A global Data Assimilation of Moisture Patterns from 21,000–0 BP (DAMP-21ka) using lake level proxy records

Christopher L. Hancock[1], Michael P. Erb[1], Nicholas P. McKay[1], Sylvia G. Dee[2], Ruza F. Ivanovic[3]

[1]School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ, USA
[2]Department of Earth, Environmental, & Planetary Sciences, Rice University, Houston, TX USA
[3]School of Earth and Environment, University of Leeds, Leeds, UK

Correspondence to: Christopher L. Hancock (clh624@nau.edu)

Abstract. Global hydroclimate significantly differed from modern climate during the mid-Holocene (6 ka) and Last Glacial Maximum (21 ka). Consequently, both periods have been described as either a partial or reverse analogue for current climate change. To reconstruct past hydroclimate, an offline paleoclimate data assimilation methodology is applied to a dataset of 216 lake status records which provide relative estimates of water level change. The proxy observations are integrated with the climate dynamics of two transient simulations (TraCE-21ka and HadCM3) using a multivariate proxy system model (PSM) which estimates relative lake status from available climate simulation variables. The resulting DAMP-21ka (Data Assimilation of Moisture Patterns 21,000–0 BP) reanalysis reconstructs annual lake status and precipitation values at 500-year resolution and represents the first application of the methodology to global hydroclimate on timescales spanning the Holocene and longer. Validation using Pearson's correlation coefficients indicates that the reconstruction (0.24) is more skillful, on average, than model simulations (0.09), particularly in portions of North America and East Africa where data density is high and proxy-model disagreement is prominent during the Holocene. Results of the PSM and assimilation are used to evaluate climatic controls on lake status, spatiotemporal patterns of moisture variability, and proxy-model disagreement. During the mid-Holocene, wetter conditions are reconstructed for North and East Africa, Asia, and southern Australia, but, in contrast to the model prior, negative anomalies are observed in North America resulting in drier than modern conditions throughout the Northern Hemisphere midlatitudes. Proxy-model disagreement in western North America may reflect a bias in model simulations to stronger sea level pressure gradients in the North Pacific during the mid-Holocene. The data assimilation framework is able to reconcile these differences by integrating the constraints of proxy observations with the dynamics of the model prior to produce a more robust estimation of hydroclimate variability during the past 21,000 years.

1 Introduction

During the past 21,000 years, hydroclimate varied significantly at global, regional, and local scales. These hydrologic changes are well-documented by paleoclimate proxy data (e.g., Herzschuh et al., 2023; Shanahan et al., 2015; Street-Perrott et al., 1989) and explored with general circulation model (GCM) simulations (e.g., Bartlein et al., 1998;

COHMAP Members, 1988; Harrison et al., 2015) and reflect thermodynamic and dynamic forcings associated with Milankovitch cycles of insolation, retreat of ice-sheets, and warming temperature since the Last Glacial Maximum (LGM) (Davis and Brewer, 2009; Harrison et al., 2014; Sun et al., 2021b; Tierney et al., 2020). As a result of these large external forcings, many regions experienced drier or wetter conditions than present during portions of either the deglacial period (21–11.7 ka) or the Holocene (11.7 ka – present). For example, during the LGM, colder temperatures suppressed the hydrologic cycle causing widespread aridity (Kageyama et al., 2021; Prentice and Jolly, 2000; Yung et al., 1996). However, these changes were not uniform. In regions such as the Southwest United States and Southwest Asia, lower evaporative demand and shifting storm tracks sustained lakes in basins that are dry today (Benson et al., 1990; Li and Zhang, 2020; McGee et al., 2018; Oster et al., 2015; Qin and Yu, 1998). During the early- and mid-Holocene, intensified monsoon circulation was widespread throughout the Northern Hemisphere tropics (COHMAP Members, 1988; Zhao and Harrison, 2012) as typified by the African Humid Period (e.g., deMenocal et al., 2000; Lézine et al., 2011). These anomalies were anti-phased with Southern Hemisphere monsoon regions due to precessional forcing and shifts in the Intertropical Convergence Zone (ITCZ) position (Cheng et al., 2012; Deininger et al., 2020).

Proxy records and GCMs provide complementary information about hydroclimate changes in Earth's recent climate history: paleoclimate proxies integrate the observed environmental response and timing of climate events, whereas models simulate the dynamics associated with prescribed boundary conditions. Although these two data types generally indicate similar large scale hydroclimate patterns, notable disagreements include Holocene trends in North America (Hermann et al., 2018; de Wet et al., 2023), underestimation of the African Humid Period by models (Brierley et al., 2020; Coe and Harrison, 2002; Jolly et al., 1998; Tierney et al., 2011), and uncertainty about the direction and magnitude of tropical precipitation anomalies during the LGM (Kageyama et al., 2021). To some extent, these differences reflect known biases in either data type such as the effect of "no analogue" environments on proxy calibration (Veloz et al., 2012) or the lack of dynamic vegetation, dust, and lakes in climate models (Lu et al., 2018; Pausata et al., 2016; Specht et al., 2022).

Paleoclimate data assimilation (DA) provides a framework to blend the independent information and strengths provided from each source of information. It does this by quantifying the covariance structure simulated by a model prior and updating the model estimates based on proxy data and their uncertainties, resulting in a gridded reconstruction which is constrained by both the proxy observations and dynamics of the GCM (Hakim et al., 2016). Previous DA applications to geologic timescales produced global datasets of Holocene and LGM–present (Osman et al., 2021) surface temperatures which integrate the temporal variability of proxy observations with the spatial completeness of model data. These results provide a variety of insights into past climate changes including the relative role of different climate forcings (Osman et al., 2021), potential biases affecting the Holocene Temperature Conundrum (Erb et al., 2022), and Earth's equilibrium climate sensitivity (Tierney et al., 2020).

Despite extensive comparisons between GCM simulations and hydroclimate proxies during the LGM (e.g., Bartlein et al., 1998; Chevalier et al., 2017; COHMAP Members, 1988) and the Holocene (e.g., Brierley et al., 2020; Hancock et al., 2023; Tarasov et al., 1998), hydroclimate has not been assimilated globally on these timescales. Although, model

simulations may underestimate the magnitude of Holocene precipitation anomalies for some regions (Harrison et al., 2014; Lu et al., 2019; Morrill et al., 2019), a new data compilation for hydroclimate indicates that models and proxies are more likely to agree about the sign of regional anomalies for precipitation than temperature during the mid-Holocene for most regions (Hancock et al., 2023). These results suggest that some models can accurately simulate the structure of past moisture anomalies when the magnitude of change is significant.

At present, the longest DA-based global hydroclimate reconstruction is the "Paleo Hydrodynamics Data Assimilation" (PHYDA; Steiger et al., 2018) product of the Common Era (1 to 2000 CE). The PHYDA dataset and other Common Era dynamical DA reconstructions (King et al., 2023) provide valuable information about the inter-annual variability of gridded precipitation and associated hydroclimate indices during a relatively stable period in Earth's climate history. However, the Common Era represents a narrow range of climate states relative to the Holocene (Deininger et al., 2020; Shuman et al., 2018), LGM (Kageyama et al., 2021; Osman et al., 2021) and current warming (Cook et al., 2022; King et al., 2023). The past 2,000 years also lacks information about changing hydrodynamics during time periods such as the early and mid-Holocene when Arctic warming was amplified (Miller et al., 2010; Thomas et al., 2018) and intensified monsoon circulation impacted the tropics (Cheng et al., 2012; Zhao and Harrison, 2012), both of which represent observed or projected features of modern climate change (Rantanen et al., 2022; Wang et al., 2021).

To fill this gap, we present a Data Assimilation of Moisture Patterns from 21,000 BP to present (DAMP-21ka) reanalysis, developed by assimilating lake status proxy records using transient simulations from two GCMs as priors. We define lake status as a relative estimate of water depth measured using percentile units and distinct from lake level which can indicate the absolute surface elevation (Street-Perrott et al., 1989). Extending the reconstruction beyond the Common Era requires the use of different proxy records and proxy system models (PSMs) compared to PHYDA, which relies heavily on tree ring proxies that rarely extend beyond the past two millennia (Steiger et al., 2018). Over our timescales of interest (21–0 ka), lake status records in particular have several advantages including: 1) broad spatial distribution, 2) long duration, and perhaps most importantly 3) a clear and well-documented connection to hydrologic variability. To assimilate the lake status data, we develop and apply a simple lake status PSM (see Sect. 2.4) based on the framework described by Li and Morrill (2010).

The DAMP-21ka reanalysis reconstructs paleo lake status, mean annual precipitation, and mean annual temperature. Results focus on the two hydrological variables and temperature is primarily used to contextualize differences between lake status and precipitation. To reconstruct climate variables other than lake status, the data assimilation methodology relies on the covariance relationship of the simulations for quantifying the lake status relationship to precipitation or temperature between varying locations. These gridded results allow for two key advances in paleoclimate science: 1) a format enabling more direct comparison between the climate anomalies reconstructed in proxy data and simulated by GCMs, and 2) the ability to examine plausible climate dynamics associated with observed proxy-model disagreement.

2 Data and methods

2.1 Proxy data

Proxy observations constrain past climate changes by recording environmental variability through a variety of potential pathways, but not all data can be utilized by the DA framework. The most significant requirement for assimilating proxy data is the ability to quantify the relationship between the observed proxy value and the climate state simulated by a GCM, and previous DA methodologies differ in their approach to this problem. Erb et al. (2022) selected data previously calibrated to temperature values which could therefore be directly compared to the output of GCMs. An alternative approach is to apply proxy system models (PSMs) to specific proxy types (Osman et al., 2021). Using PSMs can improve the reconstruction skill if the proxy system is sensitive to multiple climate variables (Dee et al., 2016), but is not always possible given the available proxy data types and PSMs.

Hydroclimate variables have been reconstructed by DA on paleoclimate timescales, but these results relied on datasets of entirely (Tierney et al., 2022) or predominantly (Steiger et al., 2018) temperature proxies; others have been performed regionally in the arctic where precipitation and temperature are strongly correlated (Badgeley et al., 2020). During the Holocene, nonstationary covariance relationships exist between temperature and precipitation and differ for proxies and models in several regions (Hancock et al., 2023; Herzschuh et al., 2022). In an attempt to extract a predominantly moisture-driven signal spanning the last 21ka, we assimilate only proxies which are primarily sensitive to hydroclimate; temperature proxies used by Erb et al. (2022) were purposefully excluded in this work.

A review of the Holocene hydroclimate dataset (Hancock et al., 2023) identified three proxy types with widespread distributions and which could theoretically be applied to DA: 1) speleothem and lake $\delta^{18}O$, 2) pollen assemblages calibrated to mean annual precipitation, and 3) lake status records. The $\delta^{18}O$ records were excluded because of the requirement for an isotope enabled model prior. Although time slice (Osman et al., 2021) and transient (Dee et al., 2015b; He et al., 2021) simulations exist for portions of the LGM–present, there is currently no available isotope-enabled transient simulation for the entire period. The pollen data could be applied to a DA based on the framework of Erb et al. (2022), which assimilates values previously calibrated to climate variable units. However, straightforward comparisons between these data and available model simulations have already resulted in an extensive literature on pollen based data-model comparisons including the evaluation of multiple generations of PMIP simulations (Bartlein et al., 1998; Brierley et al., 2020; Harrison et al., 2015; Mauri et al., 2015; Sun et al., 2021b). For this project, we sought to produce a reconstruction separate from the data used by these previous projects to provide independent results for comparison.

In summary, we assimilate lake status records inferred from shoreline and geomorphic evidence of paleo water levels. These lake status data provide a widely distributed dataset recording the direct effects of climate change on basin-wide moisture availability while avoiding a number of secondary inferences needed for interpreting isotope or pollen reconstructions related to calibration, evaporative enrichment, and circulation dynamics (Dayem et al., 2010; Dee et al.,

2015a; Liefert and Shuman, 2020; Williams and Shuman, 2008). Pollen records calibrated to reconstruct precipitation were withheld as out-of-sample data to validate the lake status-based precipitation reconstruction.

2.1.1 Lake status data

216 lake level records (Table S1) were selected by searching existing proxy compilations (Hancock et al., 2023; Liefert and Shuman, 2020; Street-Perrott et al., 1989; Qin et al. 1998) to find data with a minimum duration of 10,000 years, at least 8 data points in the timeseries, at least 6 data points during the Holocene, and a median resolution less than 1,500 years in the time period spanning 21,000–0 BP. Although the data are spatially biased towards North America (n = 62), most terrestrial regions outside of the high latitudes are represented (Fig. 1). Notably, the data density in Africa, a region with typically poor proxy coverage, is substantial. During the past 12 ka, data density remains above 93%, but the number of available records decreases prior to the Holocene to a minimum of 70 records at 20.25 ka.

The data are primarily sourced from the Oxford Lake Status (OLS) Databank (n = 86; Street-Perrott et al., 1989) which provides ordinal data measuring the vertical change in the elevation of the lake surface binned into three categories: high (70–100% of total vertical range of lake fluctuation), intermediate (15–70%), and low (0–15%). Although this qualitative metric is limiting, the OLS data provide valuable information in otherwise data sparse regions such as North Africa (Fig. 1). Prior to assimilation, all proxy values are converted to percentile units with low and high designations assigned as 0 and 100, respectively. Intermediate values differ for each record according to the percentile scale and the relative number of high and low values (Fig. 2). These data were supplemented with 74 sites from the Global Lake Status DataBase (GLSDB; Harrison et al., 2003) which provide similar ordinal data but with increased spatial coverage in Eurasia from the inclusion of additional regional data compilations (Tarasov et al., 1996; Yu, 2001; Yu and Harrison, 1995).

If available, OLS and GLSDB data were replaced by newer, higher-resolution records from the same site. An additional 52 records are provided by the more recently created North America (n = 28, Liefert and Shuman, 2020), eastern and southern Africa (n = 19; De Cort et al., 2021), and Australian (n = 5; Clerke et al., 2023) lake level databases, and another four records from West Africa (Shanahan et al., 2006), South America (Placzek et al. 2006), and Asia (Jiang et al., 2020; Xu et al., 2020b) were identified according to the defined selection criteria. These data represent a variety of measurement types, ranging between relative lake status estimates similar to the OLS to more precise values of absolute water level elevation. Generally, the newer data provide improved measurement values which are more descriptive and more precise in age compared to the OLS data (Fig. 2a).

Both data types were standardized to a 500-year common time bin with data interpolated using a nearest neighbor approach. If multiple values were recorded within the same time bin, the mean value was calculated. Finally, the data were converted to percentile units and the mean value during the 11.7–0 ka reference period was removed (Fig. 2). Although the percentile conversion removes information of more precise measurements, it provides a method of characterizing lake status

change which can be directly compared between different lake status record types and, more importantly, with the PSM developed for the project and the data assimilation framework (see Sect. 2.4).

