Science Bulletin 70 (2025) 2014–2022

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Article

Model seasonal and proxy spatial biases revealed by assimilated mid-Holocene seasonal temperatures

Shuo Hao [a,b,1], Xu Zhang [a,c,1,*], Yanwu Duan [a,1], Evan J. Gowan [d], Jiang Zhu [e], Alexandre Cauquoin [f], Jie Chen [g], Martin Werner [h], Fahu Chen [a,b,g,*]

[a] Alpine Paleoecology and Human Adaptation Group, State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China
[b] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
[c] British Antarctic Survey, Cambridge, CB3 0ET, UK
[d] Department of Earth and Environmental Sciences, Kumamoto University, Kumamoto, 860-8555, Japan
[e] Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO 80305, USA
[f] Institute of Industrial Science (IIS), The University of Tokyo, Kashiwa, 277-0882, Japan
[g] Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China
[h] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, 27570, Germany

ARTICLE INFO

Article history:
Received 1 November 2024
Received in revised form 10 March 2025
Accepted 12 March 2025
Available online 15 March 2025

Keywords:
Holocene temperature conundrum
Paleoclimate data assimilation
Holocene seasonal temperature
Temperature regional heterogeneity
Climate model deficiencies
Proxy record distribution

ABSTRACT

Proxy-based reconstructions and climate model simulations of Holocene global annual mean temperatures exhibit divergent trends, leading to the well-known "Holocene temperature conundrum (HTC)". This discrepancy is most pronounced in the mid-to-high latitudes of the Northern Hemisphere (NH) and has been attributed to either proxy seasonal bias or deficiencies in climate models. Paleoclimate data assimilation (PDA), which integrates proxy records with climate model simulations, provides an advanced method for generating global seasonal temperature reanalysis datasets for the mid-Holocene (MH). Assimilated results indicate that MH Eurasian temperatures are largely independent of the choice of model priors and exhibit significant spatial heterogeneity. Compared to the pre-industrial (PI) period, the MH is characterized by winter and annual mean warming in Europe and high-latitude Eurasia, while the rest of Eurasia experiences cooling. However, this spatial heterogeneity is not well represented in model simulations due to a pronounced winter cooling bias at high latitudes, likely resulting from inadequate representations of vegetation and sea ice feedback mechanisms. As Eurasian proxy records are predominantly concentrated in Europe, this regional imbalance introduces a warm bias in reconstructed winter and annual temperatures intended to represent broader Eurasian temperature changes. These results suggest that the HTC may stem from both the uneven spatial distribution of proxy records and the incomplete representation of internal climate feedbacks in current models.

1. Introduction

Characterizing the temperature evolution throughout the current interglacial period (the Holocene, beginning 11,700 years ago) is essential for understanding the forcings and feedback mechanisms of the Earth system. This contextualizes post-industrial global warming and contributes to reducing uncertainties in future climate projections [1]. However, a long-standing debate persists regarding changes in global mean surface temperature (GMST) throughout the Holocene. Proxy-based reconstructions indicate an early to mid-Holocene (MH) thermal maximum, occurring after the final retreat of the Northern Hemisphere (NH) ice sheets and coinciding with peak NH summer insolation, followed by a cooling trend until the pre-industrial (PI) period [2,3]. In contrast, climate models suggest a long-term warming trend attributed to the retreat of the ice sheets and increasing concentrations of greenhouse gases [4]. This discrepancy between proxy data and climate models, or "Holocene temperature conundrum (HTC)", hinders a mechanistic understanding of Holocene climate change and undermines confidence in both proxy reconstructions and climate simulations [5,6].

* Corresponding authors.
E-mail addresses: xuang@bas.ac.uk (X. Zhang), fhchen@itpcas.ac.cn (F. Chen).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.scib.2025.03.039
2095-9273/

Seasonal bias in proxy reconstructions has been proposed as a significant factor contributing to the HTC [7]. Specifically, the reconstructed seasonal temperatures are contingent upon the methods employed to empirically convert proxy records into temperature estimates. This phenomenon has been evident in the winter temperatures derived from pollen records [8,9]. Additionally, there is significant variability in the simulated MH temperature seasonality among models that apply similar climate forcing [10,11], as well as within individual models that operate under different boundary conditions (e.g. vegetation types [12]). This variability introduces substantial uncertainty regarding seasonal temperature changes throughout the Holocene, thereby complicating efforts to resolve the HTC.

Paleoclimate data assimilation (PDA) facilitates the reconstruction of seasonal temperatures by integrating both the "observed" information derived from proxy records and the constraints imposed by climate model physics [13–15]. PDA offers dual advantages for reconstructing paleo-temperatures. Firstly, the approach is constrained by observations, which mitigates the effects of model deficiencies. Secondly, it is constrained by physical principles, thereby addressing the limitations associated with the uneven and restricted distribution of proxy records by extrapolating climate reconstructions to a global scale. Theoretically, the results obtained through assimilation should surpass those derived from either reconstruction or simulation alone, a hypothesis that has been supported by previous research [13,16,17]. However, PDA results are not without limitations. For instance, the effect of different model priors on PDA results remains ambiguous, raising critical questions regarding the reliability of the assimilated global mean temperature changes throughout the Holocene [18].

Here, we apply the PDA method by integrating global marine temperature records with two climate simulation outputs forced by different vegetation configurations during the MH epoch. This produces two reanalysis datasets depicting MH global annual and seasonal surface air temperature (SAT) anomalies relative to the late Holocene (LH). Our results indicate that, in only a few regions, such as Eurasia, the assimilated temperatures are largely independent of various model priors. Furthermore, comparing proxy reconstructions, model simulations and PDA results reveals a systematic winter cooling bias in climate models, particularly in Europe and high-latitude Eurasia, which leads to an overestimation of the spatial homogeneity of winter and annual mean SAT changes across Eurasia. Additionally, the assimilated results indicate significant spatial biases in the temperature changes reconstructed from Eurasian records, primarily due to the predominance of European records in representing Eurasian temperature changes.

