Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

ANNUAL REVIEWS

Annual Review of Earth and Planetary Sciences

Advances in Paleoclimate Data Assimilation

Jessica E. Tierney,[1] Emily J. Judd,[1] Matthew B. Osman,[2] Jonathan M. King,[1,3] Olivia J. Truax,[4,5] Nathan J. Steiger,[6] Daniel E. Amrhein,[7] and Kevin J. Anchukaitis[8]

[1]Department of Geosciences, University of Arizona, Tucson, Arizona, USA; email: jesst@arizona.edu

[2]Department of Geography, University of Cambridge, Cambridge, United Kingdom

[3]Current affiliation: United States Geological Survey, Golden, Colorado, USA

[4]Te Kura Aronukurangi School of Earth and Environment, Te Whare Wānanga o Waitaha, University of Canterbury, Aotearoa, New Zealand

[5]Department of Earth and Space Sciences, University of Washington, Seattle, Washington, USA

[6]Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel

[7]Climate and Global Dynamics Laboratory, NSF National Center for Atmospheric Research, Boulder, Colorado, USA

[8]School of Geography, Development, and Environment and Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona, USA

Annu. Rev. Earth Planet. Sci. 2025. 53:24.1–24.26

The *Annual Review of Earth and Planetary Sciences* is online at earth.annualreviews.org

https://doi.org/10.1146/annurev-earth-032320-064209

Copyright © 2025 by the author(s).
All rights reserved

Keywords

paleoclimate reconstruction, data assimilation, forward modeling, statistics, climate field reconstruction

Abstract

Reconstructions of past climates in both time and space provide important insight into the range and rate of change within the climate system. However, producing a coherent global picture of past climates is difficult because indicators of past environmental changes (proxy data) are unevenly distributed and uncertain. In recent years, paleoclimate data assimilation (paleoDA), which statistically combines model simulations with proxy data, has become an increasingly popular reconstruction method. Here, we describe advances in paleoDA to date, with a focus on the offline ensemble Kalman filter and the insights into climate change that this method affords. PaleoDA has considerable strengths in that it can blend multiple types of information while also propagating uncertainty. Drawbacks of the methodology include an overreliance on the climate model and variance loss. We

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

conclude with an outlook on possible expansions and improvements in paleoDA that can be made in the upcoming years.

- Paleoclimate data assimilation blends model and proxy information to enable spatiotemporal reconstructions of past climate change.
- This method has advanced our understanding of global temperature change, Earth's climate sensitivity, and past climate dynamics.
- Future innovations could improve the method by implementing online paleoclimate data assimilation and smoothers.

1. INTRODUCTION AND MOTIVATION

With current atmospheric CO_2 concentrations at levels not seen in at least 3 million years, we now live in a climate that is fundamentally different from the preindustrial state. While climate models can be used to simulate changes in the Earth system under high CO_2, they are typically tuned to and evaluated against the recent historical climate and may not produce reliable results outside of that range (e.g., Zhu et al. 2020). Reconstructions of past climates are therefore one of the best tools for understanding how Earth's climate system behaves under radically different forcings (Tierney et al. 2020a). In particular, spatially complete paleoclimate reconstructions, also known as climate field reconstructions (CFRs) (Evans et al. 1998), are crucial for constraining a number of aspects of the Earth system. They can identify the spatial imprint of internal climate variability (Wise 2016, Emile-Geay et al. 2020) and forced change (Fernández-Donado et al. 2013, Marvel & Cook 2022) and quantify how warming patterns affect critical metrics such as equilibrium climate sensitivity (ECS) (Cooper et al. 2024). However, the geological indicators of past climate—proxy data—used in CFRs are not evenly distributed in time or space. Each type of proxy also has its own set of strengths and limitations, including seasonal biases and temporal resolution. Any method for developing a CFR therefore needs to be able to use a sparse network of uncertain proxy data to estimate a complete climate field.

The challenge of producing a CFR from a sparse network of proxies is not new (Fritts et al. 1971, Webb & Bryson 1972), and many different methods exist (Tingley et al. 2012, Smerdon et al. 2023). Early CFR approaches were anchored in principal components and canonical correlation analysis, which reduce both the predictors (the proxies) and the predictand (the climate field) into a few leading spatiotemporal modes of variance whose time series can then be used in least squares regression (Fritts et al. 1971, Cook et al. 1994). Objective analysis and optimal interpolation methods, adapted from the techniques used to produce gridded products such as historical sea surface temperatures (SSTs) (Kaplan et al. 1997), have been used to generate CFRs on both recent (i.e., last millennium) (Evans et al. 2002) and more ancient (Gill et al. 2016, Tierney et al. 2019a) timescales. Similar to earlier methods, these approaches isolate a few leading modes of climate variability to project proxy information in space but can explicitly incorporate errors in the proxy-climate calibration and the representation of the modern climate field. These approaches are efficient, linear, and unbiased (Evans et al. 2001); however, the spatial covariance patterns used to broadcast the sparse proxy data into a continuous climate field are estimated from the modern climate state. Patterns of climate variability are almost certainly different on longer geological timescales, especially when Earth's boundary conditions (ice sheets, continental positions, topography, sea level, radiative forcing, etc.) change.

Another class of CFR uses Bayesian hierarchical modeling (BHM) (Tingley & Huybers 2010, 2013; Werner et al. 2018; Ossandón et al. 2024). BHMs offer several advantages: They can include

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

PROXY SYSTEM MODELS

Proxy system models (PSMs) describe the processes and uncertainties that link the proxy values measured in the lab back to the environmental information that they encode (Evans et al. 2013). They are forward models in the sense that the environment predicts the proxy outcome [as occurs in nature; i.e., proxy $\sim f$(climate)], rather than the other way around (i.e., inverse methods). PSMs vary considerably in their complexity and degree of sophistication, ranging from simple statistical regressions to complex mechanistic models with multiple levels. A multi-level PSM might first describe how the sensor (e.g., foraminifera) is responding to environmental variables (e.g., seawater temperature and $\delta^{18}O$) and then describe how what was recorded ($\delta^{18}O$ of the foraminifera) is affected by the archive (e.g., bioturbation in a sediment core).

PSMs are ideally suited for paleoclimate data assimilation because they can be used as the observational operator $\mathcal{H}\mathbf{X}_{\text{prior}}$ to create estimates of proxy values ($\hat{\mathbf{Y}}$). This enables an apples-to-apples comparison between the proxy and the model data by allowing $\hat{\mathbf{Y}}$ (e.g., $\delta^{18}O$ of the foraminifera) to be calculated (or forward modeled) using multiple different prior climate fields (e.g., sea surface temperature, $\delta^{18}O$ of seawater).

an explicit model for spatial covariance rather than assuming stable or modern patterns of covariance, facilitate complete uncertainty propagation, and accommodate proxy system models (PSMs) (see the sidebar titled Proxy System Models) (Tingley & Huybers 2010, Tingley et al. 2012). However, BHM spatial covariance models have thus far been Gaussian and isotropic, which can be insufficient for capturing anisotropic behavior in the climate system (for example, from ocean dynamics or topography). In addition, BHMs can be computationally intensive because they involve inversion with Markov-Chain Monte Carlo methods (Tingley & Huybers 2010).

The application of these and other CFR techniques (Schneider 2001, J. Wang et al. 2014, Smerdon et al. 2023) generated several methodological dilemmas: At what distance should a proxy be able to influence the reconstructed climate field (cf. Cook et al. 1999)? How should errors, uncertainties, and biases in the proxy data be incorporated? How should the covariance of the climate field be quantified? What methods should be used to project proxy information across the target climate field? These choices, as well as the covariance of the underlying climate target, can meaningfully influence important features of the resulting reconstructions (Smerdon et al. 2011, Dannenberg & Wise 2013, Wang et al. 2015).

Paleoclimate data assimilation (paleoDA) provides an alternative solution to the problem of estimating past spatiotemporal climate variability by blending proxy information with climate model simulations. In the paleoDA framework, model simulations are used as a starting point, or a prior in Bayesian parlance, for the past climate. The proxy information is then added to the prior (assimilated), weighted by proxy uncertainty, the prior model spread, and the covariance patterns in the model simulations. Like BHMs, paleoDA is easily integrated with PSMs, but it is more computationally efficient. It also uses model simulations of the past climate state, rather than modern (observed) climate data, to estimate spatial covariance, which is an advantage for CFR on longer geological timescales when the glacial extent, greenhouse gas concentrations, and position of the continents were different from today.

In this review, we explain the mechanics of paleoDA and discuss some applications that have enabled advances in paleoclimate in recent years. We focus on the offline ensemble Kalman filter (EnKF) method, which has emerged as a leading approach in the past decade. However, other approaches, including particle filters, played an important role in the development of paleoDA (Van Leeuwen 2009, Goosse et al. 2010, Dubinkina et al. 2011), and online methods continue to gain traction (as discussed in Section 5). We also discuss the challenges and limitations of the offline EnKF method that have been made apparent as paleoDA has become more widely used,

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

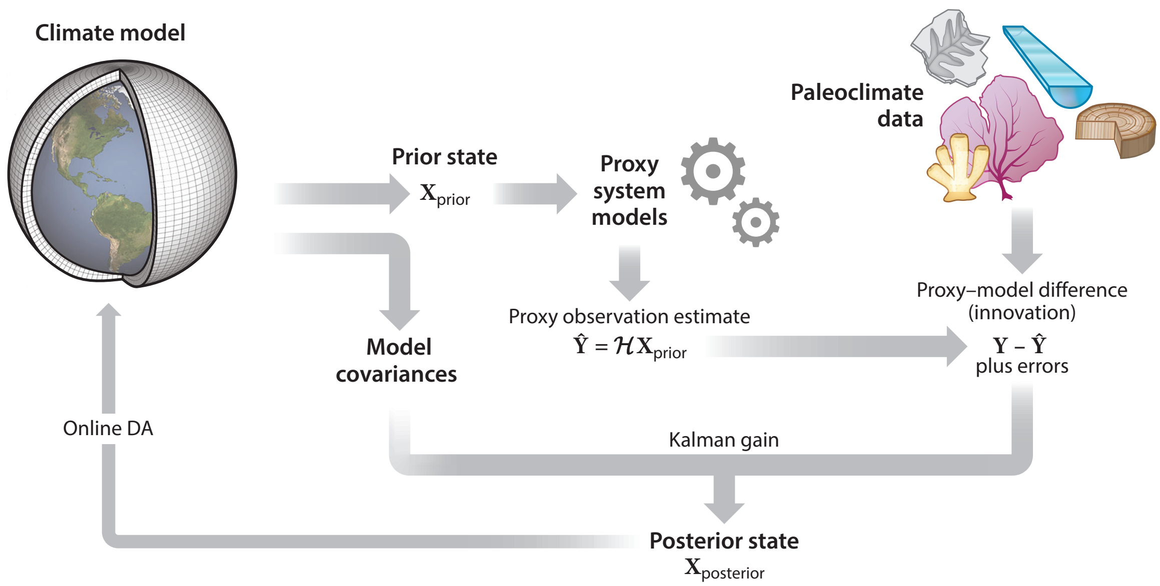

Figure 1

Overview of paleoDA. Climate models provide the initial estimate of the state ($\mathbf{X}_{\text{prior}}$), which is then mapped into an estimate of the proxy ($\hat{\mathbf{Y}}$) by proxy system models. $\hat{\mathbf{Y}}$ is compared to the actual proxy values $\mathbf{Y}$ with the difference representing the new information (the innovation) added to the reconstruction. The Kalman gain incorporates the new information into the model prior using the model covariance structures and weights the update by the uncertainties in $\mathbf{Y}$ and $\hat{\mathbf{Y}}$. This review focuses on offline paleoDA, but in online applications (discussed in Section 5), the posterior state feeds back into the climate model, so the model can react to the data update. Abbreviation: paleoDA, paleoclimate data assimilation.

and we conclude by highlighting possible improvements and extensions to paleoDA that can be made in the future.

2. ELEMENTS OF THE METHOD

Here we review the mathematics and components of the offline EnKF method (**Figure 1**), following the methodology introduced by Steiger et al. (2014). We begin with the overarching equations and then discuss each ingredient of paleoDA and some key considerations.

