., Steig et al. 2013; Stenni et al. 2017), the reconstructions perform comparatively well, with LMR seasonal generally being superior.

Seasonal verification against HadCRUT5 reveals that LMR seasonal has skill globally with positive correlations almost everywhere (Fig. 4). Compared to PHYDA, which is the only other reconstruction that includes DJF and JJA reconstructions, LMR seasonal performs relatively better during DJF than JJA, which we attribute to the sparse proxy data during DJF compared to JJA. PHYDA performs better during Northern Hemisphere summer, particularly over North America and Eurasia, which we attribute to the much larger number of tree-ring proxies that PHYDA assimilates in these locations.

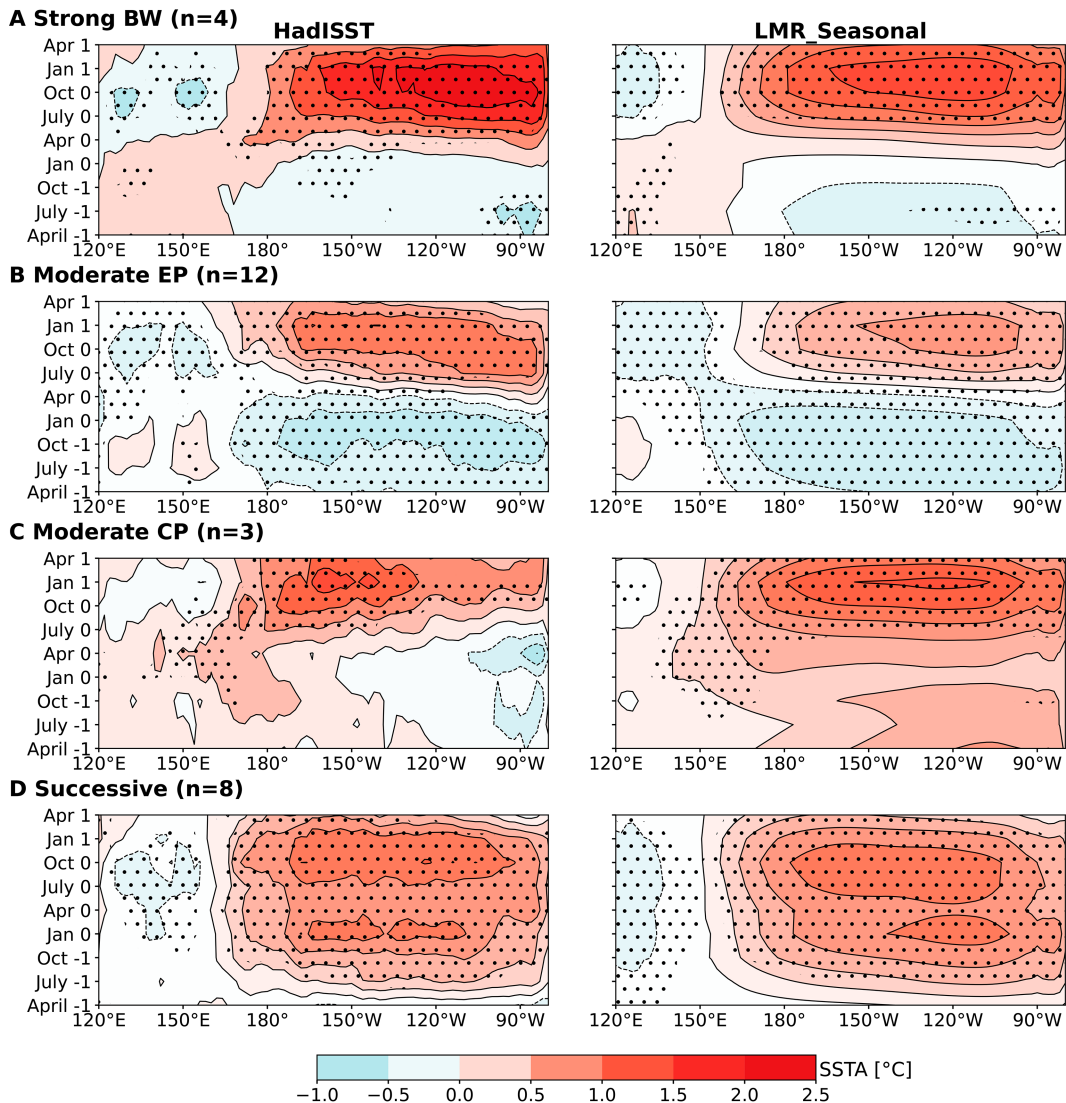


FIG. 10. Verification of the four classes of El Niño onset evolution of Wang et al. (2019) during 1900–2000. (left) Composite analyses from HadISST (Rayner et al. 2003); (right) the LMR seasonal reconstruction ensemble mean. Rows show composite averages of the four El Niño classes: (a) SBW, (b) MEP, (c) MCP, and (d) successive. Black dots show the composite SST anomaly values passing the confidence level of 95% using the Student's t test.

We also compare LMR seasonal with ModE-RA, a monthly offline DA product (Valler et al. 2024) from 1421 to 1900, as shown in Figs. S9 and S10. The correlation is largest in the locations and seasons when proxy availability is highest: JJA over western North America, central Asia, and northern Europe. This shows that the information content from the proxies is similar in both reconstructions, but how that information is spread in space and time is different. We attribute this to 1) the state-dependent prior covariances in LMR seasonal (static in ModE-RA), and 2) the way in which information is transmitted from times of high proxy availability to low (summer to winter). After 1900, ModE-RA assimilates a large number of instrumental observations, which results in higher correlations with LMR seasonal.