2. Materials and methods

2.1. Assimilated marine records and forward models

The study incorporated global marine temperature records into the PDA process, while deliberately excluding terrestrial records. This exclusion is due to the two predominant frameworks utilized in current PDA methods: one framework assimilates temperature values reconstructed from proxy records, while the other directly incorporates raw proxy values. The former approach introduces an additional layer of uncertainty in the PDA results, stemming from calibration errors that occur during the conversion of proxy values into temperature values. Therefore, the latter framework was selected, which directly assimilates raw proxy values. This approach necessitates that the proxies possess forward models. Forward models are defined as linear or nonlinear regression models that accept various environmental variables as inputs (e.g., sea surface temperature (SST) and sea surface salinity (SSS)) and subsequently predict the corresponding proxy values. Currently, forward models for marine proxies are well-established, whereas the majority of terrestrial records lack reliable forward models. This limitation restricts the incorporation of terrestrial records within the proposed PDA framework. This study utilized a SST proxy dataset in Osman et al [13]. This dataset comprises four types of proxies: alkenone $U^{K'}_{37}$, the TetraEther index of 86 carbons (TEX_{86}), the elemental ratio of magnesium to calcium in planktic foraminifera (Mg/Ca), and the oxygen isotopic composition of planktic foraminifera ($\delta^{18}O$). These proxies were selected due to the availability of associated Bayesian forward models, which are critical to the proposed PDA scheme. The research objective is to reconstruct the MH (defined as 5.5–6.5 ka in this study) temperature anomaly field relative to the LH (defined as 0–1 ka in this study). A total of 260 and 170 proxy records were selected for the MH and LH periods, respectively (Fig. S1 online). When multiple data points from a single record were present within the MH or LH periods, these values were averaged. Importantly, proxy seasonality is explicitly accounted for in the proposed PDA framework based on these forward models. Each marine organism has a specific temperature range that optimally supports its survival. For example, certain organisms can only thrive during the warm season, resulting in proxies derived from these organisms reflecting warm-season temperatures (i.e., warm seasonal bias). However, within the proposed PDA framework, the Bayesian forward models are capable of selecting the average temperature of the months that fall within the optimal temperature range for the organisms (i.e., as provided by model priors) to predict the proxy values. These predicted values are then compared with the actual proxy values to update the model priors. As environmental temperatures fluctuate, the months selected will adjust accordingly, effectively mitigating the effect of proxy seasonal biases. Further details regarding the Bayesian forward models can be found in original studies [19–22].

2.2. Eurasian seasonal temperature records

This research conducted a comparative analysis of the assimilation results with independent seasonal temperature proxy records in Zhang et al. [23]. This multiproxy temperature dataset is both comprehensive and of high quality, comprising a total of 579 annual, 465 summer, and 258 winter temperature records. These records were selectively sourced from data-sharing initiatives (e.g., National Oceanic and Atmospheric Administration (NOAA) and Publishing Network for Geoscientific & Environmental Data (PANGAEA)), previous compilations (e.g., Temp12k database), and individual studies. Each record within the dataset underwent a rigorous evaluation and screening process to ensure quality. Only those records with sample resolutions of less than 400 years and covering periods exceeding 5000 years, including the Common Era, were retained. In this study, similar to the assimilated marine records, Eurasian terrestrial temperature records were specifically selected to encompass both the 0–1 ka BP and 5.5–6.5 ka BP periods. This selection yielded 113 records of annual mean temperature, 170 records of summer temperature, and 88 records of winter temperature. Consistent with existing studies [2,8,9], the research employed a spatial gridding technique with a resolution of $5^\circ \times 5^\circ$. In this approach, all proxy records within each $5^\circ \times 5^\circ$ grid were averaged to mitigate the disproportionate effect of densely clustered records in certain regions (Fig. S11 online). This method ensures a more robust and unbiased analysis.

2.3. Climate model simulations

The climate model priors utilized in the PDA process are derived from pre-existing iCESM1.2 simulations [12]. These priors include time-slice simulations for both the PI period and the MH epoch

(6 ka BP). For the MH, two snapshot simulations with distinct vegetation conditions were employed: one with PI vegetation cover (CTL scenario) and the other featuring greening in the Sahara and Arctic regions (VEG scenario). Specifically, the VEG scenario modifies the African Sahara by replacing bare ground desert with 100% shrub cover between approximately 10°N and 25°N, and with 100% C_4 grass between 25°N and 35°N. Additionally, it transforms the Arctic by substituting all C_3 grass north of 50°N with boreal forest. In addition to the iCESM model priors, MPI-ESM model priors were also used to evaluate the impact of different climate models on assimilation results. These MPI-ESM priors are sourced from pre-existing simulations using the MPI-ESM-wiso model [24]. Detailed descriptions of the simulation configurations and results are available in previous studies [12,24]. In a standard data assimilation, the time span of the prior ensemble typically aligns with that of the proxy records. For example, a millennial average proxy would be used to update a millennial average prior. However, due to the limited duration of simulated years available in the climate model (e.g., 800 years for the PI simulation, 400 years for the 6ka-CTL simulation, and 800 years for the 6ka-VEG simulation in the iCESM model), it is necessary to utilize a different time average for the model prior. Consequently, a 10-year average was selected as the longest feasible time interval while ensuring a sufficient number of ensemble members for assimilation (40 members). Previous investigations into the time averaging of model states indicated that once the average duration exceeds an interannual timescale, the patterns observed in the assimilated fields exhibit minimal sensitivity to the length of the average [16]. Therefore, the 10-year averaged model priors are considered a reasonable choice. For both the PI and MH simulations, this study randomly selected 400 years of modeling results and subsequently calculated a 10-year average, resulting in three 40-member prior ensembles (for PI, 6ka-CTL, and 6ka-VEG). These prior ensembles are regarded as representing the climatic states of the LH period and the MH period under the two vegetation conditions.