The goal of paleoDA is to combine an ensemble of modeling simulations (the model prior) with paleoclimate data (from particular locations and times) to produce a posterior ensemble that blends the two sources of information. We start with the classical update equation for the Kalman filter (Kalnay 2003):

$$\boldsymbol{x}_{\mathrm{a}} = \boldsymbol{x}_{\mathrm{b}} + \mathbf{K}(\boldsymbol{y} - \mathcal{H}\boldsymbol{x}_{\mathrm{b}}). \qquad 1.$$

In this setup, $\boldsymbol{x}_{\mathrm{b}}$ is the prior (background) state and $\boldsymbol{x}_{\mathrm{a}}$ is the posterior (analysis) state; both contain all the climate fields to be updated collapsed into a vector. $\boldsymbol{y}$ represents the observations to be assimilated, and $\mathcal{H}$ is an observation operator that maps $\boldsymbol{x}_{\mathrm{b}}$ to observation space so that the two can be compared. In paleoDA, $\boldsymbol{y}$ represents proxy information, and $\mathcal{H}$ contains PSMs. The difference between the observations and prior estimate of them ($\boldsymbol{y} - \mathcal{H}\boldsymbol{x}_{\mathrm{b}}$) is called the innovation and represents the new information being added to the prior state. This information is weighted by the Kalman gain $\mathbf{K}$:

$$\mathbf{K} = \mathrm{cov}(\boldsymbol{x}_{\mathrm{b}}, \mathcal{H}\boldsymbol{x}_{\mathrm{b}})[\mathrm{cov}(\mathcal{H}\boldsymbol{x}_{\mathrm{b}}, \mathcal{H}\boldsymbol{x}_{\mathrm{b}}) + \mathbf{R}]^{-1}. \qquad 2.$$

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

The first term of the Kalman gain, $\mathrm{cov}(\boldsymbol{x}_\mathrm{b}, \mathcal{H}\boldsymbol{x}_\mathrm{b})$, describes the relationship between the prior and the observational estimate (in our case, proxy estimates). This term spreads the information from the observations across the different fields in the model prior according to the covariance. The second term, $[\mathrm{cov}(\mathcal{H}\boldsymbol{x}_\mathrm{b}, \mathcal{H}\boldsymbol{x}_\mathrm{b}) + \mathbf{R}]^{-1}$, represents error in the prior estimate of the observations ($\mathcal{H}\boldsymbol{x}_\mathrm{b}$) and in the actual proxy data ($\mathbf{R}$).

Because paleoDA applications usually begin with an ensemble of model prior states, we rewrite these equations slightly for updating a prior matrix rather than a vector. For clarity, we call the prior $\mathbf{X}_\mathbf{prior}$ and posterior $\mathbf{X}_\mathbf{post}$, and we refer to the estimate of the proxy values from PSMs as $\hat{\mathbf{Y}}$:

$$\mathbf{X}_\mathrm{post} = \mathbf{X}_\mathrm{prior} + \mathbf{K}(\mathbf{Y} - \hat{\mathbf{Y}}),$$
$$\mathbf{K} = (M-1)^{-1}\mathbf{X}_\mathrm{prior}\hat{\mathbf{Y}}^\top \left[(M-1)^{-1}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\top + \mathbf{R}\right]^{-1}. \quad 3.$$

In this matrix version of the update equation, $\mathbf{X}_\mathbf{prior}$ is the model prior matrix of size $N \times M$, which consists of all the model fields to be updated by the data assimilation (e.g., 2D fields such as surface air temperature) concatenated into a single dimension N for each M model ensemble member. $\mathbf{X}_\mathbf{post}$, the posterior matrix, has the same dimensions. $\mathbf{Y}$ is the proxy data matrix of size $P \times M$. Proxy data are usually a vector of data at a set of particular locations, so they are tiled M times to match the dimensions of $\hat{\mathbf{Y}}$. $\hat{\mathbf{Y}}$ is the proxy estimate matrix of size $P \times M$, which is computed from the model prior at the proxy locations using PSMs. The multiplication by $(M-1)^{-1}$ is to gain an unbiased estimate. The Kalman gain dimensions are $N \times P$.

Following Whitaker & Hamill (2002), the update equation is solved by decomposing it into solutions for the mean value and the deviations from the mean:

$$\overline{\mathbf{X}}_\mathrm{post} = \overline{\mathbf{X}}_\mathrm{prior} + \mathbf{K}(\mathbf{y} - \overline{\hat{\mathbf{Y}}}) \text{ and}$$
$$\mathbf{X}'_\mathrm{post} = \mathbf{X}'_\mathrm{prior} - \tilde{\mathbf{K}}\hat{\mathbf{Y}}', \quad 4.$$

where $\tilde{\mathbf{K}}$ is

$$\tilde{\mathbf{K}} = (M-1)^{-1}\mathbf{X}_\mathrm{prior}\hat{\mathbf{Y}}^\top \left[\sqrt{(M-1)^{-1}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\top + \mathbf{R}}^{\,-1}\right]^\top \left[\sqrt{(M-1)^{-1}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\top + \mathbf{R}} + \sqrt{\mathbf{R}}\right]^{-1}. \quad 5.$$

The full assimilated ensemble is then recovered through

$$\mathbf{X}_\mathrm{post} = \overline{\mathbf{X}}_\mathrm{post} + \mathbf{X}'_\mathrm{post}. \quad 6.$$

The posterior ensemble has dimensions $N \times M$, like the prior model ensemble. The matrix dimensions above are illustrated by King et al. (2023b).

2.1. The Model Prior

In ensemble paleoDA, the assimilation begins with $\mathbf{X}_\mathrm{prior}$, a prior estimate of plausible climate states for the time period of interest derived from model simulations. Ideally, the distribution of $\mathbf{X}_\mathrm{prior}$ should encapsulate all the uncertainties in our prior knowledge about the climate system during the time period of interest, including internal climate variability, model bias, forcings, and boundary conditions. If $\mathbf{X}_\mathrm{prior}$ is too narrow (insufficient variance), it constrains the data assimilation system's ability to integrate new information from the proxies: The covariance ($\mathbf{X}_\mathrm{prior}\hat{\mathbf{Y}}^\top$) is lower, leading to an underestimate of the Kalman gain, $\mathbf{K}$, and therefore a smaller update by the proxy data (Equation 3). Conversely, if the variance of $\mathbf{X}_\mathrm{prior}$ is too high, the innovation from $\mathbf{Y}$ will be overweighted. Both scenarios result in a posterior state, $\mathbf{X}_\mathrm{post}$, that does not optimally

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

blend information from the proxies and the climate model. Moreover, incorrect variance in $\mathbf{X}_{\text{prior}}$ also leads to an inaccurate estimation of the variance in $\mathbf{X}_{\text{post}}$, resulting in either overconfidence or excess uncertainty in the reconstructed climate state.

Constructing an appropriately varied $\mathbf{X}_{\text{prior}}$ for paleoDA depends on the research question and time period of interest. During the Common Era, internal climate variability and model biases are the largest sources of uncertainty in spatial and seasonal patterns of simulated climate variability (Goosse et al. 2005, Deser et al. 2012). PaleoDA reconstructions for this period typically employ a stationary $\mathbf{X}_{\text{prior}}$, consisting of the same ensemble (of annual or seasonal averages) randomly drawn from Coupled Model Intercomparison Project (CMIP) last millennium simulations for each time step (e.g., Hakim et al. 2016, Tardif et al. 2019, King et al. 2021). This method effectively samples the range of internally driven climate variability, and because the prior is identical for each time step, the proxy data drive the temporal structure of the reconstruction. If the prior consists of a single model, the reconstruction will inherit that model's covariance structures, including its biases [i.e., extent of the tropical Pacific cold tongue, double Intertropical Convergence Zone (ITCZ) (Amrhein et al. 2020)]. The use of multiple climate models within $\mathbf{X}_{\text{prior}}$ can help with this problem by capturing variance associated with different climate model physics, improving reconstruction skill (Parsons et al. 2021, Eswaran et al. 2024). Note however that using multiple models is typically an ad hoc way of representing structural model uncertainties and may underrepresent errors that are common across model architectures.

Constructing a $\mathbf{X}_{\text{prior}}$ that captures the magnitude and uncertainty of forced climate variability is essential on longer geological timescales (thousands to millions of years), over which changes in Earth's orbit, ice sheet and continental configurations, and greenhouse gas concentrations alter seasonal and spatial covariances. On these timescales, uncertainties in boundary conditions and model physics are the largest sources of variance (e.g., Kageyama et al. 2021, Thompson et al. 2022, Zhu et al. 2022). An ideal $\mathbf{X}_{\text{prior}}$ would consist of a multi-model ensemble of simulations that sample multiple possible boundary conditions, but this represents a substantial computational endeavor, and so such an ensemble rarely exists. As a result, longer timescale applications of paleoDA often rely on smaller $\mathbf{X}_{\text{prior}}$ ensembles drawn from simulations from a single climate model run with varied boundary conditions (e.g., Osman et al. 2021, Tierney et al. 2022, Judd et al. 2024) or on multi-model ensembles run with a single set of boundary conditions (e.g., Annan et al. 2022). For transient estimates of Earth's climate across fundamentally different climate states [e.g., Last Glacial Maximum (LGM) to present], constructing a $\mathbf{X}_{\text{prior}}$ that evolves through time using a running window provides a rough approximation of evolving covariance relationships [e.g., 4,000–5,000 years (Osman et al. 2021, Erb et al. 2022)] but can still result in unrealistic artifacts (e.g., simulations run with a Laurentide ice sheet are included in $\mathbf{X}_{\text{prior}}$ during the mid-Holocene). Another limitation of paleoDA on longer timescales is that the time step represented by the model prior and the proxy data are rarely equivalent. For example, reconstructing average LGM climate involves proxy data that span 4,000 years of time (23–19 kyr), yet the model priors are not run for 4,000 years (Tierney et al. 2020b). As a compromise, climatologies computed over 50–100 years can be used as model prior states, under the assumption that multidecadal-to-centennial climate variability in the model is equivalent to millennial-scale variability (Tierney et al. 2020b, Osman et al. 2021). As the diversity, number, and length of paleoclimate simulations continue to grow, constructing varied and representative $\mathbf{X}_{\text{prior}}$ for paleoDA on longer geological timescales will become more feasible.

2.2. The Proxy Data and Their Uncertainty

Broadly speaking, data assimilation was designed to incorporate observational information into models in order to improve simulations and forecasts. In paleoclimate applications, the purpose

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

of DA is to incorporate proxy information to improve our reconstruction of past climate over what model simulations can achieve alone. One of the advantages of the paleoDA framework is that, in theory, the proxy data $\mathbf{Y}$ enter into reconstruction as is, while the understanding of how the proxy data represent climate information is encoded in $\hat{\mathbf{Y}}$ (see the next section). In practice, this might not always be the case if certain proxy uncertainties cannot be easily included in $\hat{\mathbf{Y}}$. One of these is time uncertainty, i.e., the uncertainty in the date or span of time represented by the proxies. Aside from annually resolved records from tree rings (and other cross-dated archives with annual bands), all proxy archives have time uncertainty associated with the method of dating used, ranging from a few years for layer-counted archives such as corals and varved sediments to decades to centuries for radiocarbon-dated sediments and 1,000–100,000 years or more on longer geological timescales. If this uncertainty is larger than the target time step of the paleoDA, it should be accounted for in the process, ideally as part of the archive-level component of a PSM. However, in a transient reconstruction, the dating of $\mathbf{Y}$ might affect which time step of the DA a proxy appears in, in which case the time uncertainty has to be applied to both $\mathbf{Y}$ and $\hat{\mathbf{Y}}$. A Monte Carlo sampling process can be used to sample age model uncertainty and then conduct an ensemble of paleoDAs with the elements of $\mathbf{Y}$ falling into various different time steps accordingly (Osman et al. 2021).

The proxy uncertainty appears in the Kalman gain as the covariance matrix $\mathbf{R}$ and determines how much weight to ascribe to any given proxy estimate (Equation 3). Proxy data with high $\mathbf{R}$ values have a smaller Kalman gain and thus have less influence on the reconstruction than those with low values. In classical weather and climate data assimilation, $\mathbf{R}$ is the observational error, i.e., the error associated with the measurements. In paleoDA, laboratory measurement error of a proxy is one, but not the only, component of $\mathbf{R}$. $\mathbf{R}$ also includes random uncertainties that a proxy might experience in the natural environment. Ideally, structural uncertainties such as those associated with bioturbation or diagenesis can be encoded in $\hat{\mathbf{Y}}$, but if this is not possible, then these sources of uncertainty can be represented in $\mathbf{R}$ as well.

While measurement error is easy to quantify, the actual environmental error of a proxy is often poorly constrained. As an example, consider a record of alkenone $U^{K'}_{37}$ (a proxy for SST) measured in a marine sediment core, at a single location, that spans the LGM to present. Suppose the record indicates that the glacial time was 1.5°C cooler than the present day. The laboratory uncertainty of the record is known to be 0.006 $U^{K'}_{37}$ units (1σ), which translates to about 0.2°C. The error on the calibration model (used in $\hat{\mathbf{Y}}$) for $U^{K'}_{37}$ to SST is 1.4°C (1σ), but if this were the error on the downcore record, the entirety of the glacial/interglacial structure would be subsumed by it. Clearly, the true error lies between these extremes, but what is it?

To determine a true $\mathbf{R}$ experimentally, one would need to measure $U^{K'}_{37}$ in multiple sediment cores from approximately the same location (and also know a good deal about the depositional process of the alkenone compounds at that location). This is not commonly done, so another method is needed to inform the choice for $\mathbf{R}$. One approach is to experiment with different values within a feasible prior range and do either internal or external validation tests to find a value that performs the best (Tierney et al. 2020b, Osman et al. 2021). However, this method is indirect rather than process based, and therefore it is not guaranteed to yield a realistic value. In applications where proxy data overlap with historical climate data, $\mathbf{R}$ might be estimated through comparison of real and PSM-simulated proxy data (King et al. 2021). In all paleoDA applications, ambiguity in setting $\mathbf{R}$ remains a challenge and a clear area for improvement alongside the development of more accurate PSMs, as discussed next.