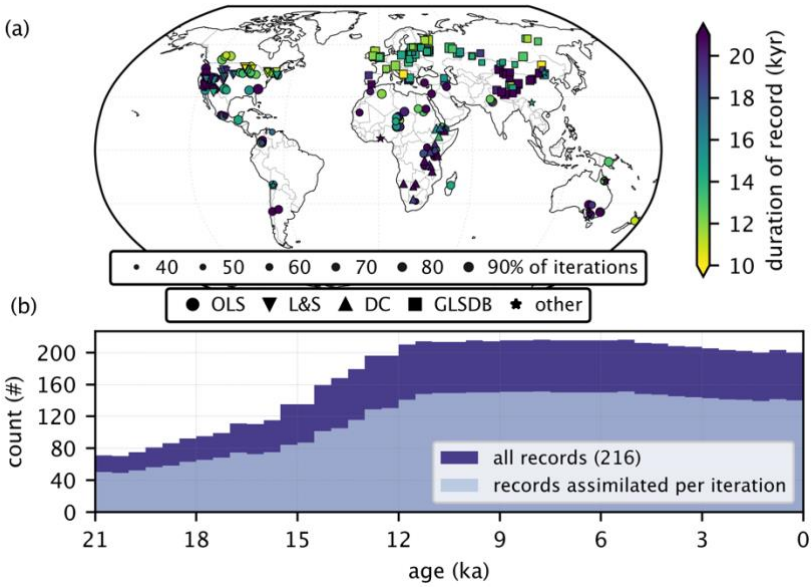


Figure 1: Lake status proxy records. (a) Map showing the spatial distribution of proxy records. Colors indicate the duration of each record, and the size is proportional to the number of iterations each record is used in the assimilation. Shape represents the source of the data including OLS for the Oxford Lake Status databank, GLS for the Global Lake Status databank, L&S for Liefert and Shuman (2020), and DC for De Cort et al. (2021). Other indicates records from smaller data compilations or individual publications. A reference list for each site is provided in Table S1. (b) Data density plot showing the total number of available proxy records at each time step (dark blue) and the mean number of records used by each iteration of the assimilation (light blue). The difference between these represents the number of proxies records withheld for validation.

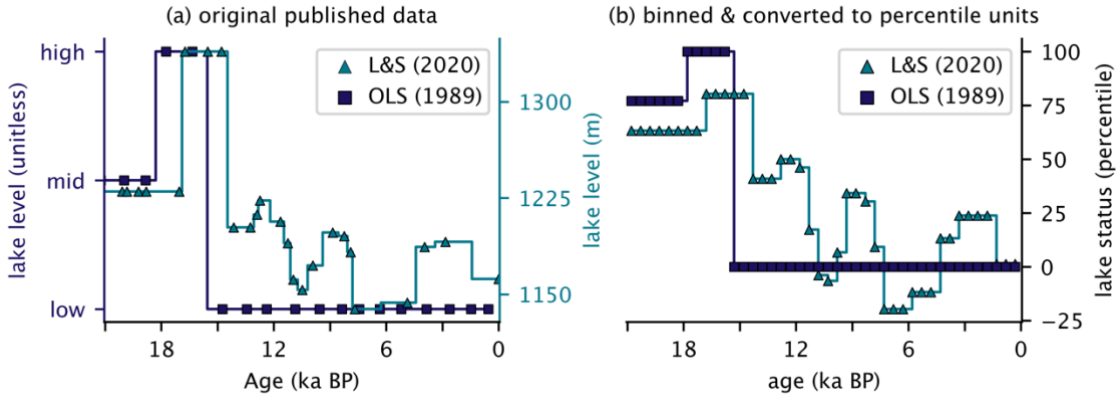


Figure 2: Data pre-processing for two lake records from Pyramid Lake, Nevada, USA. The record from the Oxford lake status data bank is shown by purple squares (Broecker and Orr, 1958; Street-Perrott et al., 1989). These data were superseded by a newer record (Benson et al., 1990; Briggs et al., 2005) compiled by Liefert and Shuman (2020) which is shown as teal triangles. (a) The original data for the two records. (b) The same records standardized to a common time axis using a nearest neighbor interpolation and converted percentile unit for assimilation. The mean value during the 11.7–0 ka reference period is then subtracted from each record.

2.1.2 Validation data

The coarse time resolution of the lake status proxy records prevents comparison to the instrumental period or even reconstructions of the Common Era. Therefore, 30% of the available lake status records were withheld during the assimilation to evaluate the resulting reconstruction following Osman et al. (2021). To account for potential biases from the selection of sites used for assimilation and validation, a different sample of 70% proxy records was used for each iteration (see Sect. 2.2). The mean validation statistic (i.e., correlation coefficient) was then calculated for each record using iterations for which the data was withheld.

For the DAMP-21ka reanalysis, lake status data were assimilated to reconstruct a variety of additional climate fields including mean annual precipitation. The DA method relies on the covariance between multiple variables to quantify this relationship, but no precipitation (mm/a) records were included. Therefore, to measure the precipitation reconstruction skill, 350 calibrated pollen records were selected from the Holocene hydroclimate data compilation (Hancock et al., 2023). These records primarily originate from the LegacyClimate 1.0 (Herzschuh et al., 2023) and are supplemented by additional records from sites not included in that dataset (Table S2). Previous studies have used calibrated pollen data to evaluate GCM simulations (e.g. Brierley et al., 2020), and our approach is conceptually similar to validate the reanalysis.

2.2 Data assimilation methodology

Lake status proxies provide valuable observations about hydrologic variations in Earth's past but are spatially limited to the basins which they infill. Data assimilation provides a mathematical technique for using these data to create a spatially complete reconstruction of specific climate variables of interest. The methods used for DAMP-21ka are based on the

temperature reconstruction of Erb et al. (2022) which follows previous DA projects for the last millennium (Hakim et al., 2016; Steiger et al., 2014). A detailed explanation of the methodology is described by these references, but a brief explanation of data assimilation, the model prior, and different parameter choices is provided here.

The two major components of data assimilation are the innovation and the Kalman gain. The innovation constrains the results using the observed data and is calculated as the difference between each proxy value and the estimated value from the prior using a PSM. This difference represents the mismatch between the model prior and the proxy at the proxy location. The Kalman gain then quantifies the relationship between the innovation and the target variable for each location using the covariance structure of the model prior. This calculation allows for the translation between proxy variables to a broader climate field. The equations for these calculations are provided in Sect. S1.

A significant change from the Erb et al., 2022 methodology is the implementation of an ensemble process, so that the DA is performed 10 times. During each iteration, a different sample of proxy records (70%) and a different sample of climate states (70%) is selected for the model prior. This technique increases the computational cost of the analysis, but provides a more robust framework for uncertainty quantification. If the posterior results for a region are consistent regardless of the specific proxies and model states sampled, the results can be considered more robust than if posterior varies substantially between iterations. We note that previous paleoclimate DA reanalyses have also used a randomized data selection process (Osman et al., 2021; Tardif et al., 2019).

Several modifications to the parameter choices were implemented to optimize the reconstruction for the hydroclimate proxy, target climate variables, and longer timeframe. The 500-year timestep of reanalysis reflects the coarser resolution of the proxies. The analysis also converts all data to anomalies relative to a longer reference period (11.7–0 ka) than Erb et al. (2022) based on the consistent data density of the lake deposits throughout the entire Holocene (Fig. 1b). A localization radius (w_{loc}) of 8,000 kilometers is implemented for the reanalysis (Gaspari and Cohn, 1999; Tardif et al., 2019). Paleoclimate DA reconstructions differ on the inclusion of w_{loc} with some applying a spatial weight (Osman et al., 2021) and others not (Erb et al., 2022; Steiger et al., 2018). Excluding w_{loc} causes the method to rely more heavily on the model covariance matrix to appropriately determine if a skillful relationship exists between distant locations. This feature may be considered a benefit or detriment depending on the research goal (Hamill et al., 2001). During the Holocene, transient trends for different regions can impose strong covariances (Herzschuh et al., 2022) which presents a challenge for the two regions with the most abundant lake status data: North America and North Africa. Proxy records indicate a negative covariance between these regions based on the Holocene wetting trend for North America and a drying trend for North Africa, but transient models simulate a drying trend for each (Hancock et al., 2023). Using a localization radius allows results in North America and Africa to update independent of each other.

To assess the sensitivity of the method to various combinations of the localization radius and R, a proxy uncertainty parameter, assimilations were performed and evaluated by comparing the results to the withheld proxy data. Based on these results (Table S3), the localization of 8,000 km and uncertainty of the 30th percentile were selected to maximize the skill of

the precipitation reconstruction. The R value of 30, which is squared for the assimilation, also corresponds to the generalized categorization of the OLS lake status records into thirds.

2.3 Model prior

We develop a time-varying multi-model prior using the methods of Erb et al. (2022). The model prior is constructed using data from two available transient simulations for the LGM to present: 1) TraCE-21k (Transient Climate Evolution over the last 21,000 years; Liu et al., 2009) which used NCAR's CCSM3 (Collins et al., 2006) climate model and 2) a HadCM3 transient deglaciation simulation which used the UK Met Office's HadCM3 (Hadley centre Coupled Model, version 3; Snoll et al., 2022). Using a multi-model prior is valuable for limiting the biases of any one model on the result (Bach and Ghil, 2023; Parsons et al., 2021), and this feature is particularly important for a hydroclimate DA because of the wider dispersion of Holocene anomalies between different models for precipitation than temperature (Hancock et al., 2023). By using simulations from multiple models, the resulting covariance structure will provide stronger covariances where both models agree that there is a strong climate connection, but weaker covariances if the two models disagree. To create the multi-model prior, data from both HadCM (2.5° by 3.75°) and TraCE (~3.71° by 3.75°) were regridded to a common spatial resolution (2.8125° by 3.75°).

The prior ensemble was developed by sampling 140 multidecadal (50 year) mean climate states within a shifting 5,000-year window centered on each timestep. This sampling was repeated randomly for each iteration. The choice of a time-varying prior differs from PHYDA (Steiger et al., 2018) but is consistent with DA reconstructions extending beyond the Common Era (Erb et al., 2022; Osman et al., 2021). For longer timescales, changing ice sheet configuration and sea level can alter circulation dynamics and add non-stationarity to the covariance relationship between different locations and climate variables (Herzschuh et al., 2022). The 5,000-year prior window length was selected to prevent climate states which include a rapid decline of the Laurentide Ice Sheet prior to 8.5 ka (Gregoire et al., 2018) from influencing the covariance structure during the mid-Holocene (6 ka).

2.4 Lake status PSM

To quantitatively compare proxy records and model data using consistent units, we estimate lake variability from GCM output variables based on a novel proxy system model (PSM). A suitably complex PSM for assimilating the available lake status data has several requirements. The PSM must: 1) convert available model output variables into a lake status percentile, 2) be applied globally to any lake, and 3) function independently of any knowledge about lake size, basin topography, runoff efficiency, or seasonal variability. This third requirement eliminates lake balance models that have been used in the past (Morrill et al., 2019). Although these characteristics lead to more mechanistic and realistic models, this information is unavailable for many datasets, and calibrating models at site-specific scales is impractical for global studies (Dee et al., 2015b; Evans et al., 2013).

The resulting PSM (Fig. 3) is derived from Li and Morrill (2010), which provides the following two equations for estimating lake status change from a GCM simulation. The first is a simple steady-state lake water-balance equation describing the surface inputs and outputs of any lake as:

$$D = A_B * Q + A_L * (P_L - E_L), \quad (1)$$

where D = discharge, A_B = area of the basin, Q = runoff from the basin into the lake, A_L = area of the lake, P_L = precipitation directly onto the lake, and E_L = evaporation directly from the lake. Because the proxy dataset largely describes closed-basin lakes, D is assumed to be 0 and the water balance equation can be rewritten as:

$$\frac{A_L}{A_B} = \frac{\mathrm{Q}}{E_L - P_L}, \quad (2)$$

Thus, three variables are needed to calculate the water balance (A_L/A_B) of any unknown lake: runoff, E_L and P_L. Runoff and precipitation are both simulated by the two climate models and are available as output variables, but lakes are not simulated by the climate models, so lake evaporation is unavailable. Therefore, we estimate E_L using the Priestley–Taylor equation (Priestley and Taylor, 1972; Vremec and Collenteur, 2022) for potential evapotranspiration ($E_{potential}$), which considers net radiation, temperature, relative humidity, and air pressure.

Previous water balance studies have used a more computationally expensive lake energy-balance model to calculate E_L from paleoclimate simulation output variables (Morrill, 2004). However, by assuming that the lake provides an unlimited supply of water, $E_{potential}$ can be equated to E_L (McGee et al., 2018). Many of the same GCM outputs are used to calculate E_L in both methods, and the Priestley–Taylor equation can be applied to the global, transient data assimilation methodology much more efficiently. We evaluated the ability of the simplified lake status estimation to emulate the more complex energy balance models used by previous studies using Community Climate System Model (CCSM4) CMIP5 simulations of the LGM and preindustrial periods (Brady et al., 2013; Gent et al., 2011). The lake status calculation was applied to these data, and the results were compared to those of Lowry and Morrill (2019).

After calculating A_L/A_B, the decadal output for each grid cell is converted into a percentile ranging between 0–100 with 0 representing the decade with the lowest lake status value during the entire 0–21 ka timeseries and 100 as the highest. Although many of the proxy records provide vertical elevation of the lake level, the classification into lake status rankings allows for the direct comparison with A_L/A_B using the sequential ranking of percentiles:

$$Lake\ Status = \mathrm{percentile}\left(\frac{\mathrm{runoff}}{E_{potential} - P_L}\right), \quad (3)$$

However, a complication arises when P_L exceeds $E_{potential}$, and the A_L/A_B becomes negative (Li and Morrill, 2010). To accommodate this scenario, the calculation of lake status for each grid cell is performed in two steps. First, the data are grouped according to the sign of $E_{potential}$-P_L value and ranked according to their relative values within each group. The data are then re-

ranked so that negative values always exceed positive values. This methodology properly sequences each value so that increases in precipitation or runoff or decreases in $E_{potential}$ will all provide a higher lake status estimate (Fig. S1).

Both the $E_{potential}$ and lake status variables are calculated independently of the data assimilation using the decadal means of the model prior which is the highest resolution publicly available for the runoff variable from TraCE. Higher resolution lake status values could be calculated by using P/E to approximate runoff as these variables are strongly correlated in regions not affected by ice-sheets (Fig. S2). However, decadal averages are sufficient for assimilating the low-resolution lake status values for this study. A sea-level mask and ice-sheet mask are also applied to remove grid cells where no lake is feasible for the prescribed boundary conditions. During the data assimilation, the PSM identifies the correct grid cell and age range for the proxy, and the values are re-ranked within the duration of the proxy record.