We verify upper 300-m ocean heat content (OHC300) against the Hadley EN4 dataset (Good et al. 2013) (Fig. 5) and Arctic SIC against the satellite observations of Fetterer et al. (2017) (Fig. 6). Despite not incorporating any direct observations of these quantities, the LMR seasonal reconstruction shows high correlation with the verification datasets. For OHC300, skill is highest in the tropical Pacific, the eastern North Pacific, and northern Atlantic Ocean regions. Skill in SIC is highest in Hudson Bay and near sea ice edges, especially around Greenland and the Barents Sea. Skill is lowest in the Beaufort Sea during DJF and MAM, seasons where proxies are least abundant. In contrast, SIC generally exhibits higher correlations in JJA compared to other seasons. We speculate that this is due to SIC having a stronger correlation

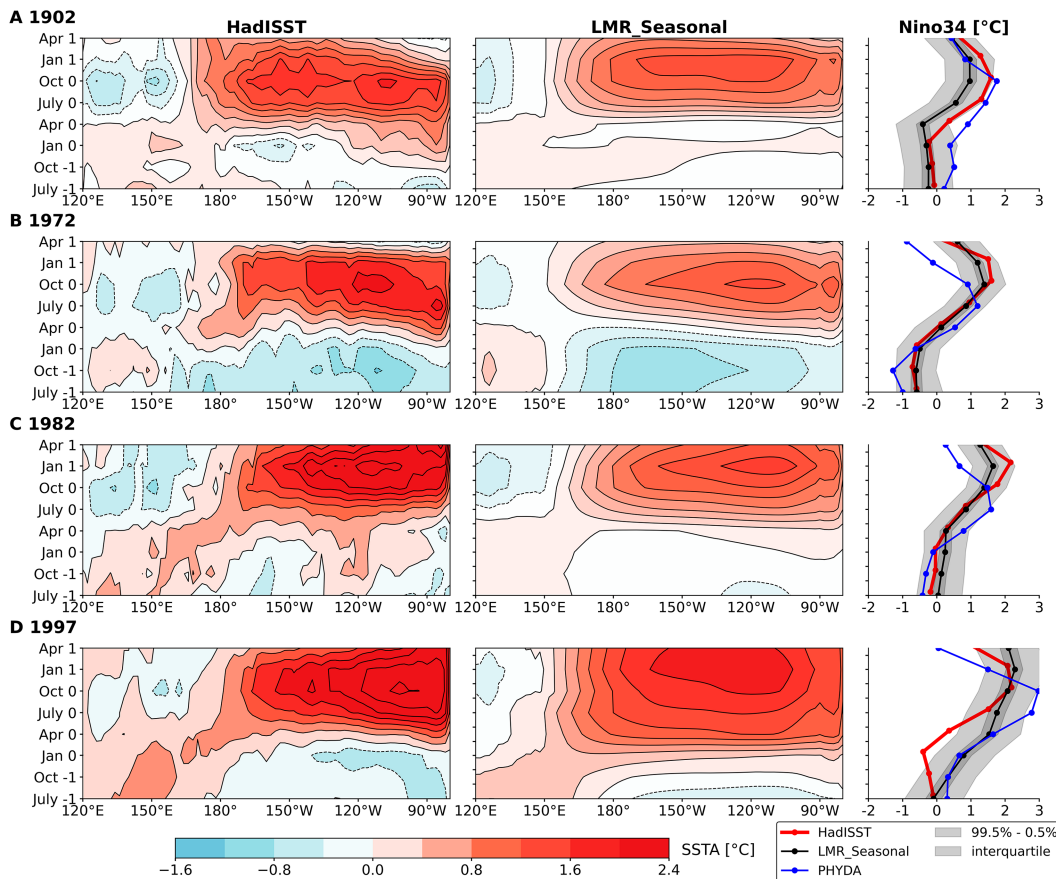


FIG. 11. Verification of four SBW (super) El Niño cases' onset evolution (1902, 1972, 1982, and 1997). (left) The evolution in HadISST (Rayner et al. 2003); (center) the LMR seasonal reconstruction ensemble mean. (right) The time series of the Niño-3.4 index from HadISST (red), PHYDA (blue), and LMR seasonal reconstruction mean (black). The dark shading denotes the interquartile range, and light shading denotes the central 99% confidence interval.

with 2-m air temperature in JJA compared to other seasons (Blanchard-Wrigglesworth et al. 2011).

Furthermore, we also compare NH sea ice area (SIA) and sea ice volume (SIV) with PIOMAS20C (Schweiger et al. 2019), which is produced by forcing a sea ice model with the ERA20C atmospheric reanalysis (Fig. 7, and Fig. S12). Discrepancies appear in the early twentieth century: PIOMAS-20C shows a rapid increase in SIA around 1900, likely due to sensitivity to the initial conditions (PIOMAS20C is initialized with HadISST, version 2.1.0.0), whereas the rest of the twentieth century shows much better agreement. The more persistent quantity, SIV, shows differences through the first half of the twentieth century. LMR seasonal, which begins data assimilation in 700 CE, benefits from proxy information before 1901 and thus avoids abrupt early century adjustments. We also compare LMR seasonal SIA with IAPICE1 (Semenov et al. 2024), which reconstructs SIC for March and September during 1901–2019 based on the statistical covariability between sea ice concentration, sea surface temperature, and sea level pressure. LMR seasonal shows less temporal variance in sea ice area, which we attribute to ensemble averaging and the

fewer numbers of proxies relative to instrumental observations (see supplemental Fig. 11). We also compare sea ice extent (SIE) with the August reconstruction from Kinnard et al. (2011) over the last millennium (supplemental Fig. 14). Their reconstruction also exhibits greater temporal variability, likely due to its focus on a single month (August), in contrast to the seasonal (JJA) resolution used in LMR seasonal, and ensemble averaging in LMR seasonal. The correlation between the two reconstructions is 0.26.