2.4. Paleoclimate data assimilation

The research utilized an offline ensemble square root Kalman filter method to conduct the PDA experiment, following the methodology described in Ref. [13,16]. For a comprehensive mathematical description, please consult these prior studies. Briefly, this method integrates simulated climate fields derived from climate models ("prior") with new information obtained from proxy observations ("innovation") to generate updated assimilated climate fields ("analysis"). The fundamental equation of this approach is expressed as follows:

$$\bar{X}_{analysis} = \bar{X}_{prior} + K\left(y_{obs} - \bar{Y}_{est}\right). \quad (1)$$

X_{prior} is an $N \times M$ matrix of prior climate states, where dimension N contains the model outputs for various variables (i.e., SST and SAT) concatenated as a vertical vector ("state vector"), and dimension M represents the number of state vector ensemble members (M = 40). The overbar in all cases denotes the average across the ensemble dimension.

$y_{obs} - \bar{Y}_{est}$ is the innovation, representing the novel information obtained from the observations. y_{obs} is a $P \times 1$ vector consisting of P proxy observations. Y_{est} is a $P \times M$ matrix containing the corresponding set of P proxy estimates, which are generated from the model output from each M state using Bayesian forward models at the same locations.

K is a $N \times P$ matrix as a weighting factor ("Kalman gain"), which can be determined as follows:

$$K = W_1 \odot cov\left(X_{prior}, Y_{est}\right) \times \left[W_2 \odot cov(Y_{est}, Y_{est}) + R\right]^{-1}, \quad (2)$$

where "cov" denotes the covariance expectation; $P \times P$ matrix R is the prescribed error covariance associated with each proxy observation. R is diagonal and user-defined. In this study, we use error estimates derived from Bayesian forward models, with distinct scaling factors applied to each proxy type, following the methodology in Osman et al. [13]. W_1 and W_2 are the localization radius weights; $\odot$ is the element-wise multiplication. Covariance localization was applied to the assimilation process to mitigate spurious correlations between distant regions. Following Ref. [13,25], a localization radius of 24,000 km was used, with localization weights determined by the distance-dependent Gaspari-Cohn function [25]. To compute the MH global annual and seasonal temperature anomaly fields relative to LH, $\bar{X}_{analysis}$ was calculated separately for both periods. For the LH, the assimilation process proceeded as follows: first, the 10-year averaged simulated fields for the LH were regridded to a 1° uniform grid to construct the prior ensemble (including SAT, SST, SSS, and $\delta^{18}O_{sw}$ fields at both monthly and annual resolutions). Second, 80% of proxy records from the LH period were randomly selected for assimilation, while the remaining 20% were withheld for internal statistical validation. $\bar{X}_{analysis}$ was then determined using Equation (1). Third, this process was repeated 500 times to generate a 500-member ensemble of analysis fields. This Monte Carlo procedure facilitated the quantification of uncertainty associated with proxy selection. A similar approach was applied to the MH period, generating 500-member $\bar{X}_{analysis}$ ensembles under different vegetation scenarios. Subsequently, 500-member MH global annual and seasonal temperature anomaly fields relative to the LH were computed for each vegetation scenarios. The 50th percentile of these ensembles was then selected as the final analysis field.

Although both the marine proxies and the PDA method employed in this study are primarily based on previous research [13,16,26], the research hypothesis and focus in this study are novel and different. Specifically, the effect of various model priors on PDA results is evaluated, identifying regions that exhibit prior independence and are primarily constrained by assimilated proxy data. More importantly, we capitalize on the distinct advantages of PDA to reconstruct temperature seasonality, which is central to addressing HTC.

2.5. Overlapping coefficient

To assess the significance of the discrepancies between PDA results under different model priors, we utilized the overlapping coefficient [27]. This coefficient quantifies the overlapping area between two probability density functions, thereby measuring the degree of agreement between probability distributions. The overlapping coefficient is defined within a range of 0 to 1, where a value of 1 indicates a perfect fit, while lower values signify a poorer fit. The analysis of 500 Monte Carlo results indicated that the assimilation temperatures for each grid generally conformed to a normal distribution. Therefore, normal distributions were fitted to the assimilation temperatures for each grid under the two priors, followed by the calculation of the overlapping coefficient between the distributions. A significance threshold of 0.05 was established to identify significant differences in PDA results across the different priors. This threshold implies that, with 95% confidence, the PDA results vary under different priors. Fig. S3 (online) illustrates the assimilated temperature distributions for various grids, along with the corresponding overlapping coefficients.

2.6. Assimilated result validation

To assess the reliability of the PDA results, both internal and external validation were conducted. Internal validation was

performed by withholding 20% of the marine proxy records across 500 Monte Carlo iterations (Section 2.4). The prior and posterior SST, SSS, and $\delta^{18}O_{sw}$ fields were used to predict these withheld proxy records by Bayesian forward models. The accuracy of the assimilation was evaluated by comparing both prior- and posterior-predicted values with the actual proxy observations. For the external validation, stalagmite $\delta^{18}O$ records from 34 caves across Eurasia, covering both the MH and LH, were selected from the SISALv2 database [28], a comprehensive speleothem isotope repository with multiple age–depth models. The simulated and assimilated $\delta^{18}O$ anomalies for the MH, relative to the LH, were compared against the observed proxy values.