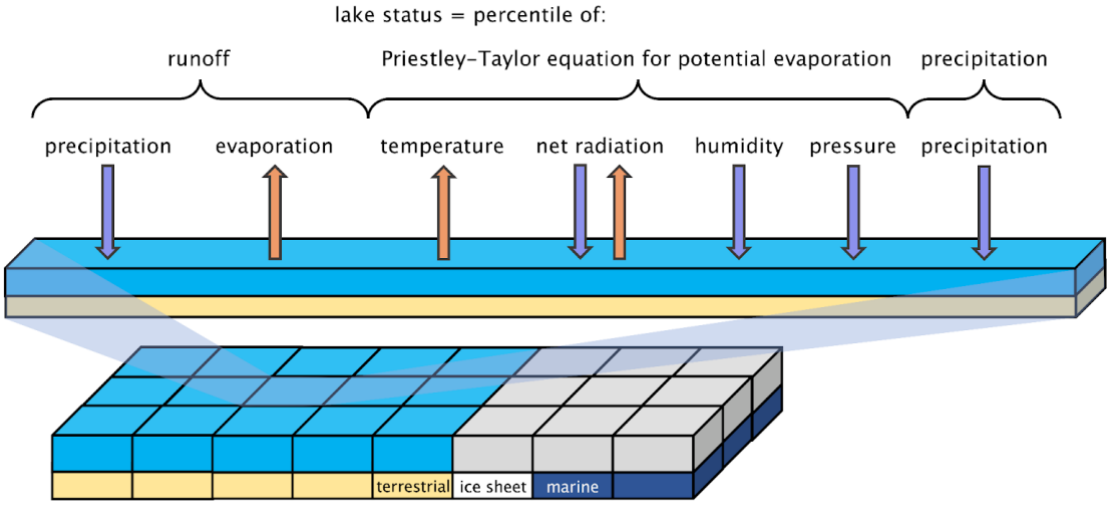


Figure 3: Conceptual diagram of our lake status PSM. Arrows indicate if an increase in the designated variable would result in greater input into the lake (down) or output from the lake (up). Each grid cell is calculated independently, and light gray shading shows that gridcells covered by ocean or substantial snow cover are masked.

2.5 Principal component analysis

To contextualize the results of the data assimilation through time, we evaluate annual precipitation changes using a principal component analysis (PCA). This technique has been used by previous data assimilation studies to synthesize reconstructed spatiotemporal climate variability (Luo et al., 2022; Okazaki and Yoshimura, 2017; Osman et al., 2021) and provides a means of examining how the spatial patterns differ between the prior and the reconstruction without the need to predefine regions. We calculate a global PC1 and PC2 timeseries (Dawson, 2016) for reconstructed mean annual terrestrial precipitation anomalies. The spatial pattern of the reconstruction is then evaluated using the covariance of precipitation values for each grid cell with each PC timeseries. We also calculate the covariance of precipitation values in the model prior with each reconstructed PC timeseries to identify where the two datasets differ.

3 Results and discussion

This study introduces a novel application of PSM and DA methodologies across various timescales, variables, and proxies not previously explored. First, we present results that validate these methodologies and show that the reconstruction is more skillful and than model simulations alone (Sect. 3.1; Sect. 3.2). Then we present timeslice results for the LGM (21 ka, Sect. 3.3) and the mid-Holocene (6 ka, Sect. 3.4), two periods of significant scientific interest commonly simulated by model intercomparison projects. Particular focus is placed on comparing how the reconstruction differs from the model prior. Spatial and temporal variability in the effects of precipitation and evaporation on lake level are also examined (Sect. 3.5). The spatial patterns of terrestrial precipitation through time are then evaluated using PCA in Sect. 3.6. Finally, climate dynamics associated with mid-Holocene anomalies in North America are discussed (Sect. 3.7).

3.1 Validation of the PSM

The PSM methodology used by DAMP-21ka represents a simplified version of previous water balance calculations made from output of paleoclimate GCM simulations (Lowry and Morrill, 2019; Morrill et al., 2019). To accommodate the available data and to efficiently assimilate the proxy data, lake evaporation was estimated by potential evaporation, and the data were converted to percentile units. This second change is necessary to compare the simulated values with the proxy values, but we evaluate the effects of substituting $E_{potential}$ for E_{lake} (Fig. 4).

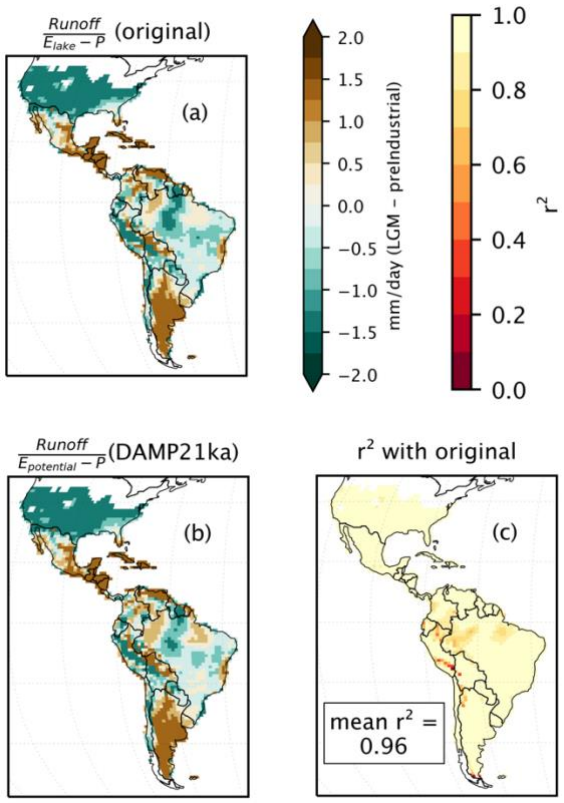


Figure 4: Comparison of lake status PSM equations. Lake status was calculated using runoff and precipitation data of CCSM4 simulations and different evaporation values. (a) Mean LGM minus preindustrial lake status values using lake evaporation data calculated from these simulations by Lowry and Morrill (2019). (b) The same values calculated using $E_{potential}$ as in this study (c) For each gridcell, an we calculate an r^2 value between the two methods using all 200 annual values. A scatterplot of these values is shown by Fig. S4.

Lowry and Morrill (2019) calculated lake evaporation from CMIP5 simulations of the LGM and preindustrial (100 years each) using the lake energy balance model of Morrill et al. (2004). The results of the CCSM4 simulations, which use a similar model to the TraCE simulation (CCSM3), were combined into a 200 year timeseries, and a series of experiments were performed to compare lake status calculations using the estimated E_{lake} and $E_{potential}$ values. The modified lake status values should be able to reproduce the different climates of the LGM and preindustrial as well as capture the interannual variability.

Potential evaporation skillfully reproduces the temporal variability and spatial patterns of LGM–PI lake evaporation ($r^2 = 0.87$; Fig. S3; Fig. S4a). By ranking the lake status as percentiles, the effect of using $E_{potential}$ is further reduced ($r^2 = 0.96$; Fig. S4a), and the spatial patterns of LGM minus preindustrial anomalies are largely unaffected by the simplified approach (Fig. 4). As a result, we conclude that the impact of substituting potential evaporation for lake evaporation in the lake status calculation is minimal, and that our lake status PSM effectively mimics the key characteristics of more sophisticated models.

3.2 Reconstruction skill

The DAMP-21ka reanalysis provides a more skillful reconstruction of paleo lake status variability than the model prior alone (Fig. 5). To quantify agreement between the reconstructed lake status and the proxy lakes, we calculate Pearson's correlation coefficients (r-values) for each proxy record using the mean r-value from iterations for which the record was withheld. Osman et al. (2021) measured skill using a variety of metrics including r^2, but we note that this value does not consider negative correlation coefficients and therefore could overestimate skill for either the prior or reconstruction by indicating agreement if two values are negatively correlated. The mean value of all sites was then calculated to provide an overall reconstruction r-value of 0.24. Although this moderate result provides room for improvement, the correlation is significant (p-value < 0.05) for 60% of records (mean p-value = 0.13), and it represents an increase from the same skill metric calculated using a composite HadCM and TraCE timeseries representative of the model prior (mean r-value of 0.09; mean p-value of 0.18; 47% significant). Furthermore, this increase is observed despite the challenges of a limited number of proxy records compared to other global data assimilation products (Erb et al., 2022; Osman et al., 2021) and the tendency for hydroclimate variables to produce lower correlation coefficient results than temperature reconstructions (Steiger et al. 2018). The two regions with the greatest improvement are East Africa and North America (Fig. 5c). The commonality between these areas is that both regions contain numerous lake status records and experienced notable proxy-model disagreement during the Holocene (Hancock et al., 2023). Updating the reanalysis in regions with fewer sites is limited because covariances naturally tend to decrease with distance (North et al., 2011), and the localization radius further reduces the Kalman gain from distant records (Valler et al., 2019). Due to both of these factors, the reconstruction more closely resembles the prior in regions with fewer records, and thus improvements are expected to be minimized.

Precipitation values are evaluated against calibrated pollen records using the same skill metric as for the lake status data (Fig. 5e-f). Overall, the mean correlation between the pollen proxies and the reconstruction (0.39) shows very little difference with that of the model simulations (0.39), but this may reflect the spatial bias of pollen data as regional skill patterns are similar to that of the lake status results. For many of these sites located in Europe and Asia, lake status skill improvement is also lower, possibly due to varied climate signals between nearby proxy records reflective of a low magnitude of change in precipitation. In North America, improvements are greater (Fig. S5), and we advise caution when interpreting precipitation values of the reanalysis outside of regions with skill improvement such as in North America and Africa.

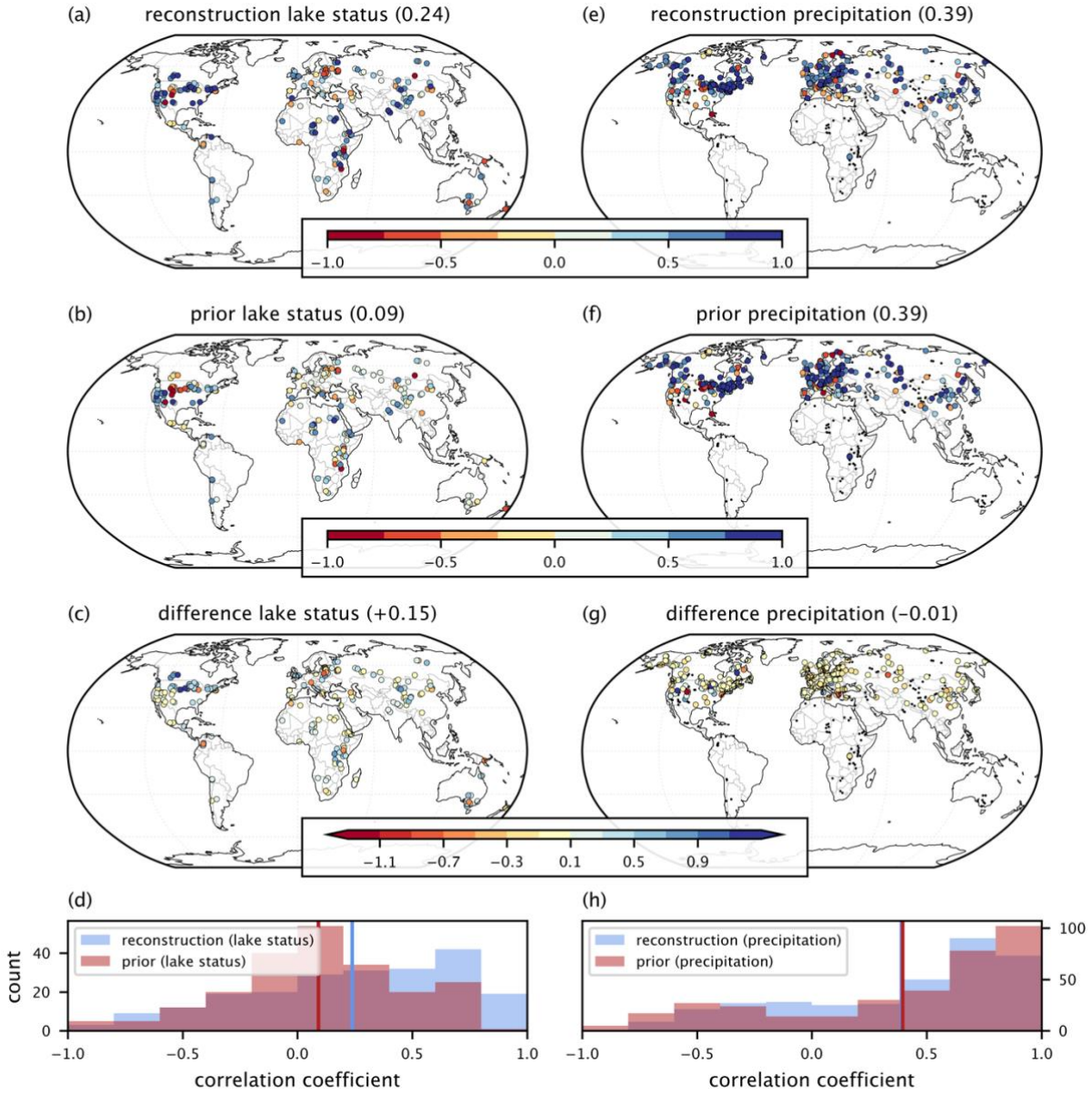


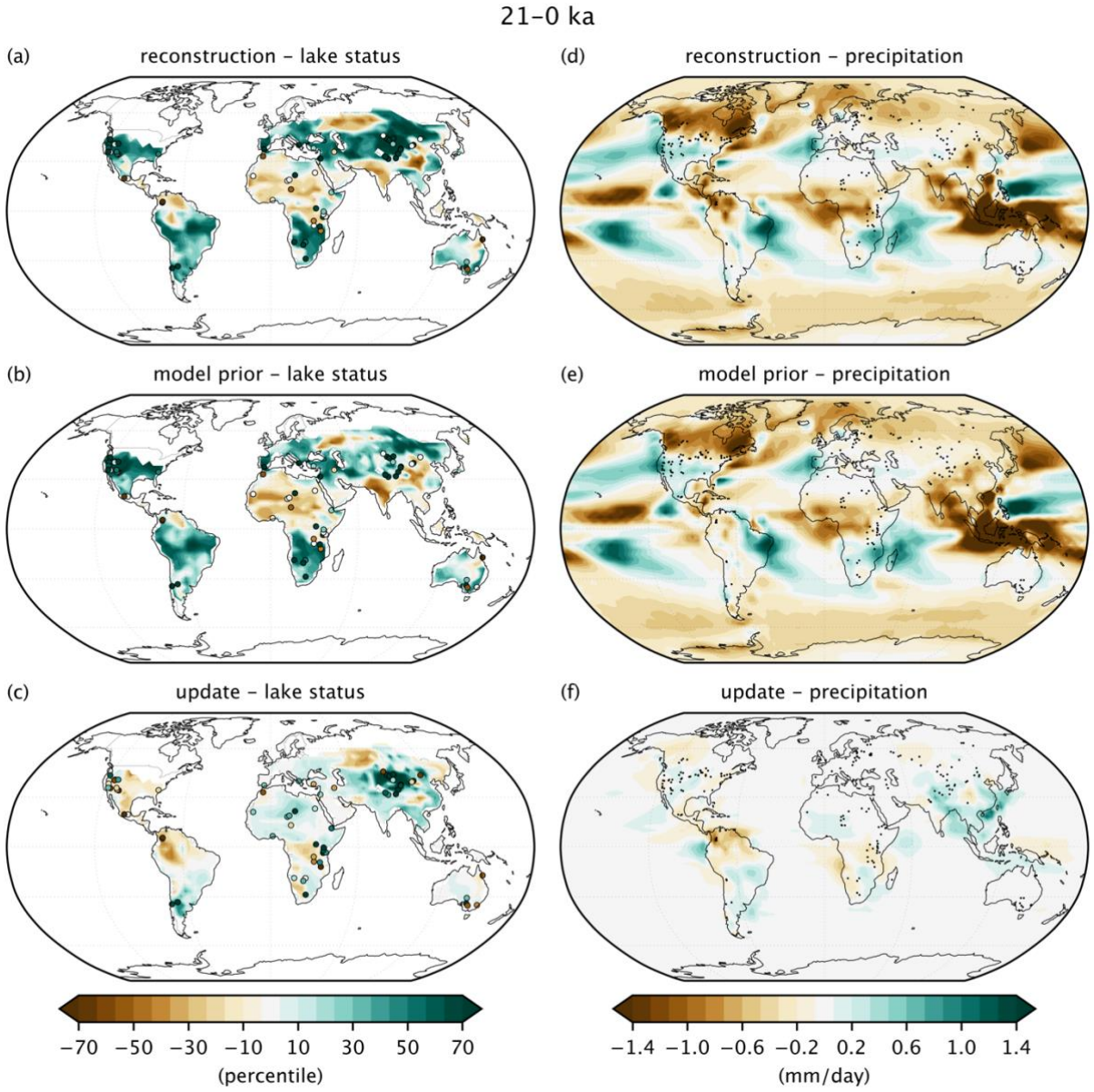
Figure 5: DAMP-21ka validation statistics. Pearson's correlation coefficients calculated for each proxy lake status (a-d) and pollen precipitation (e,h) record. The first row (a,e) shows the correlation between the reconstruction and the withheld proxy values. The second row (b,f) uses the prior values instead of the reconstruction. The third row (c,g) shows the difference so that blue colors indicate improvement, red indicates worsening agreement, and yellow represents no change between the prior and the DAMP-21ka reanalysis. The bottom row shows the histogram of the correlation coefficients for the prior (blue) and reconstruction (red) with the vertical line placed at the mean value which is also listed in the parenthesis of the panel titles.