2.3. Proxy Estimation

The blending together of proxy records and model simulations in paleoDA requires the conversion of model-derived climate variables into equivalent proxy-based units ($\hat{\mathbf{Y}}$). This is done through the

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

use of PSMs (see the sidebar titled Proxy System Models) for each respective proxy type. PSMs take climate variables [e.g., SST, sea surface salinity, the oxygen isotopic composition of seawater (δ^{18}O)] from the model prior state ensemble ($\mathbf{X}_{\text{prior}}$) and translate these into proxy units that can be compared directly with the observed proxy values in the innovation calculation (see Equation 3).

An ideal PSM captures all processes that control the proxy signature, including the preservation of the proxy response in natural archives (see the sidebar titled Proxy System Models). Practically speaking, not all of these processes are well-constrained, and some of them can be difficult to formulate into mathematical expressions. As a consequence, PSMs range in complexity from relatively simple regressions (e.g., Thompson et al. 2011, Tierney & Tingley 2018) to multivariate, layered process models (e.g., Treble et al. 2019) based on the level of available knowledge. Depending on the timescale and the location of the proxy, greater complexity is not always better, either because required model parameters might be unknown and/or because small-scale uncertainties are swamped by larger climatic signals (Hu et al. 2021).

Relative to conventional inverse models, wherein climate variables are statistically inferred from proxies, PSMs offer a clear advantage in that they can account for multiple climatic influences and nonstationarity in proxy responses, as well as provide straightforward error propagation. For example, rather than having to assume a fixed season of production for planktic foraminifera (that encode the SST proxies δ^{18}O or Mg/Ca), a PSM can predict a dynamically changing seasonality based on when modeled monthly SST suggests favorable conditions for each species (Malevich et al. 2019, Tierney et al. 2019b). Similarly, whereas the changing relative influence of multiple climatic or environmental influences (e.g., temperature versus rainfall on tree ring width) on a proxy through time can complicate inverse frameworks, PSMs can more naturally accommodate nonlinear shifts in proxy sensitivity in order to generate a plausible range of $\hat{\mathbf{Y}}$ values (e.g., Tolwinski-Ward et al. 2011).

Models that incorporate tracers that help close the gap between climate metrics and proxy units greatly simplify both the formulation and application of PSMs. In particular, simulations that include the stable isotopologues of water (e.g., $^{1}H_2{}^{16}O$, $^{1}H^{2}H^{16}O$, $^{1}H_2{}^{18}O$) in all parts of the water cycle are very useful for paleoDA. Many paleoclimate proxies encode water isotope changes (e.g., speleothem and foraminiferal calcite δ^{18}O, leaf wax δ^{2}H), but water isotopes in and of themselves are complex tracers of multiple physical processes in the climate system (Bowen et al. 2019). These processes are arguably best represented by the complexity of the climate model, rather than computed as part of a PSM offline. In addition, in the paleoDA context, an update in isotope space propagates accordingly to all of the underlying related climatic fields (e.g., temperature, humidity, rainfall source, convection, winds), providing a powerful means of inferring climate variables for which we have no direct proxies (but that strongly influence isotopic changes).

Currently, only a limited number of modeling centers run climate model simulations with water isotopes. This severely restricts the number of models and simulations available to include in an isotope-enabled $\mathbf{X}_{\text{prior}}$, which more often than not is desirable given the quantity of proxies that encode water isotopes (Tierney et al. 2020b, Osman et al. 2021, Judd et al. 2022). More widespread incorporation of water isotope tracers across different model families would be beneficial for paleoDA applications, as well as CFRs more generally.

2.4. The Kalman Gain and Localization

As described at the beginning of Section 2, the Kalman gain ($\mathbf{K}$) both spreads and weights the new information coming from the assimilation of the proxies ($\mathbf{Y} - \hat{\mathbf{Y}}$; the innovation) as it is added to the prior state. The weighting is based on the proxy error $\mathbf{R}$ (Section 2.2) and the variance of the estimates, $\hat{\mathbf{Y}}$—more uncertain proxies or proxy estimates impart a weaker update. The spreading

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

comes from the first term of the Kalman gain, which is a covariance calculation between $\hat{\mathbf{Y}}$ and the prior state ($\mathbf{X}_{\text{prior}}$). Locations and climate variables in $\mathbf{X}_{\text{prior}}$ that have strong covariance with the proxy estimate get updated more strongly than those for which this is not the case. As a practical example, consider an SST proxy in the eastern equatorial Pacific. On the interannual timescale, this proxy will have a strong covariance with SSTs in the western Pacific warm pool due to the El Niño–Southern Oscillation (ENSO), as well as remote covariances with SST off the California coast. Addition of proxy information at this location will therefore update these regions with high covariance more strongly than places that are less influenced by El Niño (e.g., the North Atlantic).

While there are modes of climate variability (such as El Niño) that connect far-flung regions of the globe, offline ensembles may have spurious covariance patterns that can introduce errors into the reconstruction. Covariance localization is a technique applied in DA that limits the spatial extent of the update provided by any given observation in order to mitigate these errors. In the context of the joint update equations described above, localization applies a set of weights to the Kalman gain, as follows (Hamill et al. 2001):

$$\mathbf{K} = \mathbf{W}_{\text{loc}} \circ \left[(M-1)^{-1}\mathbf{X}'_{\text{prior}}\hat{\mathbf{Y}}'^{\top}\right] \times \left[\mathbf{Y}_{\text{loc}} \circ \left[(M-1)^{-1}\hat{\mathbf{Y}}'\hat{\mathbf{Y}}'^{\top}\right] + \mathbf{R}\right]^{-1}. \tag{7}$$

$\mathbf{W}_{\text{loc}}$ contains the localization weights of each N state vector element relative to each proxy estimate in $\hat{\mathbf{Y}}$ and is of dimension $N \times P$. $\mathbf{Y}_{\text{loc}}$ contains the localization weights of each proxy estimate in $\hat{\mathbf{Y}}$ relative to the other estimates in $\hat{\mathbf{Y}}$ and thus is of dimension $P \times P$. M is the number of ensemble members, and multiplication by $(M-1)^{-1}$ is applied to obtain an unbiased estimate; $\circ$ denotes element-wise multiplication. The weights define the shape and extent of decorrelation. A commonly used functional form for the weights is the Gaspari-Cohn fifth-order polynomial (Gaspari & Cohn 1999), which produces a Gaussian-like isotropic region of decorrelation away from each proxy location. The cutoff radius assigned to the polynomial is the distance outside of which all covariance is eliminated.

The localization cutoff radius is a user choice, and the appropriate distance varies depending on the prior ensemble, the target fields in the analysis, the proxy network, and the origin of covariance errors. In modern DA applications, localization primarily serves to mitigate covariance sampling errors arising from small numbers of prior ensemble members. By contrast, in offline paleoDA, where it is easier to have a large prior ensemble, the primary benefit is that it mitigates covariance bias in the model prior (Amrhein et al. 2020, Parsons et al. 2021). Because Gaspari-Cohn localization imparts isotropic spatial covariance on the solution, sparse proxy networks may require longer cutoff radii (or no localization) to avoid the appearance of undesirable circular blotches in the update (**Figure 2**). Typically, paleoDA applications have used longer localization radii [e.g., 12,000–25,000 km (Tardif et al. 2019, Tierney et al. 2020b, Osman et al. 2021)], although some have used radii as small as 5,000 km (Annan et al. 2022, Masoum et al. 2024). Validation exercises (e.g., King et al. 2021) provide one approach for determining optimal localization radii for a given problem, as described in the next section.

2.5. Methods of Validation

Validation is a key component of any type of CFR. It acts as a reality check on whether the reconstruction is yielding reasonable results, provides information beyond a naive baseline, and can identify statistical overfitting. In the case of paleoDA, it is also an efficient way to evaluate whether user-defined choices such as $\mathbf{R}$ and the localization radius are appropriate.

Internal validation typically involves iteratively leaving out one or more [some applications with dense proxy networks have left out 25% (Tardif et al. 2019, Tierney et al. 2020b)] of the

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

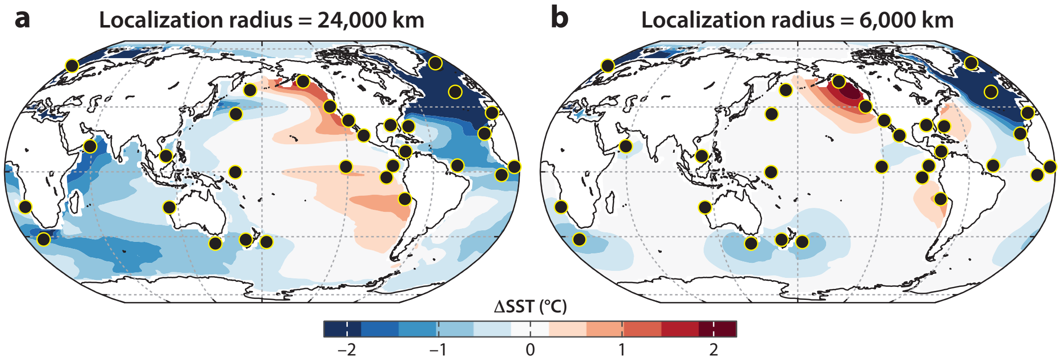

Figure 2

Example of the impact of the localization cut-off radius on the update of a climate field, with a limited proxy network ($N = 30$). (*a*) The change in SST (relative to the model prior mean) from the assimilation of SST proxy data (*black points*) with the localization radius set to 24,000 km. Note how the updates propagate along dynamical features in the SST field. (*b*) As in panel *a* but with the localization radius set to 6,000 km. Updates are more isotropic, with reduced spatial extent, following the Gaspari-Cohn function used to define the localization weights. Proxy data are a decimated version of the network used by Tierney et al. (2025), assimilated with a multi-model prior of mid-Pliocene simulations (Tierney et al. 2025). Abbreviation: SST, sea surface temperature.

proxy records from the assimilation and then predicting the omitted proxy data from the resulting posterior (effectively a $\hat{\mathbf{Y}}$ calculation). Skill is assessed through metrics such as R^2, the coefficient of efficiency (Nash & Sutcliffe 1970), or root mean square error. Internal validation provides an assessment of the internal predictability of the reconstruction and can inform baseline choices for $\mathbf{R}$ and localization radius. However, as a relatively easy test, it may not conclusively distinguish between different paleoDA parameter choices.

External validation is a much more powerful assessment of the skill of the reconstruction because it tests how well the paleoDA posterior can predict an external (nonassimilated) proxy target. Good targets for external validations are proxy systems that were not included in the assimilation and/or proxies that are sensitive to a different aspect of climate than those used in the assimilated proxy network. For example, both the LGM reconstruction of Tierney et al. (2020b) and the LGM-to-present reconstruction of Osman et al. (2021) used water isotope–enabled model priors, so they chose to validate their reconstructions against speleothem and ice core–derived $\delta^{18}O$ of precipitation ($\delta^{18}O_p$). Because these studies only assimilated SST proxies and the $\delta^{18}O_p$ proxies are on land, this is a difficult validation target. Nonetheless, both studies demonstrated that data assimilation greatly improved prediction of $\delta^{18}O_p$ over the model prior, especially for the ice cores (Tierney et al. 2020b, Osman et al. 2021). External $\delta^{18}O_p$ validation also helped refine choices for $\mathbf{R}$ and localization; in the case of the LGM reconstruction of Tierney et al. (2020b), validation improved when slightly lower values of $\mathbf{R}$ were used for the proxies than suggested by the PSMs, indicating that perhaps the calibration models in the PSMs overestimate the error variance.

Another way to evaluate paleoDA's capability to reconstruct climate fields is to use a pseudoproxy network (e.g., Steiger et al. 2014, Brennan & Hakim 2022, King et al. 2023a). In these idealized experiments, a sparse array of pseudoproxy data (drawn from a climate model and then run through a PSM) is used in the paleoDA scheme. The skill of the resulting reconstruction is then assessed by comparing it with the climate model simulation used to generate the pseudoproxies. In a perfect model experiment, both the pseudoproxies ($\mathbf{Y}$) and the prior ensemble ($\mathbf{X}_{\text{prior}}$) are drawn from the same climate model simulation, providing an upper bound on reconstruction performance for a given proxy network. In imperfect model experiments, pseudoproxies are drawn from one model and the $\mathbf{X}_{\text{prior}}$ from a different model. Imperfect model experiments

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

are a stricter test of the method and more closely resemble real paleoDA reconstructions, which rely on covariance estimates from model simulations that imperfectly represent Earth's climate.

3. EXAMPLE APPLICATIONS

In this section, we review recent applications of paleoDA over different timescales and highlight the advances in our understanding of the climate system the paleoDA method has afforded. A few themes tie these examples together: (*a*) Diverse proxy types can be combined, (*b*) multiple climate fields can be recovered, and (*c*) the full-field nature of the reconstruction enables more accurate estimates of first-order metrics of climate change, such as global mean surface temperature (GMST) and ECS, defined as the GMST change in response to a doubling of CO_2 concentrations once fast feedbacks have occurred. A distinction among these examples that illustrates the flexibility of offline DA is that Common Era reconstructions target transient variability (both forced and internal), whereas deeper-time applications typically use time-slice approaches that target long-term time means.