Comparing the LMR seasonal global mean temperature (GMT) and Niño-3.4 index with values from the instrumental datasets HadCRUT5 and HadISST shows highly skillful reconstructions (Figs. 8 and 9). Specifically, GMT has a correlation with HadCRUT5 of about 0.9 in all seasons and in the annual mean. The Niño-3.4 index, which represents the intensity of ENSO, shows a correlation (CE) around 0.8 (0.55) in all seasons. The use of seasonal coral data significantly improves the accuracy of the reconstruction compared to annualized coral data, which shows a correlation (CE) of approximately 0.7 (0.3) with the annual coral reconstructions and HadISST (not shown). In addition, compared to PHYDA, our reconstruction has

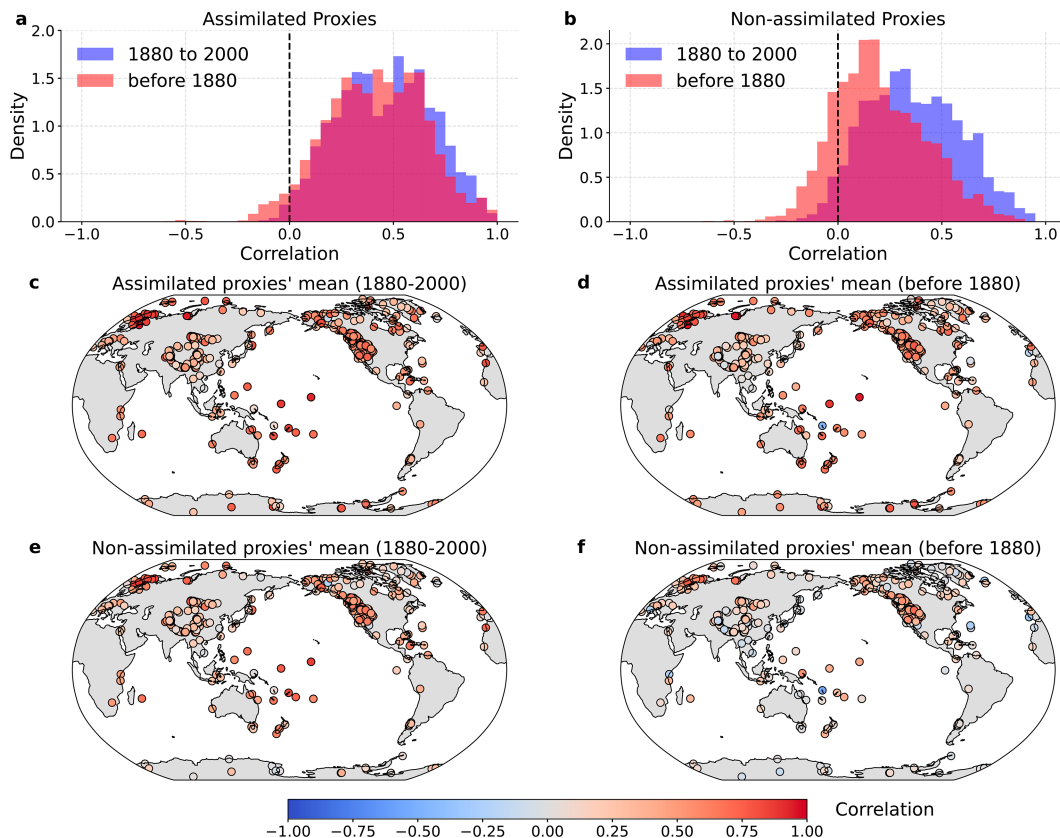


FIG. 12. Verification of LMR seasonal against assimilated and nonassimilated proxy data. (top) The distribution of correlation values between proxy values and LMR seasonal estimates for (a) assimilated and (b) nonassimilated proxy data from 1880 to 2000 (blue) and prior to 1880 (red). (middle) Proxy time-mean correlation maps for assimilated proxies (c) during 1880–2000 and (d) before 1880; (bottom) nonassimilated proxies (e) during 1880–2000 and (f) before 1880.

higher correlation and CE in all seasons except in JJA's correlation (Fig. S13). The greatest contribution to improved ENSO reconstruction likely comes from the online DA scheme since ENSO exhibits forecast skill on seasonal to annual time scales (McPhaden et al. 2006; Perkins and Hakim 2020; Meng and Hakim 2024; Meng and Li 2024); information from the previous season persists to subsequent seasons, providing a more accurate prior.

The ENSO reconstruction allows us to investigate the variability of El Niño over the last millennium with a much larger sample than is available during the instrumental period. Following Wang et al. (2019), we categorize El Niño events into four classes based on their onset-phase evolution: strong basin wide (SBW), moderate eastern Pacific (MEP), moderate central Pacific (MCP), and successive. The composite analysis of El Niño cases from 1900 to 2000 based on this classification is illustrated in Fig. 10. As described by Wang et al. (2019), SBW events are characterized by a combination of eastward SST anomalies (SSTAs) from the western Pacific and westward SSTA from the eastern Pacific, leading to strong warming events (“super El Niño”). In contrast, MEP and MCP events are defined by westward and eastward SSTA from the eastern and western Pacific, respectively, resulting in moderate

warming. Successive cases describe two consecutive years of sustained El Niño conditions. Compared to HadISST data, the LMR seasonal reconstruction successfully captures most of the seasonal evolution of these four El Niño classes, although the amplitude of the SBW and MEP cases is smaller in the reconstruction than in HadISST. The most significant discrepancies occur for MCP due to the small sample size (three events) since most cases are after the year 2000; many regions still align well with HadISST observations. Furthermore, as depicted in Fig. 11, all four SBW cases demonstrate consistent evolution with the reconstructed Niño-3.4 index evolution, closely following the HadISST time series. In comparison, the PHYDA reconstruction does not align as closely with HadISST, especially during JJA and SON.

In summary, instrumental verification shows that the LMR seasonal reconstruction faithfully captures a wide range of coupled atmosphere–ocean–sea ice climate variability in space and time during the instrumental period.

b. Independent proxy verification

To assess the robustness of LMR seasonal in the preinstrumental period (800–1850 CE), we validate against proxies left out of the assimilation process following Hakim et al. (2016).

We employ the bootstrap method, randomly omitting 20% of the proxies and conducting DA across 50 epochs. For each epoch, the proxies are forward modeled from the reconstructed climate states using the PSM (7) for each proxy, yielding a direct comparison of the LMR seasonal reconstruction to both the assimilated and independent proxy chronologies. The comparison is summarized by the time series correlation between the reconstructed and actual proxy time series. For the assimilated proxies, we find that the distribution of correlation values during the calibration and precalibration periods is very similar, suggesting a robust PSM relationship (Fig. 12). Results for nonassimilated proxies are similar, but with somewhat lower correlation values. There are no significant differences between these results and those when seasonality is defined by the expert-based seasonality defined in the PAGES2K database results (cf. Fig. 12 and Fig. S15). This indicates that the results are insensitive to the exact definition of proxy seasonality.