3. Results and discussion

3.1. The role of model priors in PDA

In PDA, simulated climatic fields can provide background information and climatic covariance relationships between different regions, often referred to as model priors [29]. Model priors can inherently affect the PDA results [30]. To assess the effect of model priors on assimilated results and their consistency, two MH simulations with different vegetation configurations were used as PDA model priors. One employs PI vegetation conditions ("CTL_model"), while the other features enhanced vegetation cover in both Saharan and Arctic regions ("VEG_model"). This latter simulation accounts for large-scale differences in vegetation cover relative to the present. During the early to mid-Holocene, increased rainfall in the Sahara Desert region supported significantly more vegetation compared to present-day conditions [31,32]. During the MH, there was a northward expansion of forest cover in the Arctic [33]. This, in combination with the marine temperature records (Fig. S1 online), enables us to produce two MH global SAT fields, of which anomalies are with respect to assimilated results from a PI model prior. These are termed "CTL_PDA" and "VEG_PDA", respectively.

The CTL_PDA and VEG_PDA results show significant discrepancies in North Africa and North America, with assimilated differences reaching more than 2 °C (Fig. 1c). In contrast, there are no significant differences across the Eurasian and Arctic regions overall, except in certain parts of East Asia. This indicates that the assimilated temperatures in the former regions are dependent on the model priors. By contrast, the assimilated temperatures in Eurasian and Arctic regions appear to be largely independent of the model priors.

The varying vegetation configurations in CTL_model and VEG_model result in different terrestrial responses in the Sahara and the high latitudes of Eurasian, as well as in regions with close dynamic interconnections, such as North America and the Arctic

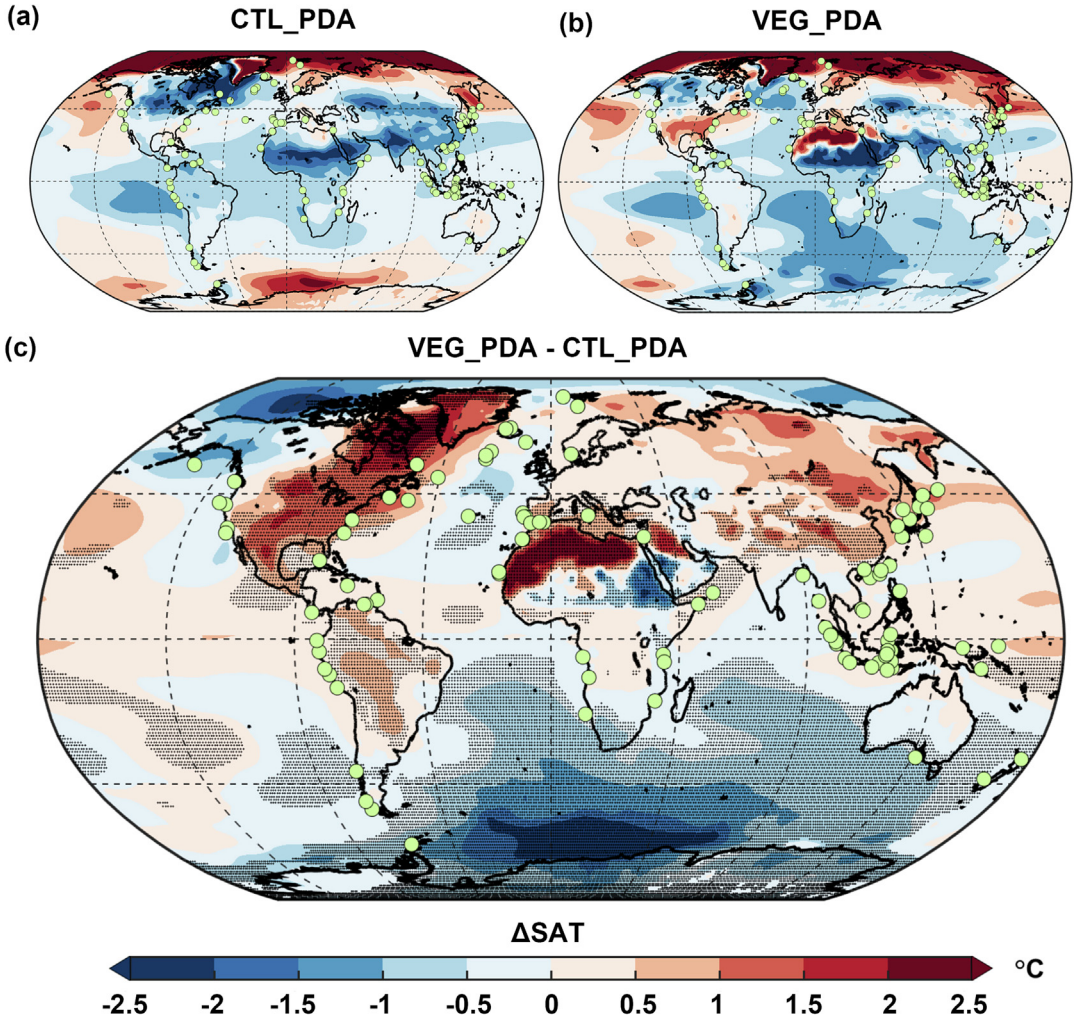

Fig. 1. Assimilated global annual mean temperature anomaly fields between MH and LH derived from PDA using (a) CTL model prior, (b) VEG model prior, and (c) the difference between the two assimilated results (VEG minus CTL). Green circles represent proxy sites with available proxy values for both the MH and LH. Black dotted areas represent the areas where the overlapping coefficient of assimilated temperature distributions from different model priors is less than 0.05, signifying statistically significant differences in the assimilated results.