3.3 Last Glacial Maximum (21 ka) anomalies

During the LGM, both thermodynamic and dynamic factors influenced large scale shifts in hydroclimate with regionally varying implications for lake levels (Lowry and Morrill, 2019). The globally reduced temperatures decreased evaporation and suppressed the hydrologic cycle (Kageyama et al., 2021), and ice-sheets redirected patterns of atmospheric circulation which increased precipitation amounts for portions of the Northern Hemisphere (Oster et al., 2015).

In the North American midlatitudes, large positive anomalies are simulated by the model prior, reflecting the significant gap between the LGM (21–20 ka) and last millennium (1–0 ka) which are estimated to be among the wettest and driest portions of the timeseries, respectively, for many grid cells (Fig. 6). By contrast, proxy lakes are more varied and lake status anomalies in the reanalysis are less positive than in the prior across most of the continent. Several proxy records also indicate that the peak high stands occurred later than what is simulated by models (Fu, 2023). Despite predominantly negative lake status innovations, more positive precipitation anomalies are reconstructed (Fig. 6c,f) for portions of this region which suggests that both temperature and precipitation are important for controlling lake status during this time.

Near the equator, dry precipitation values are reconstructed by DAMP-21ka during the LGM. Although this pattern is also true of the model prior, the update intensifies the dry anomalies across several regions including Central America, and East Africa. These decreases are consistent with the proxy-reconstructed lower lake status, but tropical locations that become wetter from the assimilation, such as eastern South America and southeast Asia, are largely unconstrained by proxy lakes. To investigate the reduced LGM precipitation values in the assimilation, the spatial patterns of the TraCE and HadCM simulations were compared to evaluate the precipitation prior (Fig. S6). This comparison reveals that small anomalies in the multi-model mean actually reflect disagreement between the two simulations. For example, wet anomalies in East Africa simulated by HadCM are offset by dry values from TraCE. The large range of climate states in the model prior increases the uncertainty in the prior, causing the assimilation to rely more heavily on the observations and produce a dry reconstruction consistent with both the lake status proxies and independent evidence from isotope records (Tierney and deMenocal, 2013). The results (and the proxy data) in East Africa support the TraCE simulation, but which model simulation best aligns with the assimilated results varies regionally. For example, in Central America and Australasia, the reanalysis produces anomalies which are more consistent with the HadCM simulation. As a result, the reconstruction produces negative precipitation anomalies which are more widespread throughout the tropics than either simulation on its own. This process also underscores the advantage of using a multi-model prior, which better captures structural uncertainties in the models and can better fit the observed values.



405 **Figure 6: Last Glacial Maximum hydroclimate anomalies. The top row shows the reconstructed mean lake status (a) and annual precipitation (d) anomalies for the Last Glacial Maximum (21–20 ka) subtracted by the last millennium (1–0 ka). (b,d) The same values for the model prior results are shown in the middle row. (c,f) The update describes the difference between these two datasets. The lake status column also shows the proxy values using the same color bar (a,b). For panel c, these values symbolize the difference between the proxy value and the mean value estimated by the prior from the PSM.**

410

3.4 Mid-Holocene (6 ka) anomalies

The mid-Holocene (~6,000 years BP) is a time period of significant scientific interest (Otto-Bliesner et al., 2017), in part because some features, such as a warmer Arctic, are common between 6 ka and modern warming (Serreze and Barry, 2011). Lake status in North America was lower at many proxy sites during the mid-Holocene (Liefert and Shuman, 2020). Both the TraCE and HadCM transient simulations show spatially variable hydroclimate anomalies, with drier conditions in the southern and central US and wetter conditions along the coastal regions (Hancock et al., 2023). After assimilating the available proxies, the gridded reconstruction indicates that both lower lake level and precipitation values were widespread throughout the entire continent (Fig. 7). Lakes throughout the Eurasian midlatitudes show a varied response to mid-Holocene climate and the prior does not differ significantly from the reconstruction. As a result, dry precipitation anomalies are observed throughout most of the terrestrial Northern Hemisphere midlatitudes which contrasts the model prior simulating varied wet (North America) and dry (Eurasia). These results suggest a uniform response throughout the midlatitudes to the orbital forcing of the mid-Holocene (Routson et al., 2019).

In the tropics, the African Humid Period characterizes a time of increased moisture and vegetation across North Africa which lasted from approximately 15 to 5 ka (deMenocal et al., 2000; Lézine et al., 2011; Shanahan et al., 2015). Both the proxy dataset and the model prior at 6 ka indicate wetter than Common Era anomalies for North America, but model simulations underestimate the magnitude of precipitation needed to sustain the proxy-inferred increase in vegetation during the early and mid-Holocene (e.g., Brierley et al., 2020). Furthermore, in the model prior, wet anomalies in the Sahara region are offset by drier conditions in Central and East Africa. Contrasting this spatial pattern, in the reanalysis, increased moisture in both North and East Africa are reconstructed in accordance with the innovation provided by the proxy lake status records. These results are in agreement with other isotopic and elemental abundance proxy results from East Africa which are independent of the reconstruction (Loomis et al., 2015; Russell and Johnson, 2005; Tierney and deMenocal, 2013). Thus, the African Humid Period may be characterized by both increased precipitation amounts and an increased spatial extent of positive moisture anomalies compared to model simulations.

Available lake records in the Southern Hemisphere are concentrated in Australia. Both lake proxies and models indicate a moisture dipole, with drier conditions in the northeast and wetter conditions in the southeast during the mid-Holocene. The reconstruction strengthens this moisture gradient and suggests that central Australia became wetter rather than drier at 6 ka. Results for the continental interior cannot be validated due to the lack of proxy records from the region (De Deckker, 2022), but the expanded wet anomalies are consistent with the coastal Swallow Lake (Barr et al., 2019).

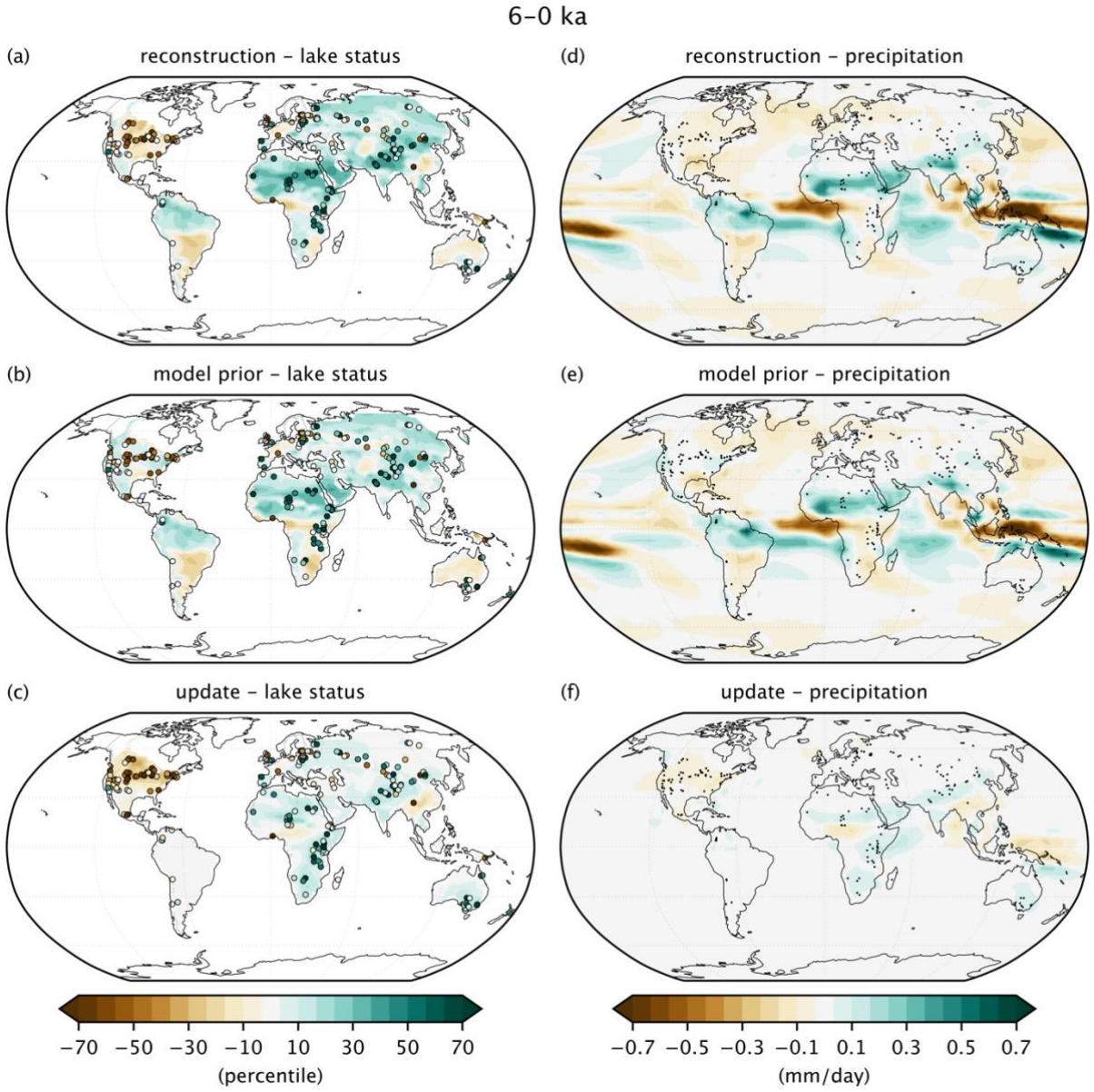


Figure 7. Mid-Holocene hydroclimate anomalies. The top row shows the reconstructed mean lake status (a) and annual precipitation (d) anomalies for the mid-Holocene (6.5–5.5 ka) subtracted by the last millennium (1–0 ka). (b,d) The same values for the model prior results are shown in the middle row. (c,f) The update describes the difference between these two datasets. The lake status column also shows the proxy values using the same color bar (a,b). For panel c, these values symbolize the difference between the proxy value and the mean value estimated by the prior from the PSM.

3.5 Spatial and temporal variability in the effects of precipitation and evaporation on lake level

The lake status PSM represents a balance between hydrologic inputs (precipitation and runoff) and outputs (lake/potential evaporation) within each grid cell. Here we explore the relationship between these variables from the TraCE simulations to better understand controls on lake-level variability and to contextualize the results of the assimilation. Sensitivity experiments were performed by comparing the lake status percentiles to similar values calculated if runoff and precipitation ($E_{varying}$) or runoff and potential evaporation ($P_{varying}$) remained constant. For both experiments runoff was set to a constant value because this variable incorporates both precipitation and evapotranspiration. Calculations were performed for the entire timeseries and separately for three segments representing the LGM (21–14 ka), deglacial (14–7 ka), and middle to late Holocene (7–0 ka) to explore non-stationary relationships among these variables.

Throughout the tropics, variability in precipitation dominates the simulated changes in lake level. This is shown by lower mean RMSE values for the $P_{varying}$ than the $E_{varying}$ experiments (Fig. 8). For example, in the Sahara, high error values in the $E_{varying}$ experiment reflect the importance of increased precipitation amounts during the African Humid Period for estimating lake status changes (Pausata et al., 2020; Specht et al., 2022). Elsewhere in the tropics, lower evaporative demand during the LGM are counter balanced with low precipitation anomalies which result in low lake status anomalies (Fig. 8a,d).

Outside of the tropics, the role of evaporation increases so that both variables become important for quantifying lake status variability. In the midlatitudes, RMSE values decrease for the $E_{varying}$ and increase for the $P_{varying}$ experiments which is consistent with previous lake status modeling of the LGM (Lowry and Morrill, 2019) and mid-Holocene (Li and Morrill, 2010). At high latitudes, high RMSE values are observed for both experiments which could reflect the positive correlation between precipitation and temperature (Badgeley et al., 2020; Hancock et al., 2023; Thomas et al., 2018). As evaporation and precipitation both increase, the water balance equation integrates these counterbalancing forces, but by only including one variable, errors increase by assuming a more significant decrease or increase in lake status.

The spatial patterns of lake status sensitivity remain relatively stable between the Holocene, deglacial, and LGM segments. However, several notable features are evident. A subtle poleward shift in RMSE values for the 7–0 ka experiments appears to have occurred in the midlatitudes (30–50°) relative to the earlier time periods. In the Arctic (>60°N) lower $E_{varying}$ RMSE values during the deglacial period also represent the emphasized role of lake evaporation during a time of increased temperature variability relative to precipitation (Badgeley et al., 2020). Overall, these results demonstrate that the simple PSM used for DAMP-21ka can account for the competing influence of temperature-driven PET and precipitation as these variables change spatially and throughout the duration of the reconstruction.

By using a multivariate PSM, the effects of both evaporation and precipitation are quantified by the Kalman gain. This may explain the spatially varied precipitation update for eastern North America during the LGM (Fig. 6f). Lower proxy lake status values are expressed in the reconstruction through temperature dependence on lake status at this time and the precipitation relationship is less meaningful. Accordingly, the LGM lake status innovation in eastern North America is reconstructed with warmer (less cold) temperature anomalies in central and eastern United States than those simulated by the

model prior (Fig. S7), and the resulting positive precipitation innovation reflects the positive covariance between precipitation and temperature.