3.1. Common Era Applications

The first and most common use of paleoDA has been to reconstruct the climate of the Common Era. Goosse et al. (2010) used a particle filter paleoDA approach, assimilating a traditional gridded reconstruction (Mann et al. 2008) as if it were temperature observations. Since this first paleoDA application, a number of Common Era reconstructions have been created on annual, seasonal, and monthly timescales (e.g., Hakim et al. 2016, Franke et al. 2017, Steiger et al. 2018, King et al. 2021). Commonly reconstructed climate variables include near-surface air temperature, geopotential height, and hydroclimate indices such as the Palmer Drought Severity Index (PDSI). This range of variables makes it possible to assess the dynamical causes of climate phenomena throughout the Common Era. Below, we briefly highlight two representative examples.

King et al. (2023a) used offline EnKF paleoDA to reconstruct the Southern Annular Mode (SAM) over the Common Era at annual resolution, extending prior DA work reconstructing the SAM during the industrial era (O'Connor et al. 2021, Dalaiden et al. 2021). In this application, the reconstruction target was an index of climate rather than a spatial field: the austral summer SAM index $P^*_{40^\circ S} - P^*_{65^\circ S}$ (Gong & Wang 1999), where P^*_X represents the normalized zonal mean sea level pressure at X latitude, averaged from December to February. The model prior consisted of the austral summer SAM index drawn from last millennium simulations conducted with four different models [Community Climate System Model 4 (CCSM4), Community Earth System Model (CESM) Last Millennium Ensemble (LME), Mac Planck Institute (MPI), and Meteorlogical Research Institute (MRI)], so that different physical representations of the SAM were sampled. Multiple types of proxies were assimilated, including temperature-sensitive tree rings, corals, and ice cores, as well as hydroclimate-sensitive drought atlas gridpoints (**Figure 3*a***). The computation of the posterior followed the same mathematics described in this review (Equation 3).

The King et al. (2023a) study found that the modern multidecadal positive trend in the SAM (indicating movement toward Antarctica) is significantly outside the range of the past 2,000 years, supporting the conclusion that this recent behavior is a response to anthropogenic climate change. The paleoDA approach was advantageous for this reconstruction target for several reasons. First, the DA framework allowed the assimilation of a range of diverse proxy types. Additionally, the use of paleoDA circumvented the need to calibrate the reconstruction directly to an instrumental SAM record. Prior to this work, all multi-century SAM reconstructions relied on such calibrations, which implicitly assume the stationarity of the SAM's teleconnections over the last few decades (e.g., Dätwyler et al. 2018). Various studies cast doubt on the validity of this assumption

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

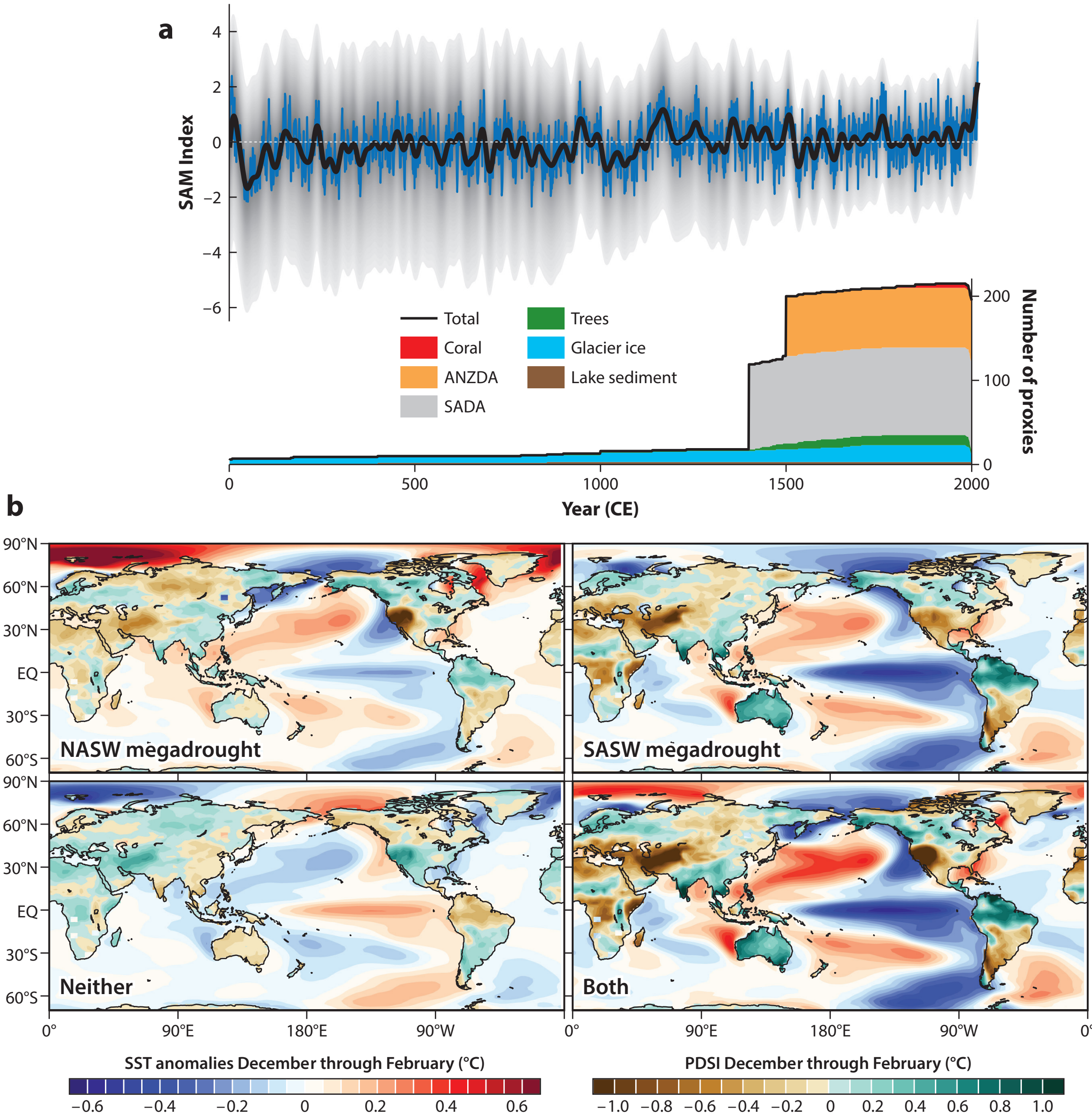

Figure 3

Examples of Common Era applications of paleoDA. (*a*) Reconstruction of the SAM Index and the proxy network used in the data assimilation. Panel adapted from King et al. (2023a). (*b*) Mean climate states during megadrought and non-megadrought years (Steiger et al. 2021) from composites of DJF SST and PDSI from PHYDA (Steiger et al. 2018). Individual panels show all years corresponding to NASW megadrought conditions, SASW megadrought conditions, when megadroughts do not exist in either location (neither) and when megadroughts exist in both locations (both). Temperature and PDSI data are anomalies with respect to the analysis period 1000–1925 CE. Abbreviations: ANZDA, Australia and New Zealand Drought Atlas; DJF, December–January–February; NASW, North American Southwest; paleoDA, paleoclimate data assimilation; PDSI, Palmer Drought Severity Index; PHYDA, Paleo Hydrodynamics Data Assimilation; SADA, South American Drought Atlas; SAM, Southern Annual Mode; SASW, South American Southwest; SST, sea surface temperature.

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

(Silvestri & Vera 2009, Gallant et al. 2013), and the nonstationarity of these teleconnections remains a major source of uncertainty in non-DA reconstructions. By contrast, the approach of King et al. (2023a) only requires the stationarity of proxy response to their local climate variables, a more conservative assumption that is typical of nearly all paleoclimate reconstructions. Furthermore, the climate covariances derived for King et al. (2023a) rely on hundreds of years of climate model output, reducing the potential sensitivity to anomalous decadal- or centennial-scale variability. The reconstruction also made use of optimal sensors embedded in the paleoDA framework (see Section 3.4). Finally, King et al. (2023a) also developed a method to address spurious variance reduction due to declining proxy data back through time (see Section 4.2).

Another powerful use of paleoDA is to reconstruct the climatic conditions associated with hydroclimate extremes. Prior to the use of paleoDA, hydroclimate reconstructions have been limited to single-variable reconstructions, particularly PDSI. Interpreting these reconstructions required analyzing independent model simulations that may or may not capture the features seen in the reconstructions (e.g., Cook et al. 2007). But with paleoDA, both hydroclimate variables and climate-dynamical variables can be reconstructed simultaneously. The dynamical variables can then be used to directly diagnose the cause of hydroclimate extremes. The Paleo Hydrodynamics Data Assimilation (PHYDA) (Steiger et al. 2018) is one example of a paleoDA product that simultaneously reconstructed both drought and temperature fields alongside indices of North Atlantic variability, ITCZ location, and ENSO. This reconstruction used the CESM LME (Otto-Bliesner et al. 2016) as its model prior and assimilated multiple types of proxies (tree rings, corals, ice cores, speleothems, and lake sediments) using the offline EnKF approach described in this review. Because the reconstructed fields and indices are all related through the climate dynamics of the CESM model, PHYDA can be used to diagnose the drivers of patterns of temperature or hydroclimate change. For example, Steiger et al. (2021) investigated the ocean dynamics associated with severe megadroughts in both the North American Southwest and the South American Southwest (**Figure 3*b***). They found that there are a similar number of megadroughts in both locations and that they occur simultaneously more often than would be expected by chance. These coupled megadroughts were associated with strong and frequent La Niña conditions (**Figure 3*b***).

3.2. The Last Glacial Maximum and Holocene

The LGM (ca. 23–19 ka BP) and the subsequent Holocene epoch have long served as benchmarks for understanding the Earth system (CLIMAP Proj. Memb. 1976). The climate signals are large, and changes in external forcings and boundary conditions (e.g., ice sheets, greenhouse gas concentrations, orbital configuration) are well-understood. There is also an abundance of both proxy data (Kaufman et al. 2020, Tierney et al. 2020b) and model simulations (Brierley et al. 2020, Kageyama et al. 2021, Snoll et al. 2024). For these reasons, this time period is well-suited for paleoDA climate reconstruction.

Several past efforts have tackled the problem of optimally combining modeled estimates of the LGM state with proxy observations, some of which have used paleoDA techniques (Gebbie & Huybers 2006, Annan & Hargreaves 2013, Amrhein et al. 2018). Tierney et al. (2020b) revisited the LGM reconstruction problem with an updated database of SST proxies and, in an advance over previous work, incorporated Bayesian PSMs into the offline EnKF framework. This allowed for the propagation of proxy uncertainties and seasonal biases into the posterior solution. This new paleoDA reconstruction was used to refine our understanding of both how cold LGM GMST was relative to the late Holocene (−6.1°C, 95% CI = −6.5 to −5.7°C) (**Figure 4**) and what ECS is, based on the LGM (3.4°C, 95% CI = 2.4 to 4.5°C). The latter was a revision upward from older estimates (e.g., Schmittner et al. 2011) and is more consistent with the modern consensus range

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

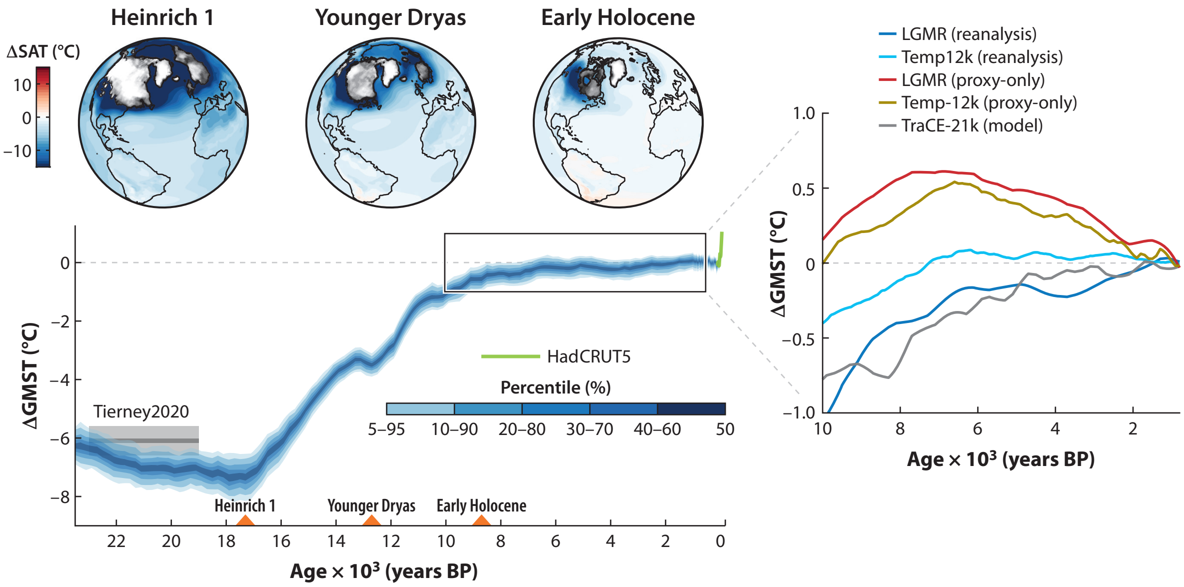

Figure 4

Reconstruction of surface temperatures since the LGM featuring the LGMR (Osman et al. 2021). GMST changes are shown as anomalies relative to the last two millennia of the Holocene. Yellow triangles indicate the respective timings of the three spatial patterns shown at top. Tierney2020 denotes the assimilated ΔGMST range from Tierney et al. (2020b) spanning 23–19 ka. (*inset*) The time series, at right, have 1,000-year smoothing applied for intercomparison; Temp12k (proxy-only) is from Kaufman et al. (2020), Temp12k (reanalysis) is from Erb et al. (2022), and the transient LGM-to-present model simulation TraCE-21k is from Liu et al. (2014). Figure adapted from figures 2 and 4 of Osman et al. (2021). Abbreviations: GMST, global mean surface temperature; HadCRUT5, Met Office Hadley Centre/Climatic Research Unit global surface temperature anomalies, version 5; LGM, Last Glacial Maximum; LGMR, Last Glacial Maximum Reanalysis; SAT, surface air temperature; Temp12k, Temperature 12k; TraCE-21k, Transient Climate Evolution of the last 21,000 years.