[34–36]. In the adjacent marine areas, the amplitude of simulated temperature anomalies exhibits spatial coherence between the CTL_model and the VEG_model, with notably larger anomalies observed in the Arctic compared to the coastal regions of North Africa and North America (Fig. S2 online). Constrained by marine proxy records, the updates by the PDA on simulated SAT in the high latitudes of Eurasian were substantially more pronounced than those in the Sahara and North America (Fig. S4 online). Consequently, the results from CTL_PDA and VEG_PDA in Eurasia are more robust (Fig. 1). These results underscore the limitations of the PDA when it relies only on marine records, thereby emphasizing the importance of integrating terrestrial proxy records to enhance the accuracy of assimilated results.

The Southern Ocean is also characterized by significant inconsistency between the CTL_PDA and VEG_PDA (Fig. 1c). The proxy records in this region are remarkably sparse (Fig. S1 online). In the absence of a substantial number of local proxies, the assimilated fields are more subject to proxies from distant regions and the climate covariance relationships (CCR) [37,38] associated with model priors. Variations in model priors may lead to differing teleconnection patterns between the Southern Ocean and areas abundant in proxies, such as the East Pacific, Western Pacific, and Western North Pacific. This variability contributes to uncertainties in CCR and, consequently, the dependence of the assimilated Southern Ocean temperatures on the chosen model priors. For example, while the CTL_model and VEG_model exhibit different temperature patterns in the NH, this distinction is not observed in the Southern Ocean (Fig. S2 online). This results in discrepancies in CCR among different model priors, ultimately leading to significant divergence in the assimilated results for the Southern Ocean. In contrast, the North Pacific does not exhibit substantial discrepancies between PDA results, even with a limited number of local records. The primary climate modes that govern spatial temperature variability in the North Pacific include the Pacific Decadal Oscillation and the El Niño-Southern Oscillation, with variance centers located in the eastern and western Equatorial Pacific as well as the northwestern North Pacific. In these regions, a substantial number of proxies are available to constrain the basin-scale temperature changes. This underscores the efficacy of PDA, which leverages climate dynamics to produce a coherent large-scale temperature anomaly field.

To address potential biases arising from specific physical structures within a singular climate model, we further conducted a PDA utilizing model outputs from MPI-ESM (MPIESM_PDA). This analysis is compared with the assimilated results derived from iCESM (iCESM_PDA). Note that model priors in iCESM_PDA incorporate both the CTL_model and VEG_model, as this combination most comprehensively represents the physical structures of iCESM. Both models exhibit consistent assimilated temperature patterns across the Eurasian continent, thereby supporting the assertion that the assimilated Eurasian temperatures are independent of model priors (Fig. S15 online).

Via our multi-prior approach, significant uncertainties in the Southern Ocean, North Africa, and North America prevent us from a further analysis of seasonal Holocene temperature changes at a global scale. Instead, in the following, we focus on Eurasia, where the differences in assimilation results across most regions are relatively insignificant. This enables us to minimize or eliminate the influence and uncertainty of varying assimilation results on our analysis of investigating potential deficiencies in model simulations and proxy reconstruction.

3.2. Simulated cooling bias in winter temperature

Given that most contemporary MH simulations are based on PI vegetation configurations [11], the CTL scenario is prioritized to ensure comparability with existing studies. The model simulation (CTL_model) reveals a significant winter cooling anomaly (Fig. 2f–h) over Europe and Eurasian high latitudes when contrasted with the assimilated results (CTL_PDA). In contrast, the discrepancy in anomalies within these regions is minimal during the summer (Fig. 2a–c).

Validation of the PDA results was conducted using randomly withheld marine records during the PDA process, along with independent stalagmite $\delta^{18}O$ records from Eurasia. This validation confirmed an improvement in PDA performance compared to the model prior, with an approximate 15% reduction in error (Figs. S5–S7 online). Additionally, an independent seasonal terrestrial temperature proxy dataset [23] was utilized to compare PDA results with proxy-based reconstructions and model outputs. This dataset comprises 88 winter and 170 summer temperature records (Section 2). However, it is characterized by an uneven spatial distribution, with the majority of records concentrated in Europe, accounting for over 80% of the Eurasian proxies (Fig. S10 online). Consequently, this concentration may limit the dataset's ability to accurately represent continental-scale changes. For the comparative analysis, Europe was chosen as the focal region owing to its higher density of proxy records. Additionally, a spatial gridding technique ($5^\circ \times 5^\circ$) was applied to mitigate the disproportionate weighting effects caused by the dense concentration of records in specific areas (Fig. S11 online).

A robust agreement regarding MH summer warming is observed in the assimilated, simulated, and reconstructed temperature results (Fig. 2d). The results for simulated and assimilated winter temperatures exhibit contrasting values, with the latter aligning more closely with proxy data (Fig. 2i). This discrepancy suggests that the model outputs exhibit a cooling bias when simulating MH winter temperatures in Europe and the high latitudes of Eurasia. Furthermore, the majority of climate models participating in the Paleoclimate Modelling Intercomparison Project Phase 4 (PMIP4) simulate colder winters in the high latitudes of Eurasia during the MH compared to the PI period (Fig. S12 online), indicating a systematic winter cooling bias in MH simulations. In contrast, the VEG_model, which incorporates MH vegetation settings, captures the winter warming in the high latitudes of Eurasia (Fig. S8 online). This result suggests that simulated winter temperatures are affected by changes in vegetation. Additionally, the warming associated with vegetation changes is closely linked to Arctic sea ice concentration, particularly in high-latitude regions [12]. Therefore, the cooling bias is attributed to deficiencies in the models' representation of vegetation and sea ice feedback mechanisms. Climate models may underestimate the retreat of summer Arctic sea ice during the MH due to inadequate consideration of vegetation changes and sea ice dynamics, including the physics of sea-ice melt ponds [10,12,39,40]. This underestimation can result in sustained warming effects on winter temperatures, driven by a reduction in surface albedo and enhanced ocean–atmosphere interactions [10]. The warming induced by sea ice processes may, in turn, exert additional effects on vegetation, thereby establishing a positive feedback loop. However, in certain mid-latitude regions of Europe, the VEG_model still struggles with simulating MH winter warming, as indicated by assimilated and reconstructed data (Figs. S8, S9 online). This observation indicates a localized simulation bias, likely associated with the omission of anthropogenic land use in MH simulations. Since the early Holocene, the expansion of agriculture has significantly altered terrestrial landscapes and ecosystems [41,42]. This transformation may have particularly affected mid-latitude Europe, where intensive human land use was already well established at that time [43].