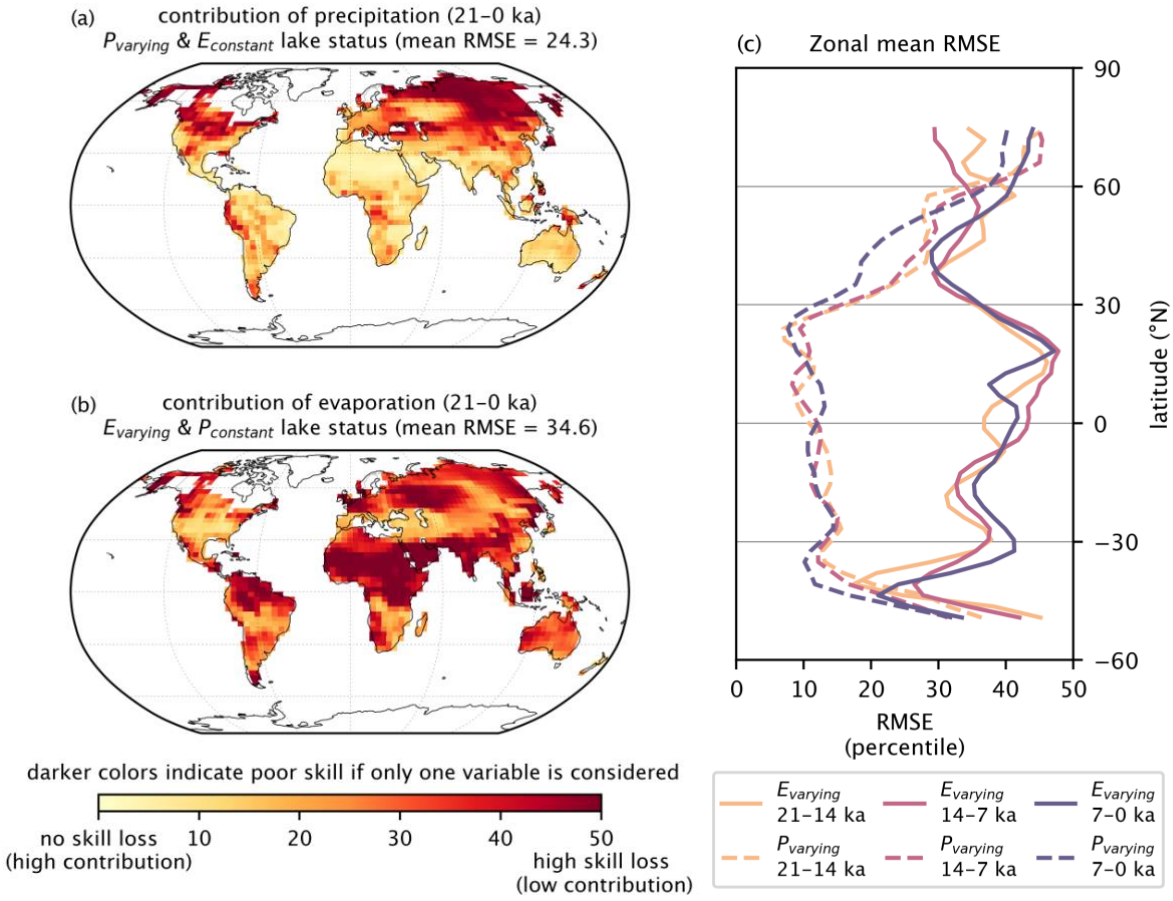


Figure 8. Lake status PSM sensitivity tests. Lake status values used in the assimilation were compared to similar values calculated using constant runoff and potential evaporation (a) or constant runoff and precipitation (b). Low RMSE values indicate areas where lake status can be accurately predicted using only the single variable. Panels a and b show values calculated for the entire timeseries. Panel c shows zonal mean RMSE values for sensitivity tests performed for three time periods.

3.6 Reconstructed PC1 and PC2

To examine spatiotemporal patterns of terrestrial precipitation, we use PCA to identify common modes of variability within the data without predefining geographic boundaries. Overall, the reconstruction and prior PCA results broadly agree, as expected given the role of the model's covariance structure in the assimilation process. Nonetheless, the spatiotemporal patterns revealed by the PCA, as well as several notable differences between the prior and reconstruction, provide insight into global hydroclimate evolution over the past 21 kyr (Fig. 9).

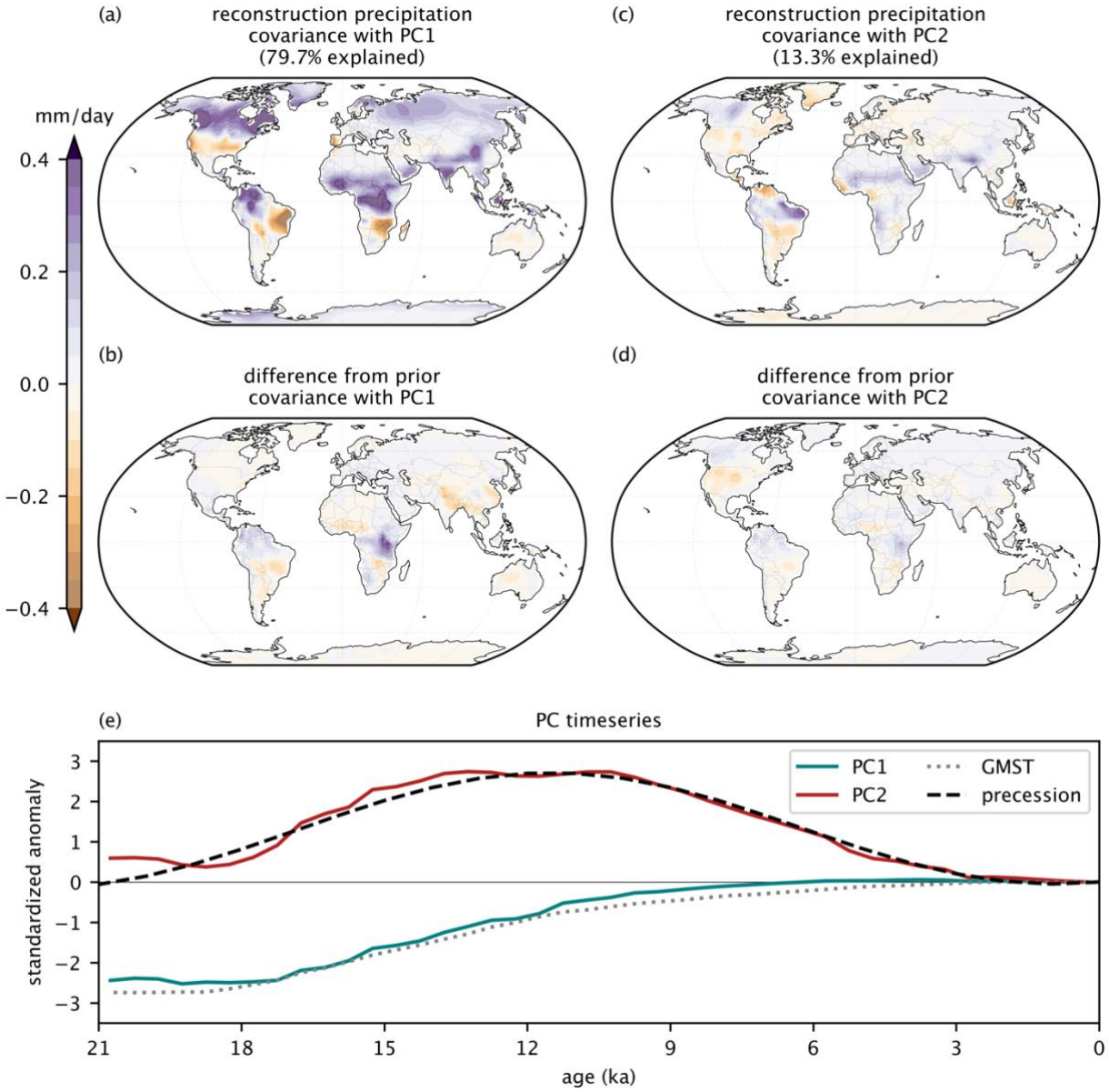


Figure 9. PCA of terrestrial annual precipitation (21–0 ka). in the reconstruction. The covariance between each grid cell and PC1 (a) and PC2 (c) is shown for the reconstruction. (b,d) The difference between panels a and c compared to the covariance calculated from the precipitation model prior with the reconstruction PC timeseries. (e) The PC timeseries are plotted (solid lines) along with the global mean surface temperature (GMST) of the reconstruction (black) and the precessional forcing (grey) (Berger and Loutre, 1991). The global mean timeseries are similar to those calculated for the model prior (Fig. S10). The timeseries values are standardized (z-scores) to allow for plotting on a common y axis.

PC1, which explains 80% of the variance in terrestrial precipitation (Fig. S8), characterizes the LGM to Holocene transition and integrates the thermodynamic and dynamical impacts of GMST warming and ice sheet retreat. Positive loadings at high latitudes are strongest over the Laurentide, Cordilleran, and Fennoscandian ice-sheets signifying the substantial increase in precipitation during the LGM-to-Holocene transition across these regions. Conversely, negative loadings in the midlatitudes, particularly in North America, correspond to diminishing precipitation amounts as circulation patterns adjust to the retreating ice-sheets (McGee et al., 2018; Oster et al., 2015). Although noteworthy, these patterns are not meaningfully distinct from those of the model prior (Fig. 9b; Fig. S9).

In tropical regions, the covariance of PC1 is consistent with a suppressed hydrological cycle and southward position of the ITCZ during the LGM (Gasse et al., 2008; Wang et al., 2023). The reconstruction differs from the prior most notably in East Africa where drier LGM anomalies for the reconstruction extend throughout both North and East Africa. This regional change is also apparent in both the 21 ka (Fig. 6) and, to a lesser extent, the 6 ka (Fig. 7) timeslices and is possibly attributable to the absence of vegetation feedback in the model prior (Dallmeyer et al., 2021).

Positive loadings for both PC1 and PC2 suggest overall wetting trends during the deglacial period throughout the Northern Hemisphere tropics, but PC2 also incorporates a drying trend from the mid-Holocene to the present which is strongly correlated with orbital precession. Differing PC2 loadings in North America suggest that regional disagreement is related to the climate's response to orbital forcing. Negative covariance in the reconstruction may indicate that the model prior, which shows greater precipitation amounts during the mid-Holocene, overestimates the role of precession for the North American midlatitudes. Therefore, long term trends in annual insolation and the strengthening of the annual latitudinal insolation gradient (Davis and Brewer, 2009; Routson et al., 2019) could be more significant than the seasonality suggested by precession alone.

3.7 Holocene hydroclimate in North America

Mid-Holocene hydroclimate anomalies in the North America midlatitudes reflect a well-established disagreement between generally dry proxies and wet models (Hermann et al., 2018; de Wet et al., 2023). In the DAMP-21ka precipitation reconstruction, the update results in dry annual precipitation outside of the monsoonal US Southwest. To investigate potential mechanisms for this divergence from the model prior, mid-Holocene climate anomalies in the TraCE and HadCM simulations are explored in this section.

When the latitudinal temperature gradient (LTG) is greater, zonal winds strengthen and more moisture is transported across the midlatitudes (Shaw et al., 2016). During the early and mid-Holocene, amplified warming in the Arctic reduced the Northern Hemisphere LTG which shifted the position and strength of the mid-latitude jet stream and associated storm tracks impacting North America hydroclimate (Routson et al., 2019; Xu et al., 2020a). Previous multi-model evaluations of the mid-Holocene show weaker westerly winds during summer months when the insolation anomalies were greatest (Routson et al., 2022). However, a more varied annual zonal wind strength response is noteworthy because the position and strength of the westerly jet during winter is the more significant control on annual precipitation anomalies for much of western North America (Hermann et al., 2018). Furthermore, a recent study exploring winter sea level pressure (SLP) gradients during the last

interglacial (~120 ka) indicates that a subset of models with the greatest proxy agreement simulate weaker winter SLP gradients across the Pacific (de Wet et al., 2023).

In the HadCM and TraCE simulations, mid-Holocene SLP anomalies show an intensified Aleutian Low and steeper gradient during winter and spring, relative to preindustrial conditions (Fig. 10). Correspondingly, the models also simulate strengthening westerly circulation in the upper atmosphere and increased precipitation amounts along the US West Coast. Because a majority of the region's precipitation occurs during winter, these changes represent substantial alteration to annual hydroclimate (Fig. S10; Hermann et al., 2018). Consequently, this covariance structure causes the dry western North America anomalies in DAMP-21ka during the mid-Holocene to suggest the opposite pattern: weaker winter midlatitude westerly flow and a shallower SLP gradient in the mid-Holocene. In support of this hypothesis, independent $\delta^{18}O$ (Anderson et al., 2005; Broadman et al., 2020; Jones et al., 2014; Nagashima et al., 2022), paleoenvironmental (Heusser et al., 1985; Nazarova et al., 2021), and marine (Barron and Anderson, 2011; Katsuki et al., 2009; Sun et al., 2021a) proxy records are interpreted to signify an intensification of the Aleutian Low throughout the Holocene.

A weaker winter SLP gradient during the mid-Holocene is also consistent with proxy-model comparisons for the last interglacial (de Wet et al., 2023) indicating a robust North Pacific and western North America response to amplified warming in the Arctic (Turney et al., 2020). Notably, simulations of modern climate change project an intensification of the wintertime Aleutian Low in response to future sea ice loss (Screen et al., 2018), and further research could provide additional insight into the precise mechanisms controlling the strength of the Aleutian Low in climate models during different time periods.

Mid-Holocene proxy-model disagreement in North America cannot be resolved by examining winter anomalies alone. In the central and eastern United States, simulated wet summer anomalies offset dry winter precipitation values to result in wet annual precipitation anomalies in the model prior. Accordingly, proxy-model disagreement may reflect winter sensitivity in the lake status proxies (Morrill et al., 2019) which are assumed to have annual seasonality for the PSM. Alternatively, the transient model prior, which simulates wetter conditions than the PMIP4 multi-model mean (Brierley et al., 2020; Hancock et al., 2023), could misrepresent the hydroclimate response of this region to mid-Holocene boundary conditions by incorrectly simulating the seasonality of precipitation anomalies so that dry winter and wet summer anomalies are not properly weighted in the annual mean.

Overall, the DAMP-21ka reanalysis reconstructs dry mid-Holocene precipitation anomalies throughout the North American midlatitudes, aligning the results with the coherent dry conditions inferred from both the assimilated and independent proxy data. This represents a key strength of the data assimilation methodology which incorporates the dynamics simulated by the model prior into a multivariate reconstruction constrained by the proxy observations. The reconstruction provides an improved understanding of past hydroclimate variability which can be used to benchmark future research and evaluate both proxy records and model simulations.