4. Last millennium seasonal temperature trends, medieval climate anomaly, and little ice age

From 850 to 1850 CE, most proxy evidence suggests that Earth experienced a cooling climate trend driven by orbital forcing and significant volcanic eruptions (McGregor et al. 2015). We compare our reconstructed seasonal temperature trends with seasonal trends from the CCSM4 Last Millennium simulation, as depicted in Fig. 13 and Fig. S16. Both the reconstructed and modeled trends show enhanced cooling in DJF and SON relative to MAM and JJA, which have smaller trends. This seasonal difference is attributed to the delayed climate response to orbital forcing (Fig. 13b, red dashed curve), as discussed by Lücke et al. (2021). It is important to note that the LIM trained on this model simulation has no trend and no season-specific variability (i.e., a single LIM is used for all seasons); the reconstructed seasonal trends arise solely from assimilation of proxy data.

Over the Last Millennium, two significant periods of multicentennial climate variability are the MCA and the LIA. Following Mann et al. (2009), we define the MCA as the period from 950 to 1250 CE and the LIA from 1400 to 1700 CE. Comparing the annual-mean global-mean temperature in LMR seasonal with three other reconstructions, we find that LMR seasonal has larger multicentennial variability, especially with respect to the MCA–LIA difference (Fig. 14a). The spatial pattern of temperature differences between these two time periods (MCA–LIA) reveals a common pattern of Arctic-amplified warming among the reconstructions (Figs. 14b–e). Notable differences are the much larger amplitude signal in LMR seasonal and the opposite sign of tropical Pacific temperature difference when compared with PYHDA. Moreover, LMR seasonal shows an MCA–LIA pattern over parts of Antarctica and the Southern Ocean that is mostly absent in the other reconstructions. This pattern is consistent with the independent temperature reconstruction of Orsi et al. (2012) based on a borehole thermometry analysis in West Antarctica (Fig. S17). We attribute these high-latitude differences to polar amplification having larger amplitude on

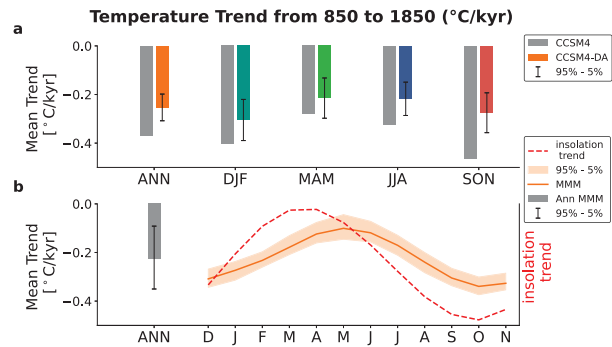


FIG. 13. Comparison of LM GMT trends for LMR seasonal and the climate model LM simulations. (a) The GMT trend for the annual mean and each season. The gray bar denotes the CCSM4 LM simulation, and the colors denote LMR seasonal. Error bars indicate the 90% ensemble confidence interval. (b) The GMT trends for the annual mean (gray bar) and each month (orange solid line) from the Last Millennium CMIP5 multimodel large ensemble (CCSM4, CESM-LME, CSIRO-Mk3L-1-2, MPI-ESM-P, IPSL-CM5A-LR, and HadCM3) (Taylor et al. 2012). The top-of-atmosphere insolation trend (right y axis) is shown as the red dashed line (Laskar et al. 2004). Error bars and orange shading represent the central 90% confidence interval.

seasonal time scales, which leads to more signal in the annual mean (Fig. S18).

As discussed above with regard to seasonal temperature trends, summer exhibits the least cooling trend over the last millennium, attributable to differences in insolation trends and seasonal lag due to ocean heat content (Lücke et al. 2021). Most proxies, especially NH tree-ring width and late-wood density, predominantly record JJA temperatures (e.g., Briffa et al. 1992; Anchukaitis et al. 2017). Annual-mean data assimilation dilutes the influence of JJA proxies, which likely reduces cooling trends over the Last Millennium. Our seasonal-update strategy ensures that connections between seasons are dynamically connected by the LIM, rather than static as in offline DA approaches. We hypothesize that these factors collectively contribute to the distinct differences we observe between the MCA and LIA. To test this hypothesis, we perform another experiment, allowing seasonal proxies to update only the annual mean during assimilation. Results show that the global-mean temperature difference between the MCA and LIA decreases by 30%, from 0.15° to 0.10°C. In this case, our seasonal-update strategy appears to be essential to reconstructing the magnitude of the MCA–LIA difference.

Differences between the MCA and LIA are also evident in sea ice area, sea ice volume, and OHC300 (Fig. 15). Specifically, sea ice area increases around 5% from 1.1×10^{13} to 1.15×10^{13} m² from the MCA to the LIA. Sea ice volume increases by about 11% from 3.5×10^{13} to 3.9×10^{13} m³, which we attribute to the longer persistence time of sea ice volume relative to area (Guemas et al. 2016). Compared to the sea ice area reconstruction of Brennan and Hakim (2022), we find largely similar centennial-scale results (Fig. 15a). In particular, both reconstructions show a decline in sea ice area that began in the early 19th century and continues to the present. LMR