In a seminal paper concerning HTC [4], a key debate revolves around why the reconstructed annual mean temperature is characterized by an evident contrast in simulated annual mean

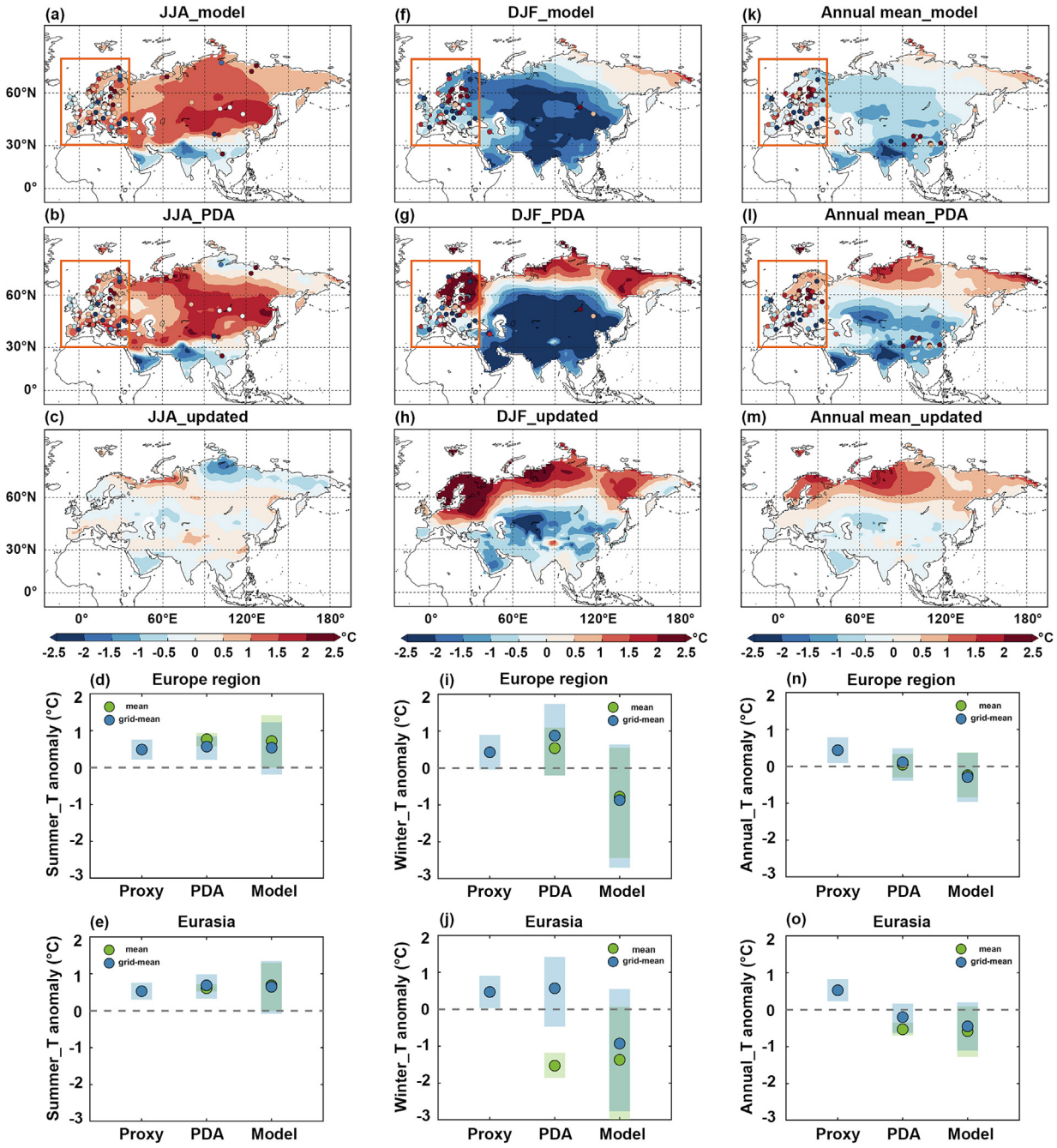

Fig. 2. PDA-proxy-model comparisons of Eurasian temperatures under CTL vegetation scenario. (a) Background field represents simulated MH Eurasian summer (JJA) temperature anomaly fields under CTL scenario, circles represent proxy-based reconstructed summer temperatures, and the box represents the Europe region. (b) As in (a), but for assimilated fields. (c) Updated field of the assimilated field relative to the simulated field, namely (b) minus (a). (d, e) PDA-proxy-model comparisons of summer temperature for Europe region and Eurasia, respectively. Blue circles represent the mean of $5^\circ \times 5^\circ$ grids with proxy records, while green circles represent the mean of all grids in the fields. Shadows represent the uncertainty range, with 1σ for reconstructed results, and 90th-percentile for assimilated and simulated results. (f–j), (k–o) As in (a)–(e), but for winter (DJF) and annual mean temperatures, respectively.