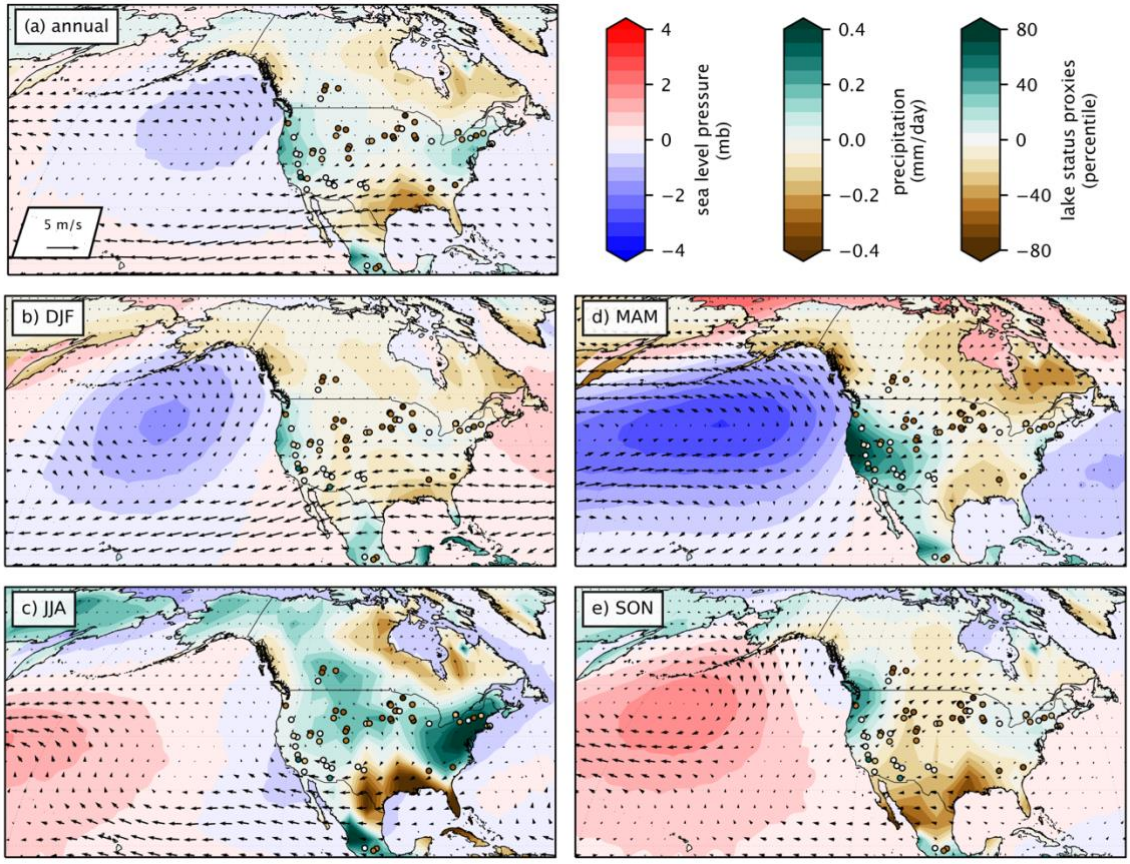


Figure 10. Simulated mid-Holocene climate anomalies for North America. Gridded values are representative of the model prior. Precipitation (green-brown) is shown for terrestrial regions and SLP (red-blue) is mapped over oceans. Wind vector anomalies for the 200 mb geopotential height are plotted with arrows. The proxy lake values used in the assimilation are plotted with circles and are the same for each panel. Anomalies are calculated as the mean of 6.5-5.5 ka minus 1-0 ka values for both the HadCM and TraCE simulations.

4 Conclusions

Data assimilation integrates the information from proxy observations and model simulations into a single gridded reconstruction of annual precipitation and lake status for the past 21,000 years, the first global application of the DA methodology to hydroclimate on both Holocene and LGM timescales. A mechanistic lake level PSM was developed to quantitatively compare model simulations with proxy records using runoff, precipitation, and potential evaporation. The multivariate PSM is capable of incorporating changing sensitivities to these variables across space and time, a critical component for midlatitude and arctic regions where both precipitation amount and evaporative demand are needed to properly estimate water balance variability.

The DAMP-21ka reanalysis more skillfuly estimates hydroclimate patterns than the model simulations alone, particularly in North America and East Africa where data density is high and proxy-model disagreement is substantial. Several notable changes are observed for the reconstruction compared to the model prior. During the LGM, tropical aridity intensified and was more geographically widespread than in either simulation alone. For the mid-Holocene, increased precipitation is reconstructed for monsoon regions across North Africa, East Africa and Asia. Outside of the tropics, dry precipitation anomalies are reconstructed throughout the midlatitudes in North America and Eurasia which suggests a uniform response to orbital forcing in the midlatitudes. As a result, positive covariance in the model prior between North America and North Africa becomes negative in the DAMP-21ka reanalysis as constrained by the lake status proxies.

The reconstruction also reconciles the well described proxy-model disagreement in North America for the mid-Holocene. Investigation into the climate dynamics of the model prior identifies stronger SLP gradients in the North Pacific as a potential source of simulated wet anomalies in western North America. Consequently, dry mid-Holocene anomalies for DAMP-21ka are consistent with the weaker Aleutian Low recorded by marine and terrestrial proxies throughout the region, and aligns the results mid-Holocene with data-model comparisons for the last interglacial, indicating a robust response in the North Pacific across two periods used as partial analogues for modern climate change.

Discrepancies between the model prior and reconstruction about the magnitude of change in the Northern Hemisphere tropics, direction of change in North America, and covariance between regions have implications for future climate change projections simulated by GCMs. Although neither the LGM or mid-Holocene are perfect analogues for modern climate change, these results demonstrate the importance of spatial structure of temperature anomalies and boundary conditions on regional precipitation and water budgets. Therefore, accurate understanding of the patterns of temperature anomalies and the resulting dynamic and thermodynamic processes on regional circulation changes and evaporative demand is necessary for predicting the hydroclimate response to warming temperatures.

Data availability:

A NetCDF file including the DAMP-21ka reconstruction, the model prior, and proxy values will be made available on Zenodo prior to publication. It can be accessed now at https://github.com/clhancock/DAMP21ka/tree/main. The proxy lake level data compilation will also be made available in LiPD format at https://lipdverse.org/. A comprehensive list of lake level and pollen proxy records and their references is provided by Table S1 and Table S2 respectively.

Code availability:

The code used by this study is a modified version of that used by Erb et al. (2022) and is available at https://github.com/Holocene-Reconstruction/Holocene-code/tree/development_hydroclimate. The code used for the DAMP-21ka reconstruction will be archived on Zenodo prior to publication.

Author contribution:

CH performed the formal analysis and writing of the original draft. MPE and CH were involved in the programming. CH and NM were involved in the data curation. MPE, SD, and NM contributed to funding acquisition, supervision. CH, MPE, SD, and NM participated in the conceptualization and methodology. RI provided model output. All authors contributed to the review of the manuscript.

Acknowledgements:

The authors would like to acknowledge Paul Valdez and Lauren Gregoire, for providing data from the HadCM simulation used to help construct the multi-model prior. We also thank Carrie Morrill for providing the lake evaporation data used to evaluate the sensitivity of the lake status equation to different formulations.

Competing interests:

The authors declare that they have no conflict of interest.

References

Anderson, L., Abbott, M. B., Finney, B. P., and Burns, S. J.: Regional atmospheric circulation change in the North Pacific during the Holocene inferred from lacustrine carbonate oxygen isotopes, Yukon Territory, Canada, Quat. Res., 64, 21–35, https://doi.org/10.1016/j.yqres.2005.03.005, 2005.

Bach, E. and Ghil, M.: A multi-model ensemble Kalman filter for data assimilation and forecasting, J. Adv. Model. Earth Syst., 15, e2022MS003123, https://doi.org/10.1029/2022MS003123, 2023.

Badgeley, J. A., Steig, E. J., Hakim, G. J., and Fudge, T. J.: Greenland temperature and precipitation over the last 20,000 years using data assimilation, Clim. Past, 16, 1325–1346, https://doi.org/10.5194/cp-16-1325-2020, 2020.

Barr, C., Tibby, J., Leng, M. J., Tyler, J. J., Henderson, A. C. G., Overpeck, J. T., Simpson, G. L., Cole, J. E., Phipps, S. J., Marshall, J. C., McGregor, G. B., Hua, Q., and McRobie, F. H.: Holocene El Niño–Southern Oscillation variability reflected in subtropical Australian precipitation, Sci. Rep., 9, 1627, https://doi.org/10.1038/s41598-019-38626-3, 2019.

Barron, J. A. and Anderson, L.: Enhanced Late Holocene ENSO/PDO expression along the margins of the eastern North Pacific, Quat. Int., 235, 3–12, https://doi.org/10.1016/j.quaint.2010.02.026, 2011.

Bartlein, P. J., Anderson, K. H., Anderson, P. M., Edwards, M. E., Mock, C. J., Thompson, R. S., Webb, R. S., Webb III, T., and Whitlock, C.: Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data, Quat. Sci. Rev., 17, 549–585, 1998.

Benson, L. V., Currey, D. R., Dorn, R. I., Lajoie, K. R., Oviatt, C. G., Robinson, S. W., Smith, G. I., and Stine, S.: Chronology of expansion and contraction of four great Basin lake systems during the past 35,000 years, Palaeogeogr. Palaeoclimatol. Palaeoecol., 78, 241–286, https://doi.org/10.1016/0031-0182(90)90217-U, 1990.

Berger, A. and Loutre, M. F.: Insolation values for the climate of the last 10 million years, Quat. Sci. Rev., 10, 297–317, https://doi.org/10.1016/0277-3791(91)90033-Q, 1991.

Brady, E. C., Otto-Bliesner, B. L., Kay, J. E., and Rosenbloom, N.: Sensitivity to glacial forcing in the CCSM4, J. Clim., 26, 1901–1925, https://doi.org/10.1175/JCLI-D-11-00416.1, 2013.

Brierley, C. M., Zhao, A., Harrison, S. P., Braconnot, P., Williams, C. J. R., Thornalley, D. J. R., Shi, X., Peterschmitt, J.-Y., Ohgaito, R., Kaufman, D. S., Kageyama, M., Hargreaves, J. C., Erb, M. P., Emile-Geay, J., D'Agostino, R., Chandan, D., Carré, M., Bartlein, P. J., Zheng, W., Zhang, Z., Zhang, Q., Yang, H., Volodin, E. M., Tomas, R. A., Routson, C., Peltier, W. R., Otto-Bliesner, B., Morozova, P. A., McKay, N. P., Lohmann, G., Legrande, A. N., Guo, C., Cao, J., Brady, E., Annan, J. D., and Abe-Ouchi, A.: Large-scale features and evaluation of the PMIP4-CMIP6 midHolocene simulations, Clim. Past, 16, 1847–1872, https://doi.org/10.5194/cp-16-1847-2020, 2020.

Briggs, R. W., Wesnousky, S. G., and Adams, K. D.: Late Pleistocene and late Holocene lake highstands in the Pyramid Lake subbasin of Lake Lahontan, Nevada, USA, Quat. Res., 64, 257–263, https://doi.org/10.1016/j.yqres.2005.02.011, 2005.

Broadman, E., Kaufman, D. S., Henderson, A. C. G., Malmierca-Vallet, I., Leng, M. J., and Lacey, J. H.: Coupled impacts of sea ice variability and North Pacific atmospheric circulation on Holocene hydroclimate in Arctic Alaska, Proc. Natl. Acad. Sci., 117, 33034–33042, https://doi.org/10.1073/pnas.2016544117, 2020.

Broecker, W. S. and Orr, P. C.: Radiocarbon chronology of Lake Lahontan and Lake Bonneville, GSA Bull., 69, 1009–1032, https://doi.org/10.1130/0016-7606(1958)69[1009:RCOLLA]2.0.CO;2, 1958.

Clerke, L.: Hydrological regime of Australian lakes over the Late-Quaternary and Holocene, Doctoral dissertation, Macquarie University, Sydney, Australia, https://doi.org/10.25949/22662253.v1, 2023.

Cheng, H., Sinha, A., Wang, X., Cruz, F. W., and Edwards, R. L.: The global paleomonsoon as seen through speleothem records from Asia and the Americas, Clim. Dyn., 39, 1045–1062, https://doi.org/10.1007/s00382-012-1363-7, 2012.

Chevalier, M., Brewer, S., and Chase, B. M.: Qualitative assessment of PMIP3 rainfall simulations across the eastern African monsoon domains during the mid-Holocene and the Last Glacial Maximum, Quat. Sci. Rev., 156, 107–120, https://doi.org/10.1016/j.quascirev.2016.11.028, 2017.

Coe, M. and Harrison, S.: The water balance of northern Africa during the mid-Holocene: an evaluation of the 6 ka BP PMIP simulations, Clim. Dyn., 19, 155–166, https://doi.org/10.1007/s00382-001-0219-3, 2002.

COHMAP Members: Climatic changes of the last 18,000 years: observations and model simulations, Science, 241, 1043–1052, https://doi.org/10.1126/science.241.4869.1043, 1988.

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. D., and Smith, R. D.: The Community Climate System Model version 3 (CCSM3), J. Clim., 19, 2122–2143, https://doi.org/10.1175/JCLI3761.1, 2006.

Cook, B. I., Smerdon, J. E., Cook, E. R., Williams, A. P., Anchukaitis, K. J., Mankin, J. S., Allen, K., Andreu-Hayles, L., Ault, T. R., Belmecheri, S., Coats, S., Coulthard, B., Fosu, B., Grierson, P., Griffin, D., Herrera, D. A., Ionita, M., Lehner, F., Leland, C., Marvel, K., Morales, M. S., Mishra, V., Ngoma, J., Nguyen, H. T. T., O'Donnell, A., Palmer, J., Rao, M. P., Rodriguez-Caton, M., Seager, R., Stahle, D. W., Stevenson, S., Thapa, U. K., Varuolo-Clarke, A. M., and Wise, E. K.: Megadroughts in the Common Era and the Anthropocene, Nat. Rev. Earth Environ., 3, 741–757, https://doi.org/10.1038/s43017-022-00329-1, 2022.

Dallmeyer, A., Claussen, M., Lorenz, S. J., Sigl, M., Toohey, M., and Herzschuh, U.: Holocene vegetation transitions and their climatic drivers in MPI-ESM1.2, Clim. Past, 17, 2481–2513, https://doi.org/10.5194/cp-17-2481-2021, 2021.

Davis, B. A. S. and Brewer, S.: Orbital forcing and role of the latitudinal insolation/temperature gradient, Clim. Dyn., 32, 143–165, https://doi.org/10.1007/s00382-008-0480-9, 2009.

Dawson, A.: eofs: A library for EOF Analysis of meteorological, oceanographic, and climate data, 4, e14, https://doi.org/10.5334/jors.122, 2016.

Dayem, K. E., Molnar, P., Battisti, D. S., and Roe, G. H.: Lessons learned from oxygen isotopes in modern precipitation applied to interpretation of speleothem records of paleoclimate from eastern Asia, Earth Planet. Sci. Lett., 295, 219–230, https://doi.org/10.1016/j.epsl.2010.04.003, 2010.

De Cort, G., Chevalier, M., Burrough, S. L., Chen, C. Y., Harrison, S. P.: An uncertainty-focused database approach to extract spatiotemporal trends from qualitative and discontinuous lake-status histories. Quat. Sci. Rev., 258, 106870, https://doi.org/10.1016/j.quascirev.2021.106870, 2021.

De Deckker, P.: The Holocene hypsithermal in the Australian region, Quat. Sci. Adv., 7, 100061, https://doi.org/10.1016/j.qsa.2022.100061, 2022.

Dee, S. G., Emile-Geay, J., Evans, M. N., Allam, A., Steig, E. J., and Thompson, D. M.: PRYSM: An open-source framework for PRoxY System Modeling, with applications to oxygen-isotope systems, J. Adv. Model. Earth Syst., 7, 1220–1247, https://doi.org/10.1002/2015MS000447, 2015a.