(Sherwood et al. 2020). It directly contributed to a narrower ECS assessment in the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) [compared to the IPCC AR5 (Forster et al. 2021)], highlighting how paleoDA has provided constraints on a key climate metric.

Osman et al. (2021) extended the results of the study by Tierney et al. (2020b) to reconstruct surface temperature changes from the LGM to the present (**Figure 4**). This study used similar paleoDA methodologies but with some new innovations including a Monte Carlo method to account for age uncertainty and an evolving prior to accommodate shifts in climatic states and associated model boundary conditions spanning the last glacial to interglacial transition. The resulting reconstruction, the Last Glacial Maximum Reanalysis (LGMR), allowed for investigation of the spatial imprint of the main drivers of LGM-to-present climate (ice sheets, greenhouse gases, oceanic thermohaline circulation, and seasonal insolation). Because the LGMR provides a robust estimate of GMST, Osman et al. (2021) were able to show that the rate and magnitude of modern anthropogenically driven warming are highly unusual in the context of the past 24 kyr.

PaleoDA reconstructions of the Holocene, including the LGMR and the Erb et al. (2022) Temp12K reanalysis, have advanced understanding of the Holocene temperature conundrum (Liu et al. 2014), which refers to the discrepancy between mid- to late-Holocene global temperature trends from proxy-based reconstructions (which show a cooling) (Marcott et al. 2013, Kaufman et al. 2020) and transient climate model simulations (which show a warming) (Liu et al. 2014,

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

Erb et al. 2022). The warming trend through the Holocene produced by the models is expected based on the changes in forcing agents (namely, a decline in land ice cover plus an increase in greenhouse gases), which means the conundrum represents either a misunderstanding of proxy seasonality or a missing forcing in the model simulations (Liu et al. 2014).

By taking a latitudinally weighted average of the SST proxy records used in the LGMR, Osman et al. (2021) closely reproduced an older proxy-only reconstruction showing Holocene cooling (Marcott et al. 2013) (**Figure 4**, *inset*). Similarly, a proxy-only reconstruction that combines terrestrial and marine proxies (Temp12K) shows a Holocene thermal maximum near 6.5 ka with a cooling trend thereafter (Kaufman et al. 2020) (**Figure 4**, *inset*). However, when these same proxy datasets are used in paleoDA, different trends emerge. In the LGMR case, the long-term cooling becomes a small warming (**Figure 4**, *inset*) that resembles the Transient Climate Evolution of the last 21,000 years (TraCE-21ka) model simulation (Liu et al. 2014) (**Figure 4**, *inset*). In the Temperature 12k (Temp12K) case, the assimilated version shows no change in GMST since 6.5 ka (Erb et al. 2022)(**Figure 4**, *inset*). In both cases, paleoDA moves the proxy-only solution closer to the TraCE-21ka trajectory. For the Erb et al. (2022) reconstruction, this is expected because the model priors include TraCE-21ka and a HadCM3 transient simulation, both of which produce a Holocene warming trend. It is less expected for the LGMR because the model prior includes simulations that have both a colder and a warmer mid-Holocene [the latter due to imposing a Green Sahara (cf. Thompson et al. 2022)]. Therefore, Osman et al. (2021) concluded that the LGMR solution reflects the assimilation process: Unlike proxy-only reconstructions that rely on latitude-weighted zonal averaging, paleoDA dynamically weights each proxy based on its uncertainties and covariance with model priors, and also explicitly accounts for proxy seasonality. Although discussion around the conundrum continues (Bova et al. 2021, Thompson et al. 2022, Kaufman & Broadman 2023, Essell et al. 2024), the results of Osman et al. (2021), Erb et al. (2022), and others (Masoum et al. 2024) highlight the utility of paleoDA in contributing to an increasingly consistent picture of LGM-to-present climate change.

3.3. Deep Time

Recent work has extended the application of paleoDA into deep geologic time, including exploring spatial patterns in temperature and precipitation just prior to and during the Paleocene–Eocene Thermal Maximum (56 Ma) (Tierney et al. 2022) and reconstructing a 485-million-year history of Earth's GMST (**Figure 5*a***) (Judd et al. 2024). Long-term records of GMST are rare and have historically been based on either climate model simulations alone (Valdes et al. 2021) or analog approaches based on modern climatic zones (Scotese et al. 2021). One of the challenges of deep-time climate reconstruction is the decreasing availability of proxy data with increasing geologic age (Judd et al. 2022), which effectively precludes reliable data-only spatial reconstruction. PaleoDA thus offers a tractable solution that can leverage both model simulations and data, and advance our understanding of long-term climate change.

The Phanerozoic Data Assimilation, or PhanDA, highlights several advantages and limits of applying paleoDA in deep geological time. PhanDA combined 872 HadCM3L simulations (Valdes et al. 2021, Judd et al. 2024) and a database of over 155,000 SST proxy values (Judd et al. 2022) to reconstruct GMST at 85 time steps spanning nearly half a billion years (**Figure 5*a***). The assimilation of the proxy data substantially increased the temperature range of the posterior when compared to the model prior, primarily by increasing GMST during greenhouse climate intervals (Judd et al. 2024). These results support an emerging consensus that the warmest climates in Phanerozoic Earth history had a GMST of ca. 30–35°C (Inglis et al. 2020, Tierney et al. 2022). Unlike model-only GMST reconstructions, PhanDA is independent from the climate sensitivity

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

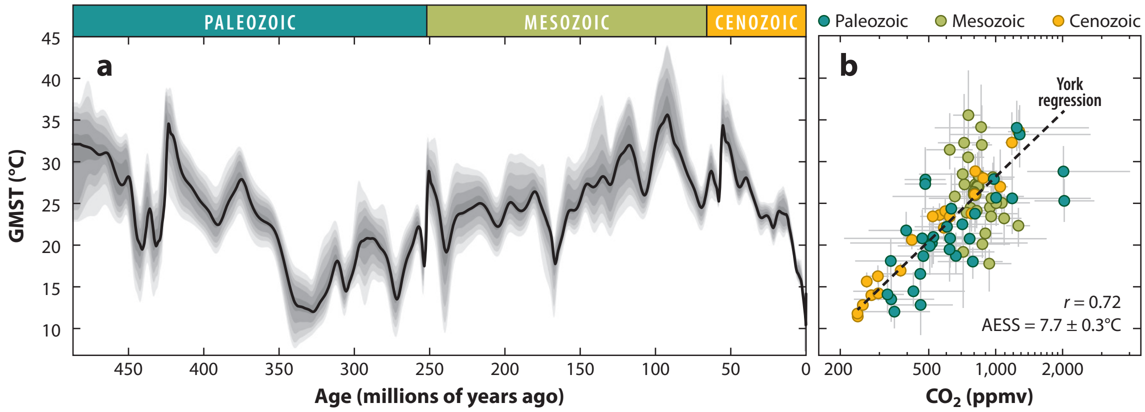

Figure 5

(*a*) Reconstructed GMST over the last 485 Myr using paleoDA (Judd et al. 2024). (*b*) Relationship between atmospheric CO_2 concentrations and reconstructed GMST over the same time interval ($r = 0.72$, $p < 0.01$), colored by geologic era. The dashed line shows the York regression (York 1968) through the data, which yields a slope, or AESS of 7.7°C per doubling of CO_2. Figure adapted from Judd et al. (2024). Abbreviations: AESS, apparent Earth system sensitivity; GMST, global mean surface temperature; paleoDA, paleoclimate data assimilation.

of the prior, enabling analysis of the long-term relationship between PhanDA GMST and the CO_2 record. CO_2 and GMST are strongly correlated across the Phanerozoic ($r = 0.72$, $p < 0.01$), particularly during the Cenozoic and Paleozoic eras (**Figure 5*b***), indicating that CO_2 has been a dominant driver of climate across the last half-billion years. The slope of this relationship (when CO_2 is in $\log_2$ space) implies a change in GMST of 7.7°C ± 0.3°C per doubling of CO_2—a metric that the authors term the apparent Earth system sensitivity (AESS). While AESS is not directly comparable with conventional estimates of ECS, it agrees well with CO_2-only Earth system sensitivity estimates from the Cenozoic (e.g., Cenozoic CO_2 Proxy Integ. Proj. (CenCO2PIP) Consort. et al. 2023). The consistency of this relationship is surprising because solar luminosity was lower during the Paleozoic (theoretically requiring higher CO_2 levels to offset it), providing a clear direction for future research.

One of the difficulties PhanDA highlights is the consequence of having a very wide prior. Unlike paleoDA applications in shallow time, where factors such as CO_2 concentration and ice sheets are better constrained and $\mathbf{X}_{\text{prior}}$ can therefore be tailored within a more narrow range of realistic conditions, there are large uncertainties surrounding boundary conditions in deep time. A wide prior is therefore required to capture different possible states, but combining very different prior states in the offline DA set up (i.e., with different ice sheet configurations or latitudinal SST gradients) can yield aphysical posteriors, such as solutions that have unrealistically steep temperature gradients between the middle latitudes and the poles (Judd et al. 2024). To address this problem, PhanDA was created by iteratively assimilating each time step, each time systematically drawing from different subsets of the prior ensemble, with a screening protocol applied to filter out dynamically infeasible results.

3.4. Optimal Sensors

In addition to reconstructing past climates, the Kalman filter can be adapted into an optimal sensor framework. In paleoclimatology, optimal sensors are typically used to assess the proxy records needed to skillfully reconstruct a climate metric and to prioritize locations for future proxy development (Evans et al. 1998, Comboul et al. 2015). The optimal sensor method can also be used to

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

assess the influence of individual proxy records on a paleoDA reconstruction, providing a level of transparency not always possible for other CFR methodologies.

The paleoDA optimal sensor examines the relationship between sets of proxy records and a climate metric, such as spatial mean temperature or a climate mode index. Specifically, the sensor quantifies the ability of a set of proxy records to reduce variance in the climate metric's posterior ensemble. This assessment is applied to multiple sets of proxy records, allowing the sets to be ranked by their ability to reduce uncertainty. The analysis follows the equation

$$\Delta\boldsymbol{\sigma}_k = (\mathbf{J}\hat{\mathbf{Y}}_k^{\top})^2[\hat{\mathbf{Y}}_k\hat{\mathbf{Y}}_k^{\top} + \mathbf{R}_k]^{-1}. \qquad 8.$$

Note the similarity to the equation for the Kalman gain (Equation 3). Here, $\mathbf{J}$ is the climate metric (replacing $\mathbf{X}_{\text{prior}}$). The k subscripts indicate values for a kth set of proxy records. $\hat{\mathbf{Y}}_k$ can be estimates for a single proxy record or a set of multiple proxies. As a rule, any proxies with covarying error statistics should be grouped into the same set. Note that the optimal sensor requires only proxy estimates; it does not require actual proxy records. As such, the analysis may be applied to extant proxy records and potential future sites alike. Note that the variance reduction computed by the optimal sensor represents the maximum potential reduction. In practice, when a proxy set is added to an existing network, the additional variance reduction will be smaller than this maximum potential. This occurs because most proxy records covary with one another, in part because of their shared sensitivity to past climate, and this covariance is down-weighted by the Kalman gain to avoid double-counting the same signal.

King et al. (2023a) applied this optimal framework to their SAM index reconstruction (see Section 3.1) to examine how different proxy records influence the reconstruction over time. To identify the records with the largest contribution to the reconstruction, the authors first assessed the potential variance reduction for each proxy in the network (**Figure 6**). Because most records were not available over the full reconstruction period, the authors applied the optimal sensor to the available records for each reconstructed time step. They found that two tree-ring records from Tasmania and New Zealand (Plateau Remote and Mt. Read) were the most influential during the early part of the reconstruction, whereas the drought atlases South American Drought Atlas and Australia and New Zealand Drought Atlas dominated during the later interval when they became available (**Figure 6**). Ultimately, the optimal sensor analysis in this study complements the assimilation by providing insight into the controls on the reconstruction's behavior and the relative contributions from each proxy record.