temperature but has a good consistency with simulated summer temperature, particularly in the mid to high latitudes of the NH. This discrepancy has led to an ongoing debate between proxy seasonal bias and model deficiency. While previous research has identified a warm-season bias in Holocene marine records as a primary contributor to the inconsistencies between proxy data and model outputs [7,9], subsequent studies have raised questions about this conclusion due to methodological issues [44,45] and the uneven spatial distribution of proxy records [8]. Recent developments in NH pollen records and East Asian snail records suggest that there is no significant seasonal bias present in the proxy data, indicating that model deficiencies are likely the primary cause of the discrepancies observed between proxies and models [8,46]. The assimilated seasonal temperature data for Eurasia serve as an independent reference to objectively assess temperature seasonality in the mid to high latitudes of the NH, a region characterized by pronounced proxy-model mismatches as highlighted in HTC [4]. Notably, the assimilated results are in strong agreement with the proxy records. Therefore, deficiencies in winter temperature simulation, rather than proxy seasonal bias, are identified as the primary drivers of discrepancies between models and proxies in this critical region. These deficiencies are attributed to overlooked

changes in surface landscapes and an underestimation of Arctic summer sea ice retreat in MH simulations. Furthermore, the significant retreat of Arctic sea ice may have considerable warming effects on lower latitudes [47]. Therefore, the simulated cooling bias may extend beyond high latitudes in the NH, potentially exerting global impacts and contributing to a cooling bias in GMST.

3.3. Regional versus continental seasonal temperature reconstruction

Holocene temperature evolution exhibits significant spatial heterogeneity [8,48–51]. This variability, coupled with the uneven spatial distribution of proxies [52], prompts an inquiry into whether the restricted spatial distribution of these proxies can adequately represent temperature changes on a larger scale, such as continental or hemispheric levels.

The assimilated Eurasian temperatures are independent of model priors, enabling a comprehensive evaluation of spatial proxy sparsity and its effect on regional temperature variations. In Eurasia, the assimilated MH temperature seasonality exhibits significant and consistent spatial heterogeneity across different vegetation settings (Fig. 3). This pattern is comparable to the proxy data but contrasts with the results from simulations. Notably, the