Dee, S. G., Noone, D., Buenning, N., Emile-Geay, J., and Zhou, Y.: SPEEDY-IER: A fast atmospheric GCM with water isotope physics, J. Geophys. Res. Atmospheres, 120, 73–91, https://doi.org/10.1002/2014JD022194, 2015b.

Dee, S. G., Steiger, N. J., Emile-Geay, J., and Hakim, G. J.: On the utility of proxy system models for estimating climate states over the common era, J. Adv. Model. Earth Syst., 8, 1164–1179, https://doi.org/10.1002/2016MS000677, 2016.

Deininger, M., McDermott, F., Cruz, F. W., Bernal, J. P., Mudelsee, M., Vonhof, H., Millo, C., Spötl, C., Treble, P. C., Pickering, R., and Scholz, D.: Inter-hemispheric synchroneity of Holocene precipitation anomalies controlled by Earth's latitudinal insolation gradients, Nat. Commun., 11, 5447, https://doi.org/10.1038/s41467-020-19021-3, 2020.

deMenocal, P., Ortiz, J., Guilderson, T., Adkins, J., Sarnthein, M., Baker, L., and Yarusinsky, M.: Abrupt onset and termination of the African Humid Period:, Quat. Sci. Rev., 19, 347–361, https://doi.org/10.1016/S0277-3791(99)00081-5, 2000.

Erb, M. P., McKay, N. P., Steiger, N., Dee, S., Hancock, C., Ivanovic, R. F., Gregoire, L. J., and Valdes, P.: Reconstructing Holocene temperatures in time and space using paleoclimate data assimilation, Clim. Past, 18, 2599–2629, https://doi.org/10.5194/cp-18-2599-2022, 2022.

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J.: Applications of proxy system modeling in high resolution paleoclimatology, Quat. Sci. Rev., 76, 16–28, https://doi.org/10.1016/j.quascirev.2013.05.024, 2013.

Fu, M.: Revisiting Western United States Hydroclimate During the Last Deglaciation, Geophys. Res. Lett., 50, e2022GL101997, https://doi.org/10.1029/2022GL101997, 2023.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Q. J. R. Meteorol. Soc., 125, 723–757, https://doi.org/10.1002/qj.49712555417, 1999.

Gasse, F., Chalié, F., Vincens, A., Williams, M. A. J., and Williamson, D.: Climatic patterns in equatorial and southern Africa from 30,000 to 10,000 years ago reconstructed from terrestrial and near-shore proxy data, Quat. Sci. Rev., 27, 2316–2340, https://doi.org/10.1016/j.quascirev.2008.08.027, 2008.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., and Zhang, M.: The Community Climate System Model Version 4, J. Clim., 24, 4973–4991, https://doi.org/10.1175/2011JCLI4083.1, 2011.

Gregoire, L. J., Ivanovic, R. F., Maycock, A. C., Valdes, P. J., and Stevenson, S.: Holocene lowering of the Laurentide ice sheet affects North Atlantic gyre circulation and climate, Clim. Dyn., 51, 3797–3813, https://doi.org/10.1007/s00382-018-4111-9, 2018.

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N., and Perkins, W. A.: The last millennium climate reanalysis project: Framework and first results, J. Geophys. Res. Atmospheres, 121, 6745–6764, https://doi.org/10.1002/2016JD024751, 2016.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, Mon. Weather Rev., 129, 2776, 2001.

Hancock, C. L., McKay, N. P., Erb, M. P., Kaufman, D. S., Routson, C. R., Ivanovic, R. F., Gregoire, L. J., and Valdes, P.: Global synthesis of regional Holocene hydroclimate variability using proxy and model data, Paleoceanogr. Paleoclimatology, 38, e2022PA004597, https://doi.org/10.1029/2022PA004597, 2023.

Harrison, S. P., Kutzbach, J. E., Liu, Z., Bartlein, P. J., Otto-Bliesner, B., Muhs, D., Prentice, I. C., and Thompson, R. S.: Mid-Holocene climates of the Americas: a dynamical response to changed seasonality. Clim. Dyn., 20, 663-688, doi:10.1007/s00382-002-0300-6, 2003.

Harrison, S. P., Bartlein, P. J., Brewer, S., Prentice, I. C., Boyd, M., Hessler, I., Holmgren, K., Izumi, K., and Willis, K.: Climate model benchmarking with glacial and mid-Holocene climates, Clim. Dyn., 43, 671–688, https://doi.org/10.1007/s00382-013-1922-6, 2014.

Harrison, S. P., Bartlein, P. J., Izumi, K., Li, G., Annan, J., Hargreaves, J., Braconnot, P., and Kageyama, M.: Evaluation of CMIP5 palaeo-simulations to improve climate projections, Nat. Clim. Change, 5, 735–743, https://doi.org/10.1038/nclimate2649, 2015.

He, C., Liu, Z., Otto-Bliesner, B. L., Brady, E. C., Zhu, C., Tomas, R., Clark, P. U., Zhu, J., Jahn, A., Gu, S., Zhang, J., Nusbaumer, J., Noone, D., Cheng, H., Wang, Y., Yan, M., and Bao, Y.: Hydroclimate footprint of pan-Asian monsoon water isotope during the last deglaciation, Sci. Adv., 7, eabe2611, https://doi.org/10.1126/sciadv.abe2611, 2021.

Hermann, N. W., Oster, J. L., and Ibarra, D. E.: Spatial patterns and driving mechanisms of mid-Holocene hydroclimate in western North America, J. Quat. Sci., 33, 421–434, https://doi.org/10.1002/jqs.3023, 2018.

Herzschuh, U., Böhmer, T., Li, C., Cao, X., Hébert, R., Dallmeyer, A., Telford, R. J., and Kruse, S.: Reversals in temperature-precipitation correlations in the Northern Hemisphere extratropics During the Holocene, Geophys. Res. Lett., 49, e2022GL099730, https://doi.org/10.1029/2022GL099730, 2022.

Herzschuh, U., Böhmer, T., Li, C., Chevalier, M., Hébert, R., Dallmeyer, A., Cao, X., Bigelow, N. H., Nazarova, L., Novenko, E. Y., Park, J., Peyron, O., Rudaya, N. A., Schlütz, F., Shumilovskikh, L. S., Tarasov, P. E., Wang, Y., Wen, R., Xu, Q., and Zheng, Z.: LegacyClimate 1.0: a dataset of pollen-based climate reconstructions from 2594 Northern Hemisphere sites covering the last 30 kyr and beyond, Earth Syst. Sci. Data, 15, 2235–2258, https://doi.org/10.5194/essd-15-2235-2023, 2023.

Heusser, C. J., Heusser, L. E., and Peteet, D. M.: Late-Quaternary climatic change on the American North Pacific coast, Nature, 315, 485–487, https://doi.org/10.1038/315485a0, 1985.

Jiang, M., Han, Z., Li, X., Wang, Y., Stevens, T., Cheng, J., Lv, C., Zhou, Y., Yang, Q., Xu, Z., Yi, S., and Lu, H.: Beach ridges of Dali Lake in Inner Mongolia reveal precipitation variation during the Holocene, J. Quat. Sci., 35, 716–725, https://doi.org/10.1002/jqs.3195, 2020.

Jolly, D., Harrison, S. P., Damnati, B., and Bonnefille, R.: Simulated climate and biomes of Africa during the late quaternary: comparison with pollen and lake status data, Quat. Sci. Rev., 17, 629–657, https://doi.org/10.1016/S0277-3791(98)00015-8, 1998.

Jones, M. C., Wooller, M., and Peteet, D. M.: A deglacial and Holocene record of climate variability in south-central Alaska from stable oxygen isotopes and plant macrofossils in peat, Quat. Sci. Rev., 87, 1–11, https://doi.org/10.1016/j.quascirev.2013.12.025, 2014.

Kageyama, M., Harrison, S. P., Kapsch, M.-L., Lofverstrom, M., Lora, J. M., Mikolajewicz, U., Sherriff-Tadano, S., Vadsaria, T., Abe-Ouchi, A., Bouttes, N., Chandan, D., Gregoire, L. J., Ivanovic, R. F., Izumi, K., LeGrande, A. N., Lhardy, F., Lohmann, G., Morozova, P. A., Ohgaito, R., Paul, A., Peltier, W. R., Poulsen, C. J., Quiquet, A., Roche, D. M., Shi, X., Tierney, J. E., Valdes, P. J., Volodin, E., and Zhu, J.: The PMIP4 Last Glacial Maximum experiments: preliminary results and comparison with the PMIP3 simulations, Clim. Past, 17, 1065–1089, https://doi.org/10.5194/cp-17-1065-2021, 2021.

Katsuki, K., Khim, B.-K., Itaki, T., Harada, N., Sakai, H., Ikeda, T., Takahashi, K., Okazaki, Y., and Asahi, H.: Land—sea linkage of Holocene paleoclimate on the Southern Bering Continental Shelf, The Holocene, 19, 747–756, https://doi.org/10.1177/0959683609105298, 2009.

King, J., Anchukaitis, K. J., Allen, K., Vance, T., and Hessl, A.: Trends and variability in the Southern Annular Mode over the Common Era, Nat. Commun., 14, 2324, https://doi.org/10.1038/s41467-023-37643-1, 2023.

Lézine, A.-M., Hély, C., Grenier, C., Braconnot, P., and Krinner, G.: Sahara and Sahel vulnerability to climate changes, lessons from Holocene hydrological data, Quat. Sci. Rev., 30, 3001–3012, https://doi.org/10.1016/j.quascirev.2011.07.006, 2011.

Li, Y. and Morrill, C.: Multiple factors causing Holocene lake-level change in monsoonal and arid central Asia as identified by model experiments, Clim. Dyn., 35, 1119–1132, https://doi.org/10.1007/s00382-010-0861-8, 2010.

Li, Y. and Zhang, Y.: Synergy of the westerly winds and monsoons in the lake evolution of global closed basins since the Last Glacial Maximum and implications for hydrological change in central Asia, Clim. Past, 16, 2239–2254, https://doi.org/10.5194/cp-16-2239-2020, 2020.

Liefert, D. T. and Shuman, B. N.: Pervasive desiccation of North American lakes during the Late Quaternary, Geophys. Res. Lett., 47, e2019GL086412, https://doi.org/10.1029/2019GL086412, 2020.

Liu, Z., Otto-Bliesner, B. L., He, F., Brady, E. C., Tomas, R., Clark, P. U., Carlson, A. E., Lynch-Stieglitz, J., Curry, W., Brook, E., Erickson, D., Jacob, R., Kutzbach, J., and Cheng, J.: Transient simulation of last deglaciation with a new mechanism for Bølling-Allerød warming, Science, 325, 310–314, https://doi.org/10.1126/science.1171041, 2009.

Loomis, S. E., Russell, J. M., and Lamb, H. F.: Northeast African temperature variability since the Late Pleistocene, Palaeogeogr. Palaeoclimatol. Palaeoecol., 423, 80–90, https://doi.org/10.1016/j.palaeo.2015.02.005, 2015.

Lowry, D. P. and Morrill, C.: Is the Last Glacial Maximum a reverse analog for future hydroclimate changes in the Americas?, Clim. Dyn., 52, 4407–4427, https://doi.org/10.1007/s00382-018-4385-y, 2019.

Lu, F., Ma, C., Zhu, C., Lu, H., Zhang, X., Huang, K., Guo, T., Li, K., Li, L., Li, B., and Zhang, W.: Variability of East Asian summer monsoon precipitation during the Holocene and possible forcing mechanisms, Clim. Dyn., 52, 969–989, https://doi.org/10.1007/s00382-018-4175-6, 2019.

Lu, Z., Miller, P. A., Zhang, Q., Zhang, Q., Wårlind, D., Nieradzik, L., Sjolte, J., and Smith, B.: Dynamic vegetation simulations of the mid-Holocene green Sahara, Geophys. Res. Lett., 45, 8294–8303, https://doi.org/10.1029/2018GL079195, 2018.

Luo, X., Dee, S., Stevenson, S., Okumura, Y., Steiger, N., and Parsons, L.: Last Millennium ENSO diversity and North American teleconnections: new insights from paleoclimate data assimilation, Paleoceanogr. Paleoclimatology, 37, e2021PA004283, https://doi.org/10.1029/2021PA004283, 2022.

Mauri, A., Davis, B. A. S., Collins, P. M., and Kaplan, J. O.: The climate of Europe during the Holocene: a gridded pollen-based reconstruction and its multi-proxy evaluation, Quat. Sci. Rev., 112, 109–127, https://doi.org/10.1016/j.quascirev.2015.01.013, 2015.

McGee, D., Moreno-Chamarro, E., Marshall, J., and Galbraith, E. D.: Western U.S. lake expansions during Heinrich stadials linked to Pacific Hadley circulation, Sci. Adv., 4, eaav0118, https://doi.org/10.1126/sciadv.aav0118, 2018.

Miller, G. H., Alley, R. B., Brigham-Grette, J., Fitzpatrick, J. J., Polyak, L., Serreze, M. C., and White, J. W. C.: Arctic amplification: can the past constrain the future?, Quat. Sci. Rev., 29, 1779–1790, https://doi.org/10.1016/j.quascirev.2010.02.008, 2010.

Morrill, C.: The influence of Asian summer monsoon variability on the water balance of a Tibetan lake, J. Paleolimnol., 32, 273–286, https://doi.org/10.1023/B:JOPL.0000042918.18798.cb, 2004.

Morrill, C., Meador, E., Livneh, B., Liefert, D. T., and Shuman, B. N.: Quantitative model-data comparison of mid-Holocene lake-level change in the central Rocky Mountains, Clim. Dyn., 53, 1077–1094, https://doi.org/10.1007/s00382-019-04633-3, 2019.

Nagashima, K., Addison, J., Irino, T., Omori, T., Yoshimura, K., and Harada, N.: Aleutian Low variability for the last 7500 years and its relation to the Westerly Jet, Quat. Res., 108, 161–179, https://doi.org/10.1017/qua.2020.116, 2022.

Nazarova, L., Sachse, D., Fuchs, H. G. E., Dirksen, V., Dirksen, O., Syrykh, L., Razjigaeva, N. G., Rach, O., and Diekmann, B.: Holocene evolution of a proglacial lake in southern Kamchatka, Russian Far East, Boreas, 50, 1011–1026, https://doi.org/10.1111/bor.12554, 2021.

North, G. R., Wang, J., and Genton, M. G.: Xorrelation models for temperature fields, J. Clim., 24, 5850–5862, https://doi.org/10.1175/2011JCLI4199.1, 2011.

Okazaki, A. and Yoshimura, K.: Development and evaluation of a system of proxy data assimilation for paleoclimate reconstruction, Clim. Past, 13, 379–393, https://doi.org/10.5194/cp-13-379-2017, 2017.

Osman, M. B., Tierney, J. E., Zhu, J., Tardif, R., Hakim, G. J., King, J., and Poulsen, C. J.: Globally resolved surface temperatures since the Last Glacial Maximum, Nature, 599, 239–244, https://doi.org/10.1038/s41586-021-03984-4, 2021.