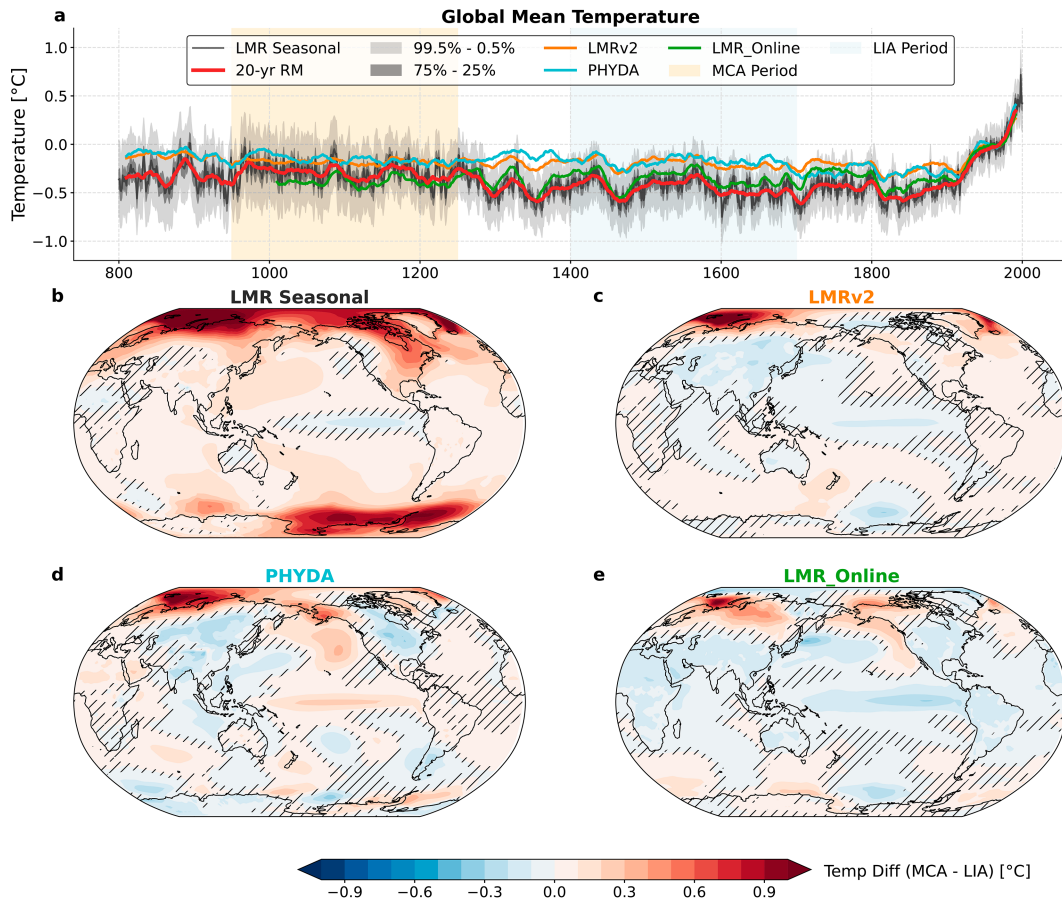


FIG. 14. Differences between the MCA (950–1250 CE) and LIA (1400–1700 CE) in four DA reconstructions. (a) GMT 20-yr running mean in LMR seasonal (red), LMRv2 (yellow), LMR online (green), and PHYDA (blue). The black solid curve represents the LMR seasonal unsmoothed GMT, dark shading denotes the interquartile range, and light shading denotes the central 99% confidence interval. (b)–(e) Global temperature pattern differences between the MCA and LIA from (b) LMR seasonal, (c) LMRv2, (d) PHYDA, and (e) LMR online. Hatching denotes regions that do not pass the 95% confidence level according to Student's t test.

seasonal has less amplitude on decadal time scales, which is especially evident during the early twentieth century warming (1920–50). We attribute this difference to weak covariability between 2-m air temperature and sea ice in the CCSM4 LM simulation, which [Brennan and Hakim \(2022\)](#) rectified with covariance inflation; here, we do not use covariance inflation. Finally, we note that differences in OHC300 show a decrease from the MCA to the LIA of about $1\text{--}1.5 \times 10^8 \text{ J m}^{-2}$, or an average of about 10 mW m^{-2} . We note that the overall trend in ocean heat content is similar to that shown in [Gebbie and Huybers \(2019\)](#).

5. Discussion and conclusions

This study introduces LMR seasonal, a new reconstruction of coupled atmosphere–ocean–sea ice climate variability over the Last Millennium, using a novel seasonal “online” data assimilation method and a new seasonal update strategy. The reconstruction is skillful in both space and time when compared with instrumental observations across the climate

variables considered. Skill is primarily attributed to the efficient utilization of proxy information, allowing for updates to model forecasts during assimilation that accurately reflect seasonal variability in the proxies. Additionally, verification against independent (nonassimilated) proxies shows the robustness of the reconstruction in the preinstrumental period.

We used the new reconstruction to examine two key measures of climate variability over the last millennium: ENSO and preinstrumental trends related to the transition from the medieval climate anomaly to the little ice age. For ENSO, LMR seasonal is able to accurately capture the space–time evolution of tropical SST for four different ENSO categories during the twentieth century. Given the large increase in sample size of ENSO over the last millennium compared to the twentieth and twenty-first centuries, LMR seasonal potentially offers a new resource for ENSO research. For temperature trends of the last millennium, we find that LMR seasonal captures seasonal variability consistent with orbital forcing, including polar amplification. Moreover, LMR seasonal demonstrates a pronounced distinction between the medieval

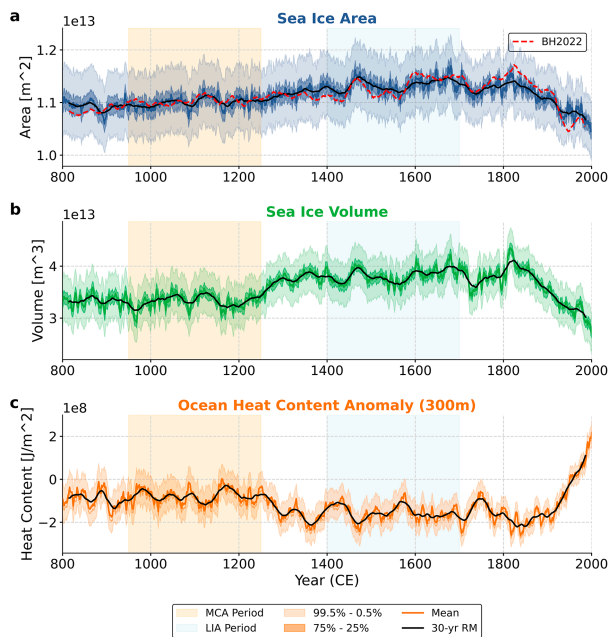


FIG. 15. Time series of Arctic sea ice (a) area and (b) volume, and (c) upper 300-m ocean heat content anomaly over the last millennium. The solid colored lines represent the ensemble mean, black solid lines denote the 30-yr running means, dark shading denotes the interquartile range, and light shading denotes the central 99% confidence interval. The red dashed line (BH2022) in (a) is the 30-yr running mean of SIA from Brennan and Hakim (2022). Light-orange shading denotes the MCA, and light-blue shading denotes the LIA.

climate anomaly (MCA) and the little ice age (LIA), consistent with established climatological studies. This distinction is significantly enhanced by the seasonal updating scheme, which ensures that summer-biased proxies do not dilute the cooler signatures of other seasons during the LIA.