4. CHALLENGES AND LIMITATIONS

PaleoDA is a powerful technique that has enabled discovery of Earth's climate system across multiple timescales. However, like all CFR methods, it has drawbacks. Here we highlight two primary issues that have challenged offline EnKF paleoDA reconstructions, but other limitations also exist, such as extremely sparse proxy networks (see Section 3.3) or the fact that the user must choose (sometimes in an ad hoc manner) how to define the proxy error (**R**) (Section 2.2) and the localization radius (Section 2.4).

4.1. Dependence on the Model Prior

PaleoDA explicitly assumes that the simulated covariance patterns of the model prior are the covariance patterns of the real climate system (**K** in Equation 3) that generated the proxy records. This is not a safe assumption, however, because most models have substantial biases when compared to historical observations (C. Wang et al. 2014). In offline DA, the covariance structures are entirely inherited from the model prior, so any biases present in the model are present in the reconstruction. This adds substantial structural uncertainty to paleoDA reconstructions, which can

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

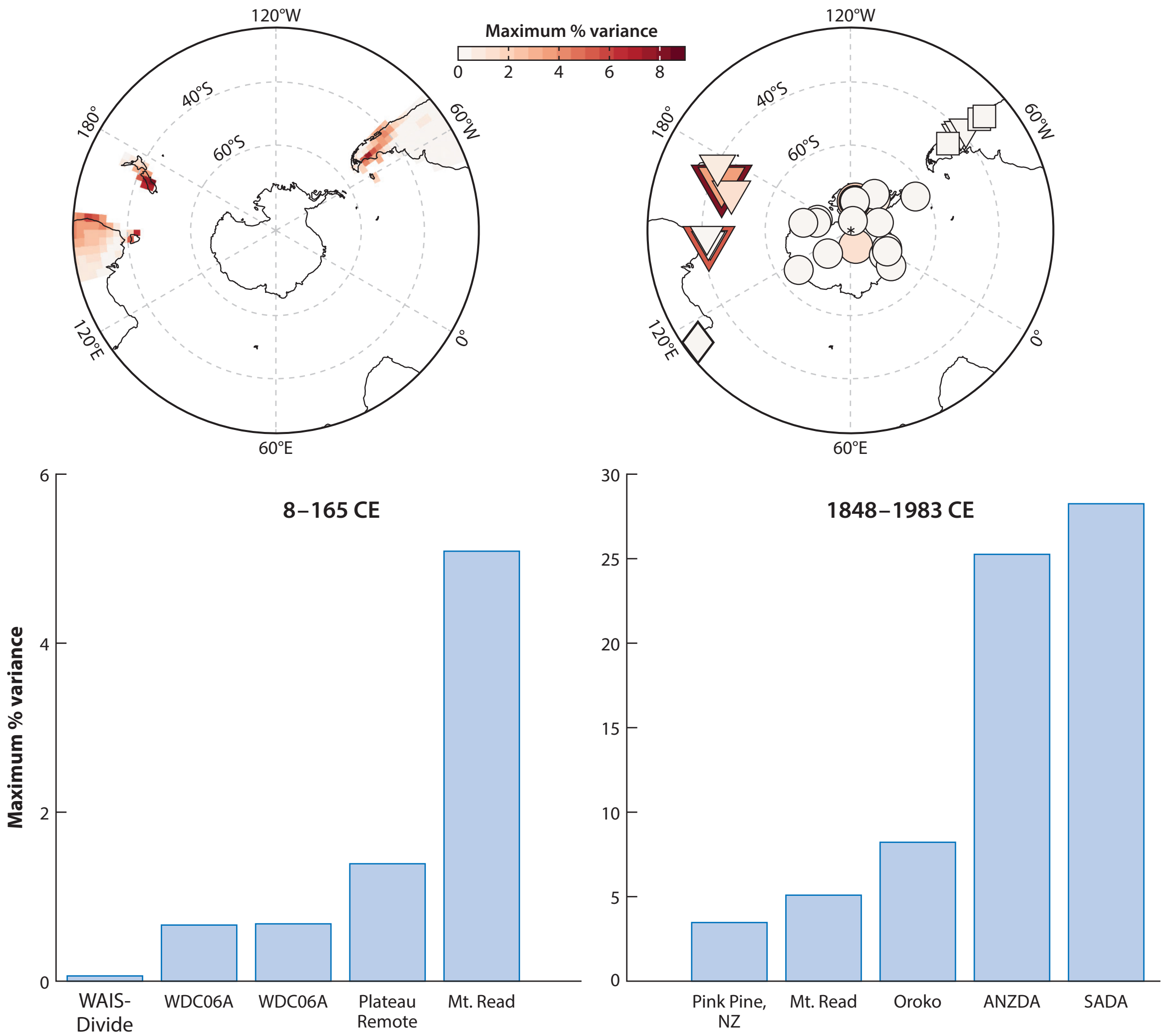

Figure 6

Optimal sensor analysis adapted from King et al. (2023a). (*top row*) The maximum percent variance constrained by the drought atlases [*left*, including the SADA (Morales et al. 2020) and ANZDA (Palmer et al. 2015)] and PAGES 2k records (*right*). (*bottom row*) Proxy sites with the greatest influence in the early part of the reconstruction (*left*; 8–135 CE) and in the postindustrial era (*right*; 1848–1983 CE). Abbreviations: ANZDA, Australia and New Zealand Drought Atlas; PAGES, Past Global Changes; SADA, South American Drought Atlas.

be large in regions far from proxy observations where models disagree in their simulated dynamics (Amrhein et al. 2020, King et al. 2021). As noted in Section 2.1, a multi-model ensemble can help address this to some extent but may present practical challenges and/or still be inadequate. For example, shared model biases in the tropical Pacific have been shown to affect data assimilation–based reconstructions of the ENSO response to large volcanic eruptions (Sanchez et al. 2021). Multi-model ensembles may be practically prohibitive due to the costs of running bespoke simulations for time periods that do not correspond to experiments from the Paleoclimate Modelling Intercomparison Project (e.g., past1000 for the period 850 to 1849 CE). As highlighted in Section 2.3, model simulations that incorporate water isotopes facilitate the assimilation of

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

isotopic proxy data, but many modeling groups do not include isotope tracers, which might necessitate the use of a single model for the prior. Thus far, the trade-offs between using an isotope-enabled single model ensemble (Tierney et al. 2020b, Osman et al. 2021) versus a multi-model ensemble that lacks isotopes (e.g., Annan et al. 2022, Erb et al. 2022) have yet to be systematically explored.

4.2. Loss of Variance

A known issue of the offline Kalman filter is that changes to the size and composition of the proxy network can artificially affect the temporal variability of the analysis ensemble (King et al. 2023a,b). As proxy records become sparse and/or more uncertain, the Kalman update is minimal, and reconstructions will approach (or collapse toward) the prior mean. This tendency can affect spectral estimates of past climate variability (Emile-Geay et al. 2025), which partition variance as a function of frequency, as well as comparisons between modern and past climate states.

To understand this effect, consider a reconstruction using the same prior for each time step. In the absence of proxy records, the mean of the prior provides the best naive estimate of past climate. Because each time step uses the same prior, the updated ensemble mean in this case will be constant through time and will have no variance. Assimilating a single proxy record provides additional information to the Kalman filter, and so the updated ensemble mean will begin to vary through time. Each additional proxy record increases the updated ensemble mean's ability to deviate from the prior mean, and so the amplitude of the ensemble mean's spectrum will increase with the size of the proxy network. This variance effect is partly by design, as the prior mean provides the best a priori estimate of past climate in the absence of additional information. In addition, a weak update upon the prior mean will be accompanied by large uncertainty bounds (**Figure 7**), which allows for the possibility of many climate trajectories. However, this effect presents a challenge for the interpretation of different periods in a reconstruction through time.

The variance effect is even more insidious for spatial field reconstructions because the information used to update each spatial point is modulated by the point's covariance with the proxies in the network. As such, points far from the proxy network may receive minimal information

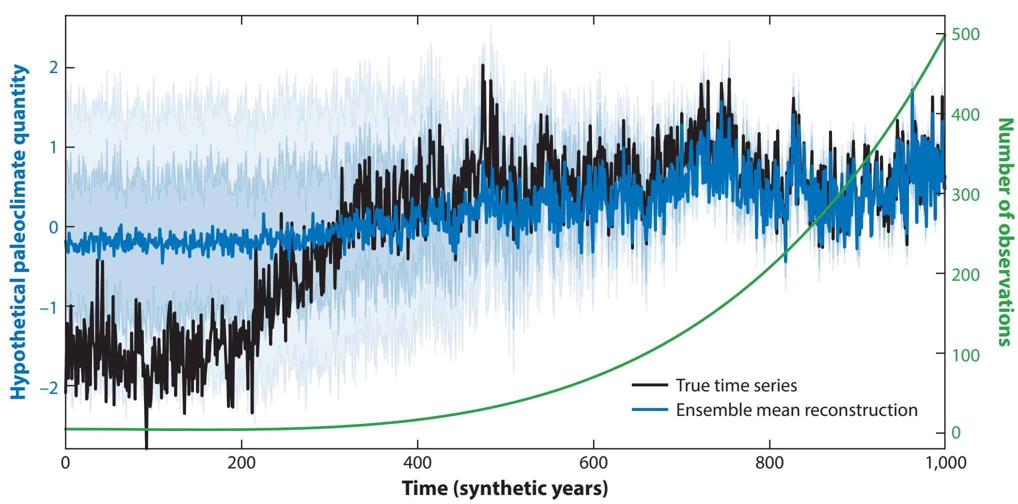

Figure 7

An example of the variance loss effect, with a toy time series and reconstruction. A synthetic autoregressive time series was reconstructed using a progressively larger number of noisy samples, leading to increasing ensemble mean variance and decreasing uncertainty in time.

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

and revert to the prior mean, whereas points close to the proxies in the network may receive sizable updates. Thus, different parts of the spatial field receive unequal updates, thereby potentially breaking expected spatial structures. We note that the variance effect is a consequence of neglecting the updated ensemble deviations and interpreting only the updated ensemble mean directly as a time series or time-varying spatial field. The updated ensemble deviations should be considered alongside the updated mean whenever analyzing climate phenomena from these reconstructions.

King et al. (2023a) provide an approach to resolving the variance issue analogous to one used in non-DA reconstruction (Cook et al. 1999, Frank et al. 2007). This method uses multiple frozen-network assimilations to assess the variance associated with different combinations of proxy records. First, only proxy records that span the full reconstruction period are assimilated, which provides a baseline time series whose variance is not affected by changes in the proxy network. Next, the method identifies each unique set of proxy records used to update at least one time step of the reconstruction. Each set is then assimilated over all time steps where it has recorded values. The standard deviation of each assimilated set is then compared to the baseline standard deviation to compute a ratio P for each set: $P(\text{set})\ \sigma_{\text{set}}/\sigma_{\text{baseline}}$. Each time step of the reconstruction then gets assigned a scaling weight based on the set of proxies used: $w(t) = P(\text{set}(t))/\max(P)$. The full reconstruction is then corrected for variance loss by multiplying its deviations from the prior mean by the scaling weights for each time step. This ensures that each time step is normalized to match the posterior variance associated with the full proxy network.

5. OUTLOOK

Offline EnKF approaches have been a leading tool for paleoDA over the past decade, but as discussed above, the method has some inherent shortcomings. However, there are other options from the pantheon of DA approaches developed over past decades, many of which have not yet been explored in depth for paleoclimate reconstruction. One way forward is to use online DA. Offline DA is simply a least squares approach informed by model covariances, and the climate model is not allowed to react to the assimilation of the data. In contrast, online DA integrates a model through time, and the model produces an adjusted forecast for each DA update step based on the proxy data. Online methods have the potential to improve paleoDA for a variety of reasons. First, they propagate innovations through time, so that improvements to the prior state are not restricted to the time of an observation (say, a measurement of deep ocean temperature, which can contain information across decades and beyond because of long ocean memory). Second, online approaches evolve prior covariances through time (known as the errors of the day in the weather literature) and can naturally handle changing statistics in the model state and its uncertainty. Third, individual ensemble members are typically more physically consistent because their evolution is governed by model processes. A data assimilation increment usually violates process relationships, including conservation of mass and energy and radiative balance, but this is not corrected for in offline DA because the model does not respond to the update.

There are several reasons why online DA methods have only been used in a limited way in paleoclimate thus far. The first is the substantial computational cost of integrating a large ensemble of models over paleoclimate timescales. Hence, online applications have typically used either models of intermediate complexity (Goosse et al. 2010, 2012; Masoum et al. 2024) or statistical emulators (Perkins & Hakim 2021). The basic idea of the emulator approach is that the spatial covariances and time evolution of a complex climate model can be empirically estimated; this estimation can then be used to forecast future states of the climate system based on the current state. For example, linear-inverse models capture the linear dynamics of a system and have been widely used in climate science to make, for example, operational forecasts of ENSO (Newman et al. 2009). Linear-inverse models have so far been used in paleoDA for reconstructing large-scale climate

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

phenomena that have strong interannual covariances, such as global mean temperature and upper ocean heat content (Perkins & Hakim 2021).