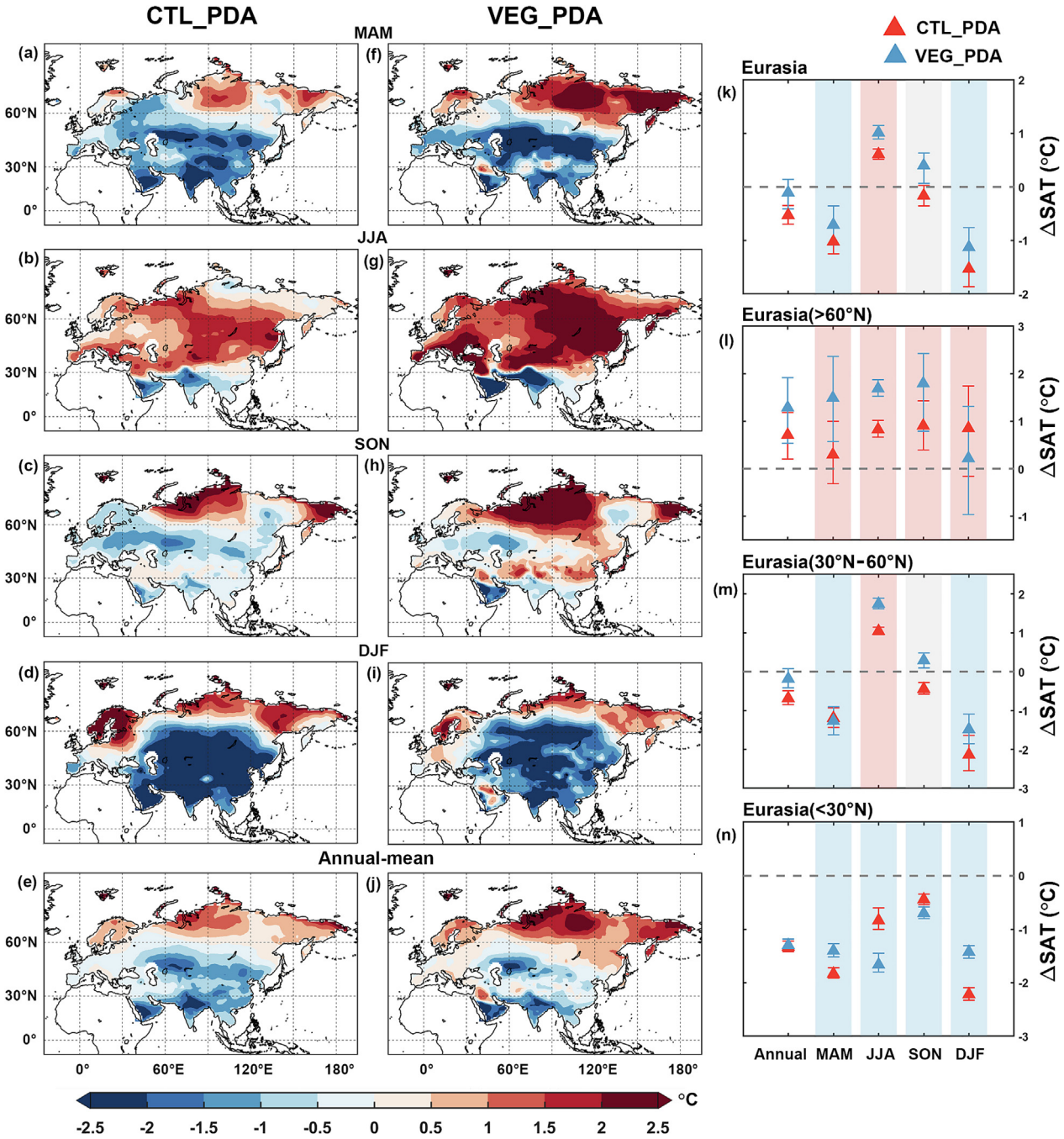

Fig. 3. Assimilated MH Eurasian seasonal and annual mean temperatures anomaly fields under different vegetation scenarios and comparison between different latitudinal zones. (a) Assimilated MH Eurasian spring (MAM) mean temperature anomaly field under CTL scenario. (b–e) As in (a), but for JJA, SON, DJF and annual mean temperature, respectively. (f, j) As in (a–e), but for the VEG scenario. (k) Assimilated mid-Holocene Eurasian annual mean and seasonal SAT anomaly, red triangles for CTL scenario, blue triangles for VEG scenario, error bars represent 90th-percentile ranges, red (blue) shade represents mid-Holocene warming (cooling) under both scenarios. (l–n) As in (k), but for different Eurasian regions, high latitudes (> 60°N), mid-latitudes (30°N–60°N), and low latitudes (< 30°N), respectively.

assimilated winter and annual mean temperatures indicate substantial warming in Europe and the high-latitude regions of Eurasia, while exhibiting cooling trends in other Eurasian areas. This stands in contrast to the homogeneous cooling observed in the simulated temperature data (Fig. 2f). The spatial heterogeneity of MH seasonal temperature changes is also reflected in the MPIESM_PDA results (Fig. S16 online). The Eurasian temperature proxies are predominantly located in Europe, which accounts for 83%, 95%, and 83% of the total continental records for summer, winter, and annual mean temperature proxies, respectively (Fig. S10 online). This uneven distribution results in regional biases in continental-scale reconstructions, particularly affecting the reconstructions of winter and annual mean temperatures (Fig. 2j, o).

In the CTL_PDA, the observed change in European MH winter temperature anomalies, characterized by a warming of 0.5 °C, contrasts sharply with the continental average, which indicates a cooling of approximately 1.5 °C (Fig. 2i, j). This discrepancy ultimately leads to a distinct spatial pattern of temperature change in annual mean temperature between Europe and Eurasia, with Europe experiencing slight warming while Eurasia undergoes significant cooling (Fig. 2n, o), despite both regions sharing a consistent pattern of summer warming. Notably, the spatial gridding technique employed does not effectively address the regional bias that arises from the uneven distribution of proxy records (Fig. 2j).

A continental-scale mechanism is suggested to drive the observed consistent summer warming, while different regional processes are required to explain the spatial heterogeneity of winter temperature changes. The summer warming can be attributed to the strong MH boreal summer insolation (Fig. S13 online). This leads to a general warming trend across Eurasia, particularly north of 30°N, with the exception of the Indian and Arabian continents, which experienced cooling due to negative cloud radiative forcing linked to the intensified boreal summer monsoon [11] (Fig. 3). During the MH winter, Europe and the high latitudes of Eurasia exhibit a warming anomaly, while the remainder of Eurasia experiences cooling. The warming observed in high latitudes is attributed to the legacy effects of insolation-driven reductions in Arctic summer sea ice [53]. The cooling observed in other regions is a direct consequence of the low MH local winter insolation, which is affected by low soil heat capacity [54] (Fig. S13 online). Furthermore, this, in addition to assimilated continental temperatures from autumn to spring (Fig. 3), suggests that the reduction of summer sea ice in the Arctic during the MH has limited effects on SAT south of approximately 60°N in Eurasia. In these areas, local seasonal and mean annual temperature changes are predominantly governed by insolation-induced seasonality and the associated climate feedback mechanisms (Fig. S14 online).

Our results suggest that the current spatial distribution of proxies, particularly the predominance of European proxies within the Eurasian dataset, is insufficient for providing a comprehensive continental-scale representation of MH temperature changes. This limitation has been overlooked in previous model-proxy comparisons [3,4,8,9,23], which have relied on overestimated homogeneous temperature changes as projected by climate models (Fig. 2f, g, j, k, l, o). Consequently, this gives rise to a warming bias in the reconstructions of continental-scale winter and annual mean temperatures, thereby complicating the understanding of Holocene temperature evolution. Furthermore, this implies that regional biases may persist in global temperature reconstructions due to the uneven distribution of records, even when employing spatial gridding techniques. The spatial sparsity of proxies also introduces uncertainty in the assimilated seasonal temperatures, particularly in under-sampled regions where PDA-proxy-model comparisons are lacking. Proxy records from these under-sampled areas, especially the mid-latitudes of Eurasia (30°N–60°N), such as the Silk Road regions, are highly valuable. They not only have the potential to fill gaps in the Eurasian database but also to qualitatively represent continental-scale changes.

4. Conclusion

This study applied MH temperature assimilation using global marine records and multiple climate model priors. The results demonstrate that different model priors can substantially affect Holocene temperature assimilation results. Due to the challenges associated with achieving globally consistent PDA results across different model priors, a comprehensive resolution of the HTC at a global scale remains unattainable. However, in the mid-to-high latitudes of the NH, where the discrepancy between model simulations and proxy reconstructions is most pronounced, our results underpin the previously outlined shortcomings of proxy reconstructions and model simulations. Addressing the HTC effectively requires improvements in the representation of surface landscapes, vegetation dynamics, and sea ice feedbacks in climate models, alongside an expansion of proxy data coverage. Despite these advancements, certain limitations persist in this study and warrant further investigation. First, the analysis relies on two independent isotope-enabled climate models, which may not fully capture the range of climate covariance structures present in contemporary models