While the LMR seasonal framework is not specifically optimized for sea ice reconstruction, it nonetheless demonstrates reasonable skill compared to previous studies in capturing large-scale sea ice variability, justifying the inclusion of sea ice in our coupled DA system.

Future studies could expand upon this work by incorporating additional proxy data and exploring regional climate events during the Last Millennium with much larger samples than are available with instrumental reanalyses. Moreover, extending this approach to reconstructions at finer spatial resolution could provide deeper insights into regional climate phenomena and their global implications.

Acknowledgments. This research was supported by NSF Awards 2402475, 2202526, and 2105805 and Heising-Simons Foundation Award 2023-4715. We thank the editor and all three reviewers for their constructive and insightful comments, which have greatly improved the manuscript. We appreciate comments and conversations related to this work with Feng Zhu (NSF National Center for Atmospheric

Research), Olivia Truax (University of Canterbury), Julien Emile-Geay (University of Southern California), Dominik Stiller (University of Washington), and Vince Cooper (University of Washington).

Data availability statement. The PSM package “CFR” (Zhu et al. 2023, 2024) and PAGES2K database are located at <https://github.com/fzhu2e/cfr>. The plotting package SACP (Meng et al. 2021, 2023) is located at <https://github.com/Zilum/sacpy>. The LMR seasonal reconstruction code is located at https://github.com/Zilum/LMR_Seasonal (Meng et al. 2025). The LMR Seasonal Reanalysis Dataset is available at <https://doi.org/10.5281/zenodo.17268597> for the ensemble mean, and the full dataset can be accessed at https://atmos.uw.edu/~zilumeng/LMR_Seasonal/.

REFERENCES

- Anchukaitis, K. J., and Coauthors, 2017: Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quat. Sci. Rev.*, **163**, 1–22, <https://doi.org/10.1016/j.quascirev.2017.02.020>.
- Anderson, J. L., 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **140**, 2359–2371, <https://doi.org/10.1175/MWR-D-11-00013.1>.
- Blanchard-Wrigglesworth, E., K. C. Armour, C. M. Bitz, and E. DeWeaver, 2011: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations. *J. Climate*, **24**, 231–250, <https://doi.org/10.1175/2010JCLI3775.1>.
- Bouttier, F., and P. Courtier, 2002: Data assimilation concepts and methods March 1999. Meteorological Training Course Lecture Series ECMWF 718, 59 pp., <https://www.ecmwf.int/sites/default/files/elibrary/2002/16928-data-assimilation-concepts-and-methods.pdf>.
- Brennan, M. K., 2022: Reconstructing Arctic sea ice in the instrumental era. Ph.D. thesis, University of Washington, 123 pp., <https://digital.lib.washington.edu/researchworks/bitstreams/94e104f7-f5bd-4e0b-a126-38ef23bfe430/download>.
- , and G. J. Hakim, 2022: Reconstructing Arctic sea ice over the Common Era using data assimilation. *J. Climate*, **35**, 1231–1247, <https://doi.org/10.1175/JCLI-D-21-0099.1>.
- Briffa, K. R., P. D. Jones, and F. H. Schweingruber, 1992: Tree-ring density reconstructions of summer temperature patterns across western North America since 1600. *J. Climate*, **5**, 735–754, [https://doi.org/10.1175/1520-0442\(1992\)005<0735:TRDROS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0735:TRDROS>2.0.CO;2).
- , T. J. Osborn, and F. H. Schweingruber, 2004: Large-scale temperature inferences from tree rings: A review. *Global Planet. Change*, **40**, 11–26, [https://doi.org/10.1016/S0921-8181\(03\)00095-X](https://doi.org/10.1016/S0921-8181(03)00095-X).
- Brown, J. R., and Coauthors, 2020: Comparison of past and future simulations of ENSO in CMIP5/PMIP3 and CMIP6/PMIP4 models. *Climate Past*, **16**, 1777–1805, <https://doi.org/10.5194/cp-16-1777-2020>.
- Christiansen, B., and F. C. Ljungqvist, 2017: Challenges and perspectives for large-scale temperature reconstructions of the past two millennia. *Rev. Geophys.*, **55**, 40–96, <https://doi.org/10.1002/2016RG000521>.
- Dee, S., Y. Okumura, S. Stevenson, and P. Di Nezio, 2020: Enhanced North American ENSO teleconnections during the little ice age revealed by paleoclimate data assimilation.