A second reason online DA has not been widely used is that some studies have found that it is of minimal added value for paleoclimate reconstruction, likely because the memory of the atmosphere in particular is short and much smaller than the length of time represented by proxy data (Matsikaris et al. 2015, Acevedo et al. 2017). However, recent work suggests that online DA offers tangible benefits for ocean variables (Okazaki et al. 2021, Perkins & Hakim 2021), which have longer memories. Additionally, the use of deep learning models is a recent and promising advancement in online DA. Deep learning models can capture complex, nonlinear relationships and have shown potential for future paleoDA applications by outperforming linear-inverse models (Meng & Hakim 2024).

A second potential way forward for paleoDA is to incorporate smoothers, which spread information both forward and backward in time (Evensen 2009). Filters (which can propagate observational information forward in time via model integration) are commonly used in weather prediction, where the primary objective is to provide a set of initial conditions for a model forecast. By contrast, reconstructing variability over a paleoclimate interval is fundamentally a smoothing problem because observations can usefully constrain both future and past properties of the state. One approach for an offline smoother is to reformulate the fundamental DA update equation, Equation 3, to apply to both time and space instead of just space (e.g., Amrhein et al. 2015); in this setup, one needs to specify time covariances in addition to the space covariances of Equation 3. Online smoothers have been implemented using adjoint-based variational approaches (commonly known as 4DVAR) for paleoceanographic applications (Dail & Wunsch 2014, Kurahashi-Nakamura et al. 2017, Amrhein et al. 2018) as well as using ensemble transform Kalman filters (García-Pintado & Paul 2018).

6. CONCLUSIONS

Relative to more traditional paleoclimate reconstruction approaches that use simple spatial interpolation or averages of proxy time series to reconstruct regional or global phenomena, paleoDA has clear advantages in terms of quantifying remote impacts, explicitly representing uncertainty, directly accommodating PSMs and optimal sensors, and readily integrating large datasets. The method has enabled a vast array of new observations about past climates, including increasing quantitation of key metrics such as GMST and ECS, and spatial patterns of temperature and precipitation change. By combining information from proxies and paleoclimate model simulations, paleoDA seeks to leverage the best of both worlds in order to improve past climate reconstruction. The offline EnKF method is flexible in terms of its ease of application across multiple timescales, but it does have limitations and drawbacks, several of which we have discussed. We anticipate that future research will adapt newer techniques from the data assimilation and weather forecasting community in order to continue optimizing how model and proxy data can be combined.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This review was supported by National Science Foundation grants AGS-1602301, AGS-1803946, and AGS-1803995; Heising-Simons Foundation grant 2016-015; and the Packard Fellowship in Science and Engineering to J.E.T. E.J.J. was supported by the PhanTASTIC Postdoctoral Fellowship, funded by Roland and Debra Sauermann through the Smithsonian Institution, and

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

the Thomas R. Brown Distinguished Chair in Integrative Science at The University of Arizona through J.E.T. N.J.S. was supported by the Israel Science Foundation grant 2654/20. This material is based upon work supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1852977.

LITERATURE CITED

Acevedo W, Fallah B, Reich S, Cubasch U. 2017. Assimilation of pseudo-tree-ring-width observations into an atmospheric general circulation model. *Clim. Past* 13(5):545–57

Amrhein DE, Gebbie G, Marchal O, Wunsch C. 2015. Inferring surface water equilibrium calcite δ^{18}O during the last deglacial period from benthic foraminiferal records: implications for ocean circulation. *Paleoceanography* 30(11):1470–89

Amrhein DE, Hakim GJ, Parsons LA. 2020. Quantifying structural uncertainty in paleoclimate data assimilation with an application to the last millennium. *Geophys. Res. Lett.* 47(22):e2020GL090485

Amrhein DE, Wunsch C, Marchal O, Forget G. 2018. A global glacial ocean state estimate constrained by upper-ocean temperature proxies. *J. Clim.* 31(19):8059–79

Annan J, Hargreaves JC. 2013. A new global reconstruction of temperature changes at the Last Glacial Maximum. *Clim. Past* 9(1):367–76

Annan JD, Hargreaves JC, Mauritsen T. 2022. A new global surface temperature reconstruction for the Last Glacial Maximum. *Clim. Past* 18(8):1883–96

Bova S, Rosenthal Y, Liu Z, Godad SP, Yan M. 2021. Seasonal origin of the thermal maxima at the Holocene and the last interglacial. *Nature* 589(7843):548–53

Bowen GJ, Cai Z, Fiorella RP, Putman AL. 2019. Isotopes in the water cycle: regional- to global-scale patterns and applications. *Annu. Rev. Earth Planet. Sci.* 47:453–79

Brennan MK, Hakim GJ. 2022. Reconstructing arctic sea ice over the Common Era using data assimilation. *J. Clim.* 35(4):1231–47

Brierley CM, Zhao A, Harrison SP, Braconnot P, Williams CJR, et al. 2020. Large-scale features and evaluation of the PMIP4-CMIP6 *midHolocene* simulations. *Clim. Past* 16(5):1847–72

Cenozoic CO_2 Proxy Integ. Proj. (CenCO2PIP) Consort. 2023. Toward a Cenozoic history of atmospheric CO_2. *Science* 382(6675):eadi5177

CLIMAP Proj. Memb. 1976. The Surface of the Ice-Age Earth: Quantitative geologic evidence is used to reconstruct boundary conditions for the climate 18,000 years ago. *Science* 191(4232):1131–37

Comboul M, Emile-Geay J, Hakim GJ, Evans MN. 2015. Paleoclimate sampling as a sensor placement problem. *J. Clim.* 28(19):7717–40

Cook ER, Briffa KR, Jones PD. 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *Int. J. Climatol.* 14(4):379–402

Cook ER, Meko DM, Stahle DW, Cleaveland MK. 1999. Drought reconstructions for the continental United States. *J. Clim.* 12(4):1145–62

Cook ER, Seager R, Cane MA, Stahle DW. 2007. North American drought: reconstructions, causes, and consequences. *Earth-Sci. Rev.* 81(1):93–34

Cooper VT, Armour KC, Hakim GJ, Tierney JE, Osman MB, et al. 2024. Last Glacial Maximum pattern effects reduce climate sensitivity estimates. *Sci. Adv.* 10(16):eadk9461

Dail H, Wunsch C. 2014. Dynamical reconstruction of upper-ocean conditions in the Last Glacial Maximum Atlantic. *J. Clim.* 27(2):807–23

Dalaiden Q, Goosse H, Rezsöhazy J, Thomas ER. 2021. Reconstructing atmospheric circulation and sea-ice extent in the West Antarctic over the past 200 years using data assimilation. *Clim. Dyn.* 57(11):3479–503

Dannenberg MP, Wise EK. 2013. Performance of climate field reconstruction methods over multiple seasons and climate variables. *J. Geophys. Res. Atmos.* 118(17):9595–610

Dätwyler C, Neukom R, Abram NJ, Gallant AJ, Grosjean M, et al. 2018. Teleconnection stationarity, variability and trends of the Southern Annular Mode (SAM) during the last millennium. *Clim. Dyn.* 51(5):2321–39

Deser C, Phillips A, Bourdette V, Teng H. 2012. Uncertainty in climate change projections: the role of internal variability. *Clim. Dyn.* 38:527–46

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

Dubinkina S, Goosse H, Sallaz-Damaz Y, Crespin E, Crucifix M. 2011. Testing a particle filter to reconstruct climate changes over the past centuries. *Int. J. Bifurc. Chaos* 21(12):3611–18

Emile-Geay J, Cobb KM, Cole JE, Elliot M, Zhu F. 2020. Past ENSO variability: reconstructions, models, and implications. In *El Niño Southern Oscillation in a Changing Climate*, ed. MJ McPhaden, A Santoso, W Cai, pp. 87–118. Hoboken, NJ: Wiley

Emile-Geay J, Hakim G, Viens F, Zhu F, Amrhein DE. 2025. Temporal comparisons involving paleoclimate data assimilation: challenges and remedies. *J. Clim.* **https://doi.org/10.1175/JCLI-D-24-0101.1**

Erb MP, McKay NP, Steiger N, Dee S, Hancock C, et al. 2022. Reconstructing Holocene temperatures in time and space using paleoclimate data assimilation. *Clim. Past* 18(12):2599–629

Essell H, Esper J, Wanner H, Büntgen U. 2024. Rethinking the Holocene temperature conundrum. *Clim. Res.* 92:61–64

Eswaran A, Truax OJ, Fudge TJ. 2024. 20th-century Antarctic sea level mitigation driven by uncertain East Antarctic accumulation history. *Geophys. Res. Lett.* 51(9):e2023GL106991

Evans MN, Kaplan A, Cane MA. 1998. Optimal sites for coral-based reconstruction of global sea surface temperature. *Paleoceanography* 13(5):502–16

Evans MN, Kaplan A, Cane MA. 2002. Pacific sea surface temperature field reconstruction from coral δ^{18}O data using reduced space objective analysis. *Paleoceanography* 17(1):7-1–7-13

Evans MN, Kaplan A, Cane MA, Villalba R. 2001. Globality and optimality in climate field reconstructions from proxy data. In *Interhemispheric Climate Linkages*, ed. V Markgraf, pp. 53–72. Cambridge, UK: Cambridge Univ. Press

Evans MN, Tolwinski-Ward SE, Thompson DM, Anchukaitis KJ. 2013. Applications of proxy system modeling in high resolution paleoclimatology. *Quat. Sci. Rev.* 76:16–28

Evensen G. 2009. *Data Assimilation: The Ensemble Kalman Filter*. Berlin: Springer

Fernández-Donado L, González-Rouco J, Raible C, Ammann C, Barriopedro D, et al. 2013. Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Clim. Past* 9(1):393–421

Forster P, Storelvmo T, Armour K, Collins W, Dufresne JL, et al. 2021. The Earth's energy budget, climate feedbacks, and climate sensitivity. 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*, ed. V Masson-Delmotte, P Zhai, A Pirani, S Connors, C Péan, et al., pp. 923–1054. Cambridge, UK: Cambridge Univ. Press

Frank D, Esper J, Cook ER. 2007. Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophys. Res. Lett.* 34(16):L16709

Franke J, Brönnimann S, Bhend J, Brugnara Y. 2017. A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations. *Sci. Data* 4:170076

Fritts HC, Blasing TJ, Hayden BP, Kutzbach JE. 1971. Multivariate techniques for specifying tree-growth and climate relationships and for reconstructing anomalies in paleoclimate. *J. Appl. Meteorol.* 10:845–64

Gallant AJ, Phipps SJ, Karoly DJ, Mullan AB, Lorrey AM. 2013. Nonstationary Australasian teleconnections and implications for paleoclimate reconstructions. *J. Clim.* 26(22):8827–49

García-Pintado J, Paul A. 2018. Evaluation of iterative Kalman smoother schemes for multi-decadal past climate analysis with comprehensive Earth system models. *Geosci. Model Dev.* 11(12):5051–84

Gaspari G, Cohn SE. 1999. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* 125(554):723–57

Gebbie G, Huybers P. 2006. Meridional circulation during the Last Glacial Maximum explored through a combination of South Atlantic δ^{18}O observations and a geostrophic inverse model. *Geochem. Geophys. Geosyst.* 7(11):Q11N07

Gill EC, Rajagopalan B, Molnar P, Marchitto TM. 2016. Reduced-dimension reconstruction of the equatorial Pacific SST and zonal wind fields over the past 10,000 years using Mg/Ca and alkenone records. *Paleoceanography* 31(7):928–52

Gong D, Wang S. 1999. Definition of Antarctic Oscillation Index. *Geophys. Res. Lett.* 26(4):459–62

Goosse H, Crespin E, de Montety A, Mann M, Renssen H, Timmermann A. 2010. Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. *J. Geophys. Res.* 115(D9):D09108

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

Goosse H, Crespin E, Dubinkina S, Loutre MF, Mann ME, et al. 2012. The role of forcing and internal dynamics in explaining the medieval climate anomaly. *Clim. Dyn.* 39(12):2847–66

Goosse H, Renssen H, Timmermann A, Bradley RS. 2005. Internal and forced climate variability during the last millennium: a model-data comparison using ensemble simulations. *Quat. Sci. Rev.* 24(12–13):1345–60

Hakim GJ, Emile-Geay J, Steig EJ, Noone D, Anderson DM, et al. 2016. The last millennium climate reanalysis project: framework and first results. *J. Geophys. Res. Atmos.* 121(12):6745–64

Hamill TM, Whitaker JS, Snyder C. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Weather Rev.* 129(11):2776–90

Hu J, Dee SG, Wong CI, Harman CJ, Banner JL, Bunnell KE. 2021. Assessing proxy system models of cave dripwater δ^{18}O variability. *Quat. Sci. Rev.* 254:106799

Inglis GN, Bragg F, Burls N, Evans D, Foster GL, et al. 2020. Global mean surface temperature and climate sensitivity of the EECO, PETM and latest Paleocene. *Clim. Past* 16:1953–68

Judd EJ, Tierney JE, Huber BT, Wing SL, Lunt DJ, et al. 2022. The PhanSST global database of Phanerozoic sea surface temperature proxy data. *Sci. Data* 9(1):753