Oster, J. L., Ibarra, D. E., Winnick, M. J., and Maher, K.: Steering of westerly storms over western North America at the Last Glacial Maximum, Nat. Geosci., 8, 201–205, https://doi.org/10.1038/ngeo2365, 2015.

Otto-Bliesner, B. L., Braconnot, P., Harrison, S. P., Lunt, D. J., Abe-Ouchi, A., Albani, S., Bartlein, P. J., Capron, E., Carlson, A. E., Dutton, A., Fischer, H., Goelzer, H., Govin, A., Haywood, A., Joos, F., LeGrande, A. N., Lipscomb, W. H., Lohmann, G., Mahowald, N., Nehrbass-Ahles, C., Pausata, F. S. R., Peterschmitt, J.-Y., Phipps, S. J., Renssen, H., and Zhang, Q.: The PMIP4 contribution to CMIP6 – Part 2: Two interglacials, scientific objective and experimental design for Holocene and Last Interglacial simulations, Geosci. Model Dev., 10, 3979–4003, https://doi.org/10.5194/gmd-10-3979-2017, 2017.

Parsons, L. A., Amrhein, D. E., Sanchez, S. C., Tardif, R., Brennan, M. K., and Hakim, G. J.: Do multi-model ensembles improve reconstruction skill in paleoclimate data assimilation?, Earth Space Sci., 8, e2020EA001467, https://doi.org/10.1029/2020EA001467, 2021.

Pausata, F. S. R., Messori, G., and Zhang, Q.: Impacts of dust reduction on the northward expansion of the African monsoon during the Green Sahara period, Earth Planet. Sci. Lett., 434, 298–307, https://doi.org/10.1016/j.epsl.2015.11.049, 2016.

Pausata, F. S. R., Gaetani, M., Messori, G., Berg, A., Maia de Souza, D., Sage, R. F., and deMenocal, P. B.: The greening of the Sahara: past changes and future implications, One Earth, 2, 235–250, https://doi.org/10.1016/j.oneear.2020.03.002, 2020.

Prentice, I. C. and Jolly, D.: Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa, J. Biogeogr., 27, 507–519, https://doi.org/10.1046/j.1365-2699.2000.00425.x, 2000.

Priestley, C. H. B. and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, Mon. Weather Rev., 100, 81–92, https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Qin, B. J., Harrison S. P., Kutzback, J. E.: Evaluation of modelled regional water balance using lake status data: a comparison of 6ka simulations with the NCAR CCM, Quat. Sci. Rev., 17, 535-548, https://doi.org/10.1016/S0277-3791(98)00011-0, 1998.

Qin, B. and Yu, G.: Implications of lake level variations at 6 ka and 18 ka in mainland Asia, Glob. Planet. Change, 18, 59–72, https://doi.org/10.1016/S0921-8181(98)00036-8, 1998.

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, Commun. Earth Environ., 3, 1–10, https://doi.org/10.1038/s43247-022-00498-3, 2022.

Routson, C. C., McKay, N. P., Kaufman, D. S., Erb, M. P., Goosse, H., Shuman, B. N., Rodysill, J. R., and Ault, T.: Mid-latitude net precipitation decreased with Arctic warming during the Holocene, Nature, 568, 83–87, https://doi.org/10.1038/s41586-019-1060-3, 2019.

Routson, C. C., Erb, M. P., and McKay, N. P.: High latitude modulation of the holocene North American monsoon, Geophys. Res. Lett., 49, e2022GL099772, https://doi.org/10.1029/2022GL099772, 2022.

Russell, J. M. and Johnson, T. C.: Late Holocene climate change in the North Atlantic and equatorial Africa: Millennial-scale ITCZ migration, Geophys. Res. Lett., 32, https://doi.org/10.1029/2005GL023295, 2005.

Screen, J. A., Deser, C., Smith, D. M., Zhang, X., Blackport, R., Kushner, P. J., Oudar, T., McCusker, K. E., and Sun, L.: Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models, Nat. Geosci., 11, 155–163, https://doi.org/10.1038/s41561-018-0059-y, 2018.

Serreze, M. C. and Barry, R. G.: Processes and impacts of Arctic amplification: A research synthesis, Glob. Planet. Change, 77, 85–96, https://doi.org/10.1016/j.gloplacha.2011.03.004, 2011.

Shanahan, T. M., Overpeck, J. T., Wheeler, C. W., Beck, J. W., Pigati, J. S., Talbot, M. R., Scholz, C. A., Peck, J., and King, J. W.: Paleoclimatic variations in West Africa from a record of late Pleistocene and Holocene lake level stands of Lake Bosumtwi, Ghana, Palaeogeogr. Palaeoclimatol. Palaeoecol., 242, 287–302, https://doi.org/10.1016/j.palaeo.2006.06.007, 2006.

Shanahan, T. M., McKay, N. P., Hughen, K. A., Overpeck, J. T., Otto-Bliesner, B., Heil, C. W., King, J., Scholz, C. A., and Peck, J.: The time-transgressive termination of the African Humid Period, Nat. Geosci., 8, 140–144, https://doi.org/10.1038/NGEO2329, 2015.

Shaw, T. A., Baldwin, M., Barnes, E. A., Caballero, R., Garfinkel, C. I., Hwang, Y.-T., Li, C., O'Gorman, P. A., Rivière, G., Simpson, I. R., and Voigt, A.: Storm track processes and the opposing influences of climate change, Nat. Geosci., 9, 656–664, https://doi.org/10.1038/ngeo2783, 2016.

Shuman, B. N., Routson, C., McKay, N., Fritz, S., Kaufman, D., Kirby, M. E., Nolan, C., Pederson, G. T., and St-Jacques, J.-M.: Placing the Common Era in a Holocene context: millennial to centennial patterns and trends in the hydroclimate of North America over the past 2000 years, Clim. Past, 14, 665–686, https://doi.org/10.5194/cp-14-665-2018, 2018.

Snoll, B., Ivanovic, R. F., Valdes, P. J., Maycock, A. C., and Gregoire, L. J.: Effect of orographic gravity wave drag on Northern Hemisphere climate in transient simulations of the last deglaciation, Clim. Dyn., 59, 2067–2079, https://doi.org/10.1007/s00382-022-06196-2, 2022.

Specht, N. F., Claussen, M., and Kleinen, T.: Simulated range of mid-Holocene precipitation changes from extended lakes and wetlands over North Africa, Clim. Past, 18, 1035–1046, https://doi.org/10.5194/cp-18-1035-2022, 2022.

Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S., and Roe, G. H.: Assimilation of time-averaged pseudoproxies for climate reconstruction, J. Clim., 27, 426–441, https://doi.org/10.1175/JCLI-D-12-00693.1, 2014.

Steiger, N. J., Smerdon, J. E., Cook, E. R., and Cook, B. I.: A reconstruction of global hydroclimate and dynamical variables over the Common Era, Sci. Data, 5, 180086, https://doi.org/10.1038/sdata.2018.86, 2018.

Street-Perrott, F. A., Marchand, D. S., Roberts, N., and Harrison, S. P.: Global lake-level variations from 18,000 to 0 years ago: A palaeoclimate analysis, Oxford Univ. (UK). Geography School, https://doi.org/10.2172/5609291, 1989.

Sun, Y., Xiao, W., Wang, R., Wu, L., and Wu, Y.: Changes in sediment provenance and ocean circulation on the northern slope of the Bering Sea since the last deglaciation, Mar. Geol., 436, 106492, https://doi.org/10.1016/j.margeo.2021.106492, 2021a.

Sun, Y., Wu, H., Kageyama, M., Ramstein, G., Li, L. Z. X., Tan, N., Lin, Y., Liu, B., Zheng, W., Zhang, W., Zou, L., and Zhou, T.: The contrasting effects of thermodynamic and dynamic processes on East Asian summer monsoon precipitation during the Last Glacial Maximum: a data-model comparison, Clim. Dyn., 56, 1303–1316, https://doi.org/10.1007/s00382-020-05533-7, 2021b.

Tarasov, P. E., Harrison, S. P., Saarse, L., Pushenko, M.Ya., Andreev, A. A., Aleshinskaya, Z. V., Davydova, N. N., Dorofeyuk, N. I., Efremov, Yu. V., Khomutova, V. I., Sevastyanov, D. V., Tamosaitis, J., Uspenskaya, O. N., Yakushko, O. F., Tarasova, I. V, Ya, M., Elina, G. A., Elovicheva, Ya. K., Filimonova, L. V., Gunova, V. S., Kvavadze, E. V., Nuestrueva, I. Yu., Pisareva, V. V., Shelekhova, T. S., Subetto, D. A., Zernitskaya, V.P .: Lake status records from the Former Soviet Union and Mongolia: documentation of the second version of the database, Paleoclimatology Publications Series Report #5, https://doi.org/10.25921/49ag-ck94, 1996.

Tarasov, P. E., Webb III, T., Andreev, A. A., Afanas'eva, N. B., Berezina, N. A., Bezusko, L. G., Blyakharchuk, T. A., Bolikhovskaya, N. S., Cheddadi, R., Chernavskaya, M. M., Chernova, G. M., Dorofeyuk, N. I., Dirksen, V. G., Elina, G. A., Filimonova, L. V., Glebov, F. Z., Guiot, J., Gunova, V. S., Harrison, S. P., Jolly, D., Khomutova, V. I., Kvavadze, E. V., Osipova, I. M., Panova, N. K., Prentice, I. C., Saarse, L., Sevastyanov, D. V., Volkova, V. S., and Zernitskaya, V. P.:

Present-day and mid-Holocene biomes reconstructed from pollen and plant macrofossil data from the former Soviet Union and Mongolia, J. Biogeogr., 25, 1029–1053, https://doi.org/10.1046/j.1365-2699.1998.00236.x, 1998.

Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay, J., Anderson, D. M., Steig, E. J., and Noone, D.: Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling, Clim. Past, 15, 1251–1273, https://doi.org/10.5194/cp-15-1251-2019, 2019.

Thomas, E. K., Castañeda, I. S., McKay, N. P., Briner, J. P., Salacup, J. M., Nguyen, K. Q., and Schweinsberg, A. D.: A wetter arctic coincident with hemispheric warming 8,000 years ago, Geophys. Res. Lett., 45, 10,637-10,647, https://doi.org/10.1029/2018GL079517, 2018.

Tierney, J. E. and deMenocal, P. B.: Abrupt shifts in Horn of Africa hydroclimate since the Last Glacial Maximum, Science, 342, 843–846, https://doi.org/10.1126/science.1240411, 2013.

Tierney, J. E., Lewis, S. C., Cook, B. I., LeGrande, A. N., and Schmidt, G. A.: Model, proxy and isotopic perspectives on the East African Humid Period, Earth Planet. Sci. Lett., 307, 103–112, https://doi.org/10.1016/j.epsl.2011.04.038, 2011.

Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., and Poulsen, C. J.: Glacial cooling and climate sensitivity revisited, Nature, 584, 569–573, https://doi.org/10.1038/s41586-020-2617-x, 2020.

Tierney, J. E., Zhu, J., Li, M., Ridgwell, A., Hakim, G. J., Poulsen, C. J., Whiteford, R. D. M., Rae, J. W. B., and Kump, L. R.: Spatial patterns of climate change across the Paleocene–Eocene Thermal Maximum, Proc. Natl. Acad. Sci., 119, e2205326119, https://doi.org/10.1073/pnas.2205326119, 2022.

Turney, C. S. M., Jones, R. T., McKay, N. P., van Sebille, E., Thomas, Z. A., Hillenbrand, C.-D., and Fogwill, C. J.: A global mean sea surface temperature dataset for the Last Interglacial (129–116 ka) and contribution of thermal expansion to sea level change, Earth Syst. Sci. Data, 12, 3341–3356, https://doi.org/10.5194/essd-12-3341-2020, 2020.

Valler, V., Franke, J., and Brönnimann, S.: Impact of different estimations of the background-error covariance matrix on climate reconstructions based on data assimilation, Clim. Past, 15, 1427–1441, https://doi.org/10.5194/cp-15-1427-2019, 2019.

Veloz, S. D., Williams, J. W., Blois, J. L., He, F., Otto-Bliesner, B., and Liu, Z.: No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models, Glob. Change Biol., 18, 1698–1713, https://doi.org/10.1111/j.1365-2486.2011.02635.x, 2012.

Vremec, M. and Collenteur, R.: PyEt: A Python package for estimating evaporation, https://doi.org/10.5281/zenodo.5896800, 2022.

Wang, B., Biasutti, M., Byrne, M. P., Castro, C., Chang, C.-P., Cook, K., Fu, R., Grimm, A. M., Ha, K.-J., Hendon, H., Kitoh, A., Krishnan, R., Lee, J.-Y., Li, J., Liu, J., Moise, A., Pascale, S., Roxy, M. K., Seth, A., Sui, C.-H., Turner, A., Yang, S., Yun, K.-S., Zhang, L., and Zhou, T.: Monsoons climate change assessment, Bull. Am. Meteorol. Soc., 102, E1–E19, https://doi.org/10.1175/BAMS-D-19-0335.1, 2021.

Wang, T., Wang, N., and Jiang, D.: Last Glacial Maximum ITCZ changes from pmip3/4 simulations, J. Geophys. Res. Atmospheres, 128, e2022JD038103, https://doi.org/10.1029/2022JD038103, 2023.

de Wet, C. B., Ibarra, D. E., Belanger, B. K., and Oster, J. L.: North American hydroclimate during past warms states: a proxy compilation-model comparison for the last interglacial and the mid-Holocene, Paleoceanogr. Paleoclimatology, 38, e2022PA004528, https://doi.org/10.1029/2022PA004528, 2023.

Williams, J. W. and Shuman, B.: Obtaining accurate and precise environmental reconstructions from the modern analog technique and North American surface pollen dataset, Quat. Sci. Rev., 27, 669–687, https://doi.org/10.1016/j.quascirev.2008.01.004, 2008.

Xu, C., Yan, M., Ning, L., and Liu, J.: Summer westerly jet in Northern Hemisphere during the mid-Holocene: a multi-model study, Atmosphere, 11, 1193, https://doi.org/10.3390/atmos11111193, 2020a.

Xu, H., Goldsmith, Y., Lan, J., Tan, L., Wang, X., Zhou, X., Cheng, J., Lang, Y., and Liu, C.: Juxtaposition of Western Pacific subtropical high on Asian summer monsoon shapes subtropical East Asian precipitation, Geophys. Res. Lett., 47, e2019GL084705, https://doi.org/10.1029/2019GL084705, 2020b.

Yu, G.: Lake status records from Europe: data base documentation, Publications Series Report #3, https://doi.org/10.25921/p5aq-0931, 1995.

Yu, G. Harrison, S.P. Xue, B.: Lake status records from China: data base documentation, MPI-BGC Tech Rep 4, 2001.

Yung, Y. L., Lee, T., Wang, C.-H., and Shieh, Y.-T.: Dust: A diagnostic of the hydrologic cycle during the Last Glacial Maximum, Science, 271, 962–963, https://doi.org/10.1126/science.271.5251.962, 1996.

Zhao, Y. and Harrison, S. P.: Mid-Holocene monsoons: a multi-model analysis of the inter-hemispheric differences in the responses to orbital forcing and ocean feedbacks, Clim. Dyn., 39, 1457–1487, https://doi.org/10.1007/s00382-011-1193-z, 2012.