- Geophys. Res. Lett.*, **47**, e2020GL087504, <https://doi.org/10.1029/2020GL087504>.
- Dirren, S., and G. J. Hakim, 2005: Toward the assimilation of time-averaged observations. *Geophys. Res. Lett.*, **32**, L04804, <https://doi.org/10.1029/2004GL021444>.
- Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, <https://doi.org/10.1007/s10236-003-0036-9>.
- Fetterer, F., K. Knowles, W. Meier, M. Savoie, and A. Windnagel, 2017: Sea Ice Index, version 3. National Snow and Ice Data Center, accessed 1 February 2025, <https://nsidc.org/data/g02135/versions/3>.
- Franke, J., S. Brönnimann, J. Bhend, and Y. Brugnara, 2017: A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations. *Sci. Data*, **4**, 170076, <https://doi.org/10.1038/sdata.2017.76>.
- Gebbie, G., and P. Huybers, 2019: The little ice age and 20th-century deep Pacific cooling. *Science*, **363**, 70–74, <https://doi.org/10.1126/science.aar8413>.
- Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Oceans*, **118**, 6704–6716, <https://doi.org/10.1126/science.aar8413>.
- Goosse, H., E. Cresspin, A. de Montety, M. E. Mann, H. Renssen, and A. Timmermann, 2010: Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. *J. Geophys. Res.*, **115**, D09108, <https://doi.org/10.1029/2009JD012737>.
- Guemas, V., and Coauthors, 2016: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **142**, 546–561, <https://doi.org/10.1002/qj.2401>.
- Hakim, G. J., J. Emile-Geay, E. J. Steig, D. Noone, D. M. Anderson, R. Tardif, N. Steiger, and W. A. Perkins, 2016: The last millennium climate reanalysis project: Framework and first results. *J. Geophys. Res. Atmos.*, **121**, 6745–6764, <https://doi.org/10.1002/2016JD024751>.
- , C. Snyder, S. G. Penny, and M. Newman, 2022: Subseasonal forecast skill improvement from strongly coupled data assimilation with a linear inverse model. *Geophys. Res. Lett.*, **49**, e2022GL097996, <https://doi.org/10.1029/2022GL097996>.
- Hu, W., and Coauthors, 2024: Reconstructing tropical monthly sea surface temperature variability by assimilating coral proxy datasets. *npj Climate Atmos. Sci.*, **7**, 261, <https://doi.org/10.1038/s41612-024-00816-w>.
- Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, <https://doi.org/10.1175/JCLI-D-16-0836.1>.
- Kinnard, C., C. M. Zdanowicz, D. A. Fisher, E. Isaksson, A. de Vernal, and L. G. Thompson, 2011: Reconstructed changes in Arctic sea ice over the past 1,450 years. *Nature*, **479**, 509–512, <https://doi.org/10.1038/nature10581>.
- Laskar, J., P. Robutel, F. Joutel, M. Gastineau, A. C. M. Correia, and B. Levrard, 2004: A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.*, **428**, 261–285, <https://doi.org/10.1051/0004-6361:20041335>.
- Lenssen, N. J. L., G. A. Schmidt, J. E. Hansen, M. J. Menne, A. Persin, R. Ruedy, and D. Zys, 2019: Improvements in the GISTEMP uncertainty model. *J. Geophys. Res. Atmos.*, **124**, 6307–6326, <https://doi.org/10.1029/2018JD029522>.
- Lücke, L. J., A. P. Schurer, R. Wilson, and G. C. Hegerl, 2021: Orbital forcing strongly influences seasonal temperature trends during the last millennium. *Geophys. Res. Lett.*, **48**, e2020GL088776, <https://doi.org/10.1029/2020GL088776>.
- Luo, X., S. Dee, S. Stevenson, Y. Okumura, N. Steiger, and L. Parsons, 2022: Last millennium ENSO diversity and North American teleconnections: New insights from paleoclimate data assimilation. *Paleoceanogr. Paleoclimatol.*, **37**, e2021PA004283, <https://doi.org/10.1029/2021PA004283>.
- Mann, M. E., and Coauthors, 2009: Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science*, **326**, 1256–1260, <https://doi.org/10.1126/science.1177303>.
- Mantua, N. J., and S. R. Hare, 2002: The Pacific decadal oscillation. *J. Oceanogr.*, **58**, 35–44, <https://doi.org/10.1023/A:1015820616384>.
- McGregor, H. V., and Coauthors, 2015: Robust global ocean cooling trend for the pre-industrial Common Era. *Nat. Geosci.*, **8**, 671–677, <https://doi.org/10.1038/ngeo2510>.
- McPhaden, M. J., S. E. Zebiak, and M. H. Glantz, 2006: ENSO as an integrating concept in Earth science. *Science*, **314**, 1740–1745, <https://doi.org/10.1126/science.1132588>.
- Meng, Z., and G. J. Hakim, 2024: Reconstructing the tropical Pacific upper ocean using online data assimilation with a deep learning model. *J. Adv. Model. Earth Syst.*, **16**, e2024MS004422, <https://doi.org/10.1029/2024MS004422>.
- , and T. Li, 2024: Why is the Pacific meridional mode most pronounced in boreal spring? *Climate Dyn.*, **62**, 459–471, <https://doi.org/10.1007/s00382-023-06914-4>.
- , Z. Hu, Z. Ai, Y. Zhang, and K. Shan, 2021: Research on planar double compound pendulum based on RK-8 algorithm. *J. Big Data*, **3**, 11–20, <https://doi.org/10.32604/jbd.2021.015208>.
- , F. Zhu, and G. J. Hakim, 2023: Sacy—A Python package for statistical analysis of climate. *AGU Fall Meeting 2023*, San Francisco, CA, Amer. Geophys. Union, Abstract PP13D-1252, <https://ui.adsabs.harvard.edu/abs/2023AGUFMPP13D1252M/abstract>.
- , G. J. Hakim, and E. J. Steig, 2025: Coupled Seasonal data assimilation of sea ice, ocean, and atmospheric dynamics over the last millennium. arXiv, 2501.14130v1, <https://doi.org/10.48550/arXiv.2501.14130>.
- Morice, C. P., and Coauthors, 2021: An updated assessment of near-surface temperature change from 1850: The HadCRUT5 data set. *J. Geophys. Res. Atmos.*, **126**, e2019JD032361, <https://doi.org/10.1029/2019JD032361>.
- Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.*, **10**, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Newman, M., 2013: An empirical benchmark for decadal forecasts of global surface temperature anomalies. *J. Climate*, **26**, 5260–5269, <https://doi.org/10.1175/JCLI-D-12-00590.1>.
- Okazaki, A., T. Miyoshi, K. Yoshimura, S. J. Grebush, and F. Zhang, 2021: Revisiting online and offline data assimilation comparison for paleoclimate reconstruction: An idealized OSSE study. *J. Geophys. Res. Atmos.*, **126**, e2020JD034214, <https://doi.org/10.1029/2020JD034214>.
- Orsi, A. J., B. D. Cornuelle, and J. P. Severinghaus, 2012: Little ice age cold interval in West Antarctica: Evidence from borehole temperature at the West Antarctic Ice Sheet (WAIS) divide. *Geophys. Res. Lett.*, **39**, L09710, <https://doi.org/10.1029/2012GL051260>.