Judd EJ, Tierney JE, Lunt DJ, Montañez IP, Huber BT, et al. 2024. A 485 million-year history of Earth's surface temperature. *Science* 385(6715):eadk3705

Kageyama M, Harrison SP, Kapsch ML, Lofverstrom M, Lora JM, et al. 2021. The PMIP4 Last Glacial Maximum experiments: preliminary results and comparison with the PMIP3 simulations. *Clim. Past* 17(3):1065–89

Kalnay E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge, UK: Cambridge Univ. Press

Kaplan A, Kushnir Y, Cane MA, Blumenthal MB. 1997. Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J. Geophys. Res.* 102(C13):27835–60

Kaufman DS, Broadman E. 2023. Revisiting the Holocene global temperature conundrum. *Nature* 614(7948):425–35

Kaufman DS, McKay N, Routson C, Erb M, Dätwyler C, et al. 2020. Holocene global mean surface temperature, a multi-method reconstruction approach. *Sci. Data* 7:201

King J, Anchukaitis KJ, Allen K, Vance T, Hessl A. 2023a. Trends and variability in the Southern Annular Mode over the Common Era. *Nat. Commun.* 14(1):2324

King J, Anchukaitis KJ, Tierney JE, Hakim GJ, Emile-Geay J, et al. 2021. A data assimilation approach to last millennium temperature field reconstruction using a limited high-sensitivity proxy network. *J. Clim.* 34(17):7091–111

King J, Tierney J, Osman M, Judd EJ, Anchukaitis KJ. 2023b. DASH: a MATLAB toolbox for paleoclimate data assimilation. *Geosci. Model Dev.* 16(19):5653–83

Kurahashi-Nakamura T, Paul A, Losch M. 2017. Dynamical reconstruction of the global ocean state during the Last Glacial Maximum. *Paleoceanography* 32(4):326–50

Liu Z, Zhu J, Rosenthal Y, Zhang X, Otto-Bliesner BL, et al. 2014. The Holocene temperature conundrum. *PNAS* 111(34):E3501–5

Malevich SB, Vetter L, Tierney JE. 2019. Global core top calibration of δ^{18} O in planktic foraminifera to sea surface temperature. *Paleoceanogr. Paleoclimatol.* 34(8):1292–315

Mann ME, Zhang Z, Hughes MK, Bradley RS, Miller SK, et al. 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *PNAS* 105(36):13252–57

Marcott SA, Shakun JD, Clark PU, Mix AC. 2013. A reconstruction of regional and global temperature for the past 11,300 years. *Science* 339(6124):1198–201

Marvel K, Cook BI. 2022. Using machine learning to identify novel hydroclimate states. *Philos. Trans. R. Soc. A* 380(2238):20210287

Masoum A, Nerger L, Willeit M, Ganopolski A, Lohmann G. 2024. Paleoclimate data assimilation with CLIMBER-X: an ensemble Kalman filter for the last deglaciation. *PLOS ONE* 19(4):e0300138

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

Meng Z, Hakim GJ. 2024. Reconstructing the tropical Pacific upper ocean using online data assimilation with a deep learning model. *J. Adv. Model. Earth Syst.* 16(11):e2024MS004422

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

Morales MS, Cook ER, Barichivich J, Christie DA, Villalba R, et al. 2020. Six hundred years of South American tree rings reveal an increase in severe hydroclimatic events since mid-20th century. *PNAS* 117(29):16816–23

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models part I—a discussion of principles. *J. Hydrol.* 10(3):282–90

Newman M, Sardeshmukh PD, Penland C. 2009. How important is air–sea coupling in ENSO and MJO evolution? *J. Clim.* 22(11):2958–77

O'Connor GK, Steig EJ, Hakim GJ. 2021. Strengthening Southern Hemisphere westerlies and Amundsen Sea low deepening over the 20th century revealed by proxy-data assimilation. *Geophys. Res. Lett.* 48(24):e2021GL095999

Okazaki A, Miyoshi T, Yoshimura K, Greybush SJ, Zhang F. 2021. Revisiting online and offline data assimilation comparison for paleoclimate reconstruction: an idealized OSSE study. *J. Geophys. Res. Atmos.* 126(16):e2020JD034214

Osman MB, Tierney JE, Zhu J, Tardif R, Hakim GJ, et al. 2021. Globally resolved surface temperatures since the Last Glacial Maximum. *Nature* 599(7884):239–44

Ossandón Á, Gual J, Rajagopalan B, Kleiber W, Marchitto T. 2024. Spatial and temporal Bayesian hierarchical model over large domains with application to Holocene Sea surface temperature reconstruction in the equatorial Pacific. *Paleoceanogr. Paleoclimatol.* 39(12):e2024PA004844

Otto-Bliesner BL, Brady EC, Fasullo J, Jahn A, Landrum L, et al. 2016. Climate variability and change since 850 CE: an ensemble approach with the Community Earth System Model. *Bull. Am. Meteorol. Soc.* 97(5):735–54

Palmer JG, Cook ER, Turney CS, Allen K, Fenwick P, et al. 2015. Drought variability in the eastern Australia and New Zealand summer drought atlas (ANZDA, CE 1500–2012) modulated by the Interdecadal Pacific Oscillation. *Environ. Res. Lett.* 10(12):124002

Parsons LA, Amrhein DE, Sanchez SC, Tardif R, Brennan MK, Hakim GJ. 2021. Do multi-model ensembles improve reconstruction skill in paleoclimate data assimilation? *Earth Space Sci.* 8(4):e2020EA001467

Perkins WA, Hakim GJ. 2021. Coupled atmosphere–ocean reconstruction of the last millennium using online data assimilation. *Paleoceanogr. Paleoclimatol.* 36(5):e2020PA003959

Sanchez SC, Hakim GJ, Saenger CP. 2021. Climate model teleconnection patterns govern the Niño-3.4 response to early nineteenth-century volcanism in coral-based data assimilation reconstructions. *J. Clim.* 34(5):1863–80

Schmittner A, Urban NM, Shakun JD, Mahowald NM, Clark PU, et al. 2011. Climate sensitivity estimated from temperature reconstructions of the Last Glacial Maximum. *Science* 334(6061):1385–88

Schneider T. 2001. Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. *J. Clim.* 14(5):853–71

Scotese CR, Song H, Mills BJ, van der Meer DG. 2021. Phanerozoic paleotemperatures: the Earth's changing climate during the last 540 million years. *Earth-Sci. Rev.* 215:103503

Sherwood SC, Webb MJ, Annan JD, Armour KC, Forster PM, et al. 2020. An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys.* 58(4):e2019RG000678

Silvestri G, Vera C. 2009. Nonstationary impacts of the southern annular mode on Southern Hemisphere climate. *J. Clim.* 22(22):6142–48

Smerdon JE, Cook ER, Steiger NJ. 2023. The historical development of large-scale paleoclimate field reconstructions over the Common Era. *Rev. Geophys.* 61(4):e2022RG000782

Smerdon JE, Kaplan A, Zorita E, González-Rouco JF, Evans M. 2011. Spatial performance of four climate field reconstruction methods targeting the Common Era. *Geophys. Res. Lett.* 38(11):L11705

Snoll B, Ivanovic R, Gregoire L, Sherriff-Tadano S, Menviel L, et al. 2024. A multi-model assessment of the early last deglaciation (PMIP4 LDv1): a meltwater perspective. *Clim. Past* 20(4):789–815

Steiger NJ, Hakim GJ, Steig EJ, Battisti DS, Roe GH. 2014. Assimilation of time-averaged pseudoproxies for climate reconstruction. *J. Clim.* 27(1):426–41

Steiger NJ, Smerdon JE, Cook ER, Cook BI. 2018. A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Sci. Data* 5:180086

Steiger NJ, Smerdon JE, Seager R, Williams AP, Varuolo-Clarke AM. 2021. ENSO-driven coupled megadroughts in North and South America over the last millennium. *Nat. Geosci.* 14(10):739–44

Downloaded from www.annualreviews.org. Cambridge University (ar-182579) IP: 131.111.5.161 On: Mon, 28 Apr 2025 11:36:10

Tardif R, Hakim GJ, Perkins WA, Horlick KA, Erb MP, et al. 2019. Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. *Clim. Past* 15(4):1251–73

Thompson AJ, Zhu J, Poulsen CJ, Tierney JE, Skinner CB. 2022. Northern Hemisphere vegetation change drives a Holocene thermal maximum. *Sci. Adv.* 8(15):eabj6535

Thompson D, Ault T, Evans M, Cole J, Emile-Geay J. 2011. Comparison of observed and simulated tropical climate trends using a forward model of coral δ^{18}O. *Geophys. Res. Lett.* 38(14):L14706

Tierney JE, Haywood AM, Feng R, Bhattacharya T, Otto-Bliesner BL. 2019a. Pliocene warmth consistent with greenhouse gas forcing. *Geophys. Res. Lett.* 46(15):9136–44

Tierney JE, King J, Osman M, Abell J, Burls N, et al. 2025. Pliocene warmth and patterns of climate change inferred from paleoclimate data assimilation. *AGU Adv.*

Tierney JE, Malevich SB, Gray W, Vetter L, Thirumalai K. 2019b. Bayesian calibration of the Mg/Ca paleothermometer in planktic foraminifera. *Paleoceanogr. Paleoclimatol.* 34(12):2005–30

Tierney JE, Poulsen CJ, Montañez IP, Bhattacharya T, Feng R, et al. 2020a. Past climates inform our future. *Science* 370(6517):eaay3701

Tierney JE, Tingley MP. 2018. BAYSPLINE: a new calibration for the alkenone paleothermometer. *Paleoceanogr. Paleoclimatol.* 33(3):281–301

Tierney JE, Zhu J, King J, Malevich SB, Hakim GJ, Poulsen CJ. 2020b. Glacial cooling and climate sensitivity revisited. *Nature* 584(7822):569–73

Tierney JE, Zhu J, Li M, Ridgwell A, Hakim GJ, et al. 2022. Spatial patterns of climate change across the Paleocene–Eocene Thermal Maximum. *PNAS* 119(42):e2205326119

Tingley MP, Craigmile PF, Haran M, Li B, Mannshardt E, Rajaratnam B. 2012. Piecing together the past: statistical insights into paleoclimatic reconstructions. *Quat. Sci. Rev.* 35:1–22

Tingley MP, Huybers P. 2010. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: development and applications to paleoclimate reconstruction problems. *J. Clim.* 23(10):2759–81

Tingley MP, Huybers P. 2013. Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature* 496(7444):201–5

Tolwinski-Ward SE, Evans MN, Hughes MK, Anchukaitis KJ. 2011. An efficient forward model of the climate controls on interannual variation in tree-ring width. *Clim. Dyn.* 36:2419–39

Treble PC, Mah M, Griffiths A, Baker A, Deininger M, et al. 2019. Separating isotopic impacts of karst and in-cave processes from climate variability using an integrated speleothem isotope-enabled forward model. EarthArXiv. **https://doi.org/10.31223/osf.io/j4kn6**

Valdes PJ, Scotese CR, Lunt DJ. 2021. Deep ocean temperatures through time. *Clim. Past* 17(4):1483–506

Van Leeuwen PJ. 2009. Particle filtering in geophysical systems. *Mon. Weather Rev.* 137(12):4089–114

Wang C, Zhang L, Lee SK, Wu L, Mechoso CR. 2014. A global perspective on CMIP5 climate model biases. *Nat. Clim. Change* 4(3):201–5

Wang J, Emile-Geay J, Guillot D, McKay NP, Rajaratnam B. 2015. Fragility of reconstructed temperature patterns over the Common Era: implications for model evaluation. *Geophys. Res. Lett.* 42(17):7162–70

Wang J, Emile-Geay J, Guillot D, Smerdon J, Rajaratnam B. 2014. Evaluating climate field reconstruction techniques using improved emulations of real-world conditions. *Clim. Past* 10(1):1–19

Webb T, Bryson RA. 1972. Late- and postglacial climatic change in the northern Midwest, USA: quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. *Quat. Res.* 2(1):70–115

Werner JP, Divine DV, Charpentier Ljungqvist F, Nilsen T, Francus P. 2018. Spatio-temporal variability of Arctic summer temperatures over the past 2 millennia. *Clim. Past* 14(4):527–57

Whitaker JS, Hamill TM. 2002. Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.* 130(7):1913–24

Wise EK. 2016. Five centuries of US West Coast drought: occurrence, spatial distribution, and associated atmospheric circulation patterns. *Geophys. Res. Lett.* 43(9):4539–46

York D. 1968. Least squares fitting of a straight line with correlated errors. *Earth Planet. Sci. Lett.* 5:320–24

Zhu J, Otto-Bliesner BL, Brady EC, Gettelman A, Bacmeister JT, et al. 2022. LGM paleoclimate constraints inform cloud parameterizations and equilibrium climate sensitivity in CESM2. *J. Adv. Model. Earth Syst.* 14(4):e2021MS002776

Zhu J, Poulsen CJ, Otto-Bliesner BL. 2020. High climate sensitivity in CMIP6 model not supported by paleoclimate. *Nat. Clim. Change* 10(5):378–79