- Pachauri, R. K., and Coauthors, 2007: *IPCC Fourth Assessment Report*. IPCC, 673 pp.
- PAGES2k Consortium, 2013: Continental-scale temperature variability during the past two millennia. *Nat. Geosci.*, **6**, 339–346, <https://doi.org/10.1038/ngeo1797>.
- , 2017: A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data*, **4**, 170088, <https://doi.org/10.1038/sdata.2017.88>.
- Penland, C., and T. Magorian, 1993: Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Climate*, **6**, 1067–1076, [https://doi.org/10.1175/1520-0442\(1993\)006<1067:PONSST>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1067:PONSST>2.0.CO;2).
- , and L. Matrosova, 1994: A balance condition for stochastic numerical models with application to the El Niño–Southern Oscillation. *J. Climate*, **7**, 1352–1372, [https://doi.org/10.1175/1520-0442\(1994\)007<1352:ABCFNS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1352:ABCFNS>2.0.CO;2).
- , and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2024, [https://doi.org/10.1175/1520-0442\(1995\)008<1999:TOGOTS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<1999:TOGOTS>2.0.CO;2).
- Perkins, W. A., and G. Hakim, 2020: Linear inverse modeling for coupled atmosphere–ocean ensemble climate prediction. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001778, <https://doi.org/10.1029/2019MS001778>.
- , and G. J. Hakim, 2021: Coupled atmosphere–ocean reconstruction of the last millennium using online data assimilation. *Paleoceanogr. Paleoclimatol.*, **36**, e2020PA003959, <https://doi.org/10.1029/2020PA003959>.
- Poli, P., and Coauthors, 2016: ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Climate*, **29**, 4083–4097, <https://doi.org/10.1175/JCLI-D-15-0556.1>.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. R. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, <https://doi.org/10.1029/2002JD002670>.
- Roach, L. A., and Coauthors, 2020: Antarctic sea ice area in CMIP6. *Geophys. Res. Lett.*, **47**, e2019GL086729, <https://doi.org/10.1029/2019GL086729>.
- Rohde, R. A., and Z. Hausfather, 2020: The Berkeley Earth land/ocean temperature record. *Earth Syst. Sci. Data*, **12**, 3469–3479, <https://doi.org/10.5194/essd-12-3469-2020>.
- Schulzweida, U., L. Kornbluh, and R. Quast, 2019: CDO user guide. CDO Doc., 263 pp., <https://code.mpimet.mpg.de/projects/cdo/embedded/cdo.pdf>.
- Schweiger, A. J., K. R. Wood, and J. Zhang, 2019: Arctic sea ice volume variability over 1901–2010: A model-based reconstruction. *J. Climate*, **32**, 4731–4752, <https://doi.org/10.1175/JCLI-D-19-0008.1>.
- Semenov, V. A., T. A. Aldonina, F. Li, N. S. Keenlyside, and L. Wang, 2024: Arctic sea ice variations in the first half of the 20th century: A new reconstruction based on hydrometeorological data. *Adv. Atmos. Sci.*, **41**, 1483–1495, <https://doi.org/10.1007/s00376-024-3320-x>.
- Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA’s historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296, <https://doi.org/10.1175/2007JCLI2100.1>.
- Steig, E. J., and Coauthors, 2013: Recent climate and ice-sheet changes in West Antarctica compared with the past 2,000 years. *Nat. Geosci.*, **6**, 372–375, <https://doi.org/10.1038/ngeo1778>.
- Steiger, N. J., E. J. Steig, S. G. Dee, G. H. Roe, and G. J. Hakim, 2017: Climate reconstruction using data assimilation of water isotope ratios from ice cores. *J. Geophys. Res. Atmos.*, **122**, 1545–1568, <https://doi.org/10.1002/2016JD026011>.
- , J. E. Smerdon, E. R. Cook, and B. I. Cook, 2018: A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Sci. Data*, **5**, 180086, <https://doi.org/10.1038/sdata.2018.86>.
- Stenni, B., and Coauthors, 2017: Antarctic climate variability on regional and continental scales over the last 2000 years. *Climate Past*, **13**, 1609–1634, <https://doi.org/10.5194/cp-13-1609-2017>.
- Sun, H., L. Lei, Z. Liu, L. Ning, and Z.-M. Tan, 2022: An analog offline EnKF for paleoclimate data assimilation. *J. Adv. Model. Earth Syst.*, **14**, e2021MS002674, <https://doi.org/10.1029/2021MS002674>.
- Tardif, R., and Coauthors, 2019: Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. *Climate Past*, **15**, 1251–1273, <https://doi.org/10.5194/cp-15-1251-2019>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Valler, V., and Coauthors, 2024: ModE-RA: A global monthly paleo-reanalysis of the modern era 1421 to 2008. *Sci. Data*, **11**, 36, <https://doi.org/10.1038/s41597-023-02733-8>.
- Wang, B., X. Luo, Y.-M. Yang, W. Sun, M. A. Cane, W. Cai, S.-W. Yeh, and J. Liu, 2019: Historical change of El Niño properties sheds light on future changes of extreme El Niño. *Proc. Natl. Acad. Sci. USA*, **116**, 22 512–22 517, <https://doi.org/10.1073/pnas.1911130116>.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2).
- Widmann, M., H. Goosse, G. van der Schrier, R. Schnur, and J. Barkmeijer, 2010: Using data assimilation to study extratropical Northern Hemisphere climate over the last millennium. *Climate Past*, **6**, 627–644, <https://doi.org/10.5194/cp-6-627-2010>.
- Williams, E. J., 1959: The comparison of regression variables. *J. Roy. Stat. Soc.*, **21B**, 396–399, <https://doi.org/10.1111/j.2517-6161.1959.tb00346.x>.
- Zhu, F., J. Emile-Geay, K. J. Anchukaitis, G. J. Hakim, A. T. Wittenberg, M. S. Morales, M. Toohey, and J. King, 2022: A re-appraisal of the ENSO response to volcanism with paleoclimate data assimilation. *Nat. Commun.*, **13**, 747, <https://doi.org/10.1038/s41467-022-28210-1>.
- , —, N. P. McKay, S. Stevenson, and Z. Meng, 2023: A pseudoproxy emulation of the PAGES 2k database using a hierarchy of proxy system models. *Sci. Data*, **10**, 624, <https://doi.org/10.1038/s41597-023-02489-1>.
- , —, G. J. Hakim, D. Guillot, D. Khider, R. Tardif, and W. A. Perkins, 2024: Cfr (v2024.1.26): A Python package for climate field reconstruction. *Geosci. Model Dev.*, **17**, 3409–3431, <https://doi.org/10.5194/gmd-17-3409-2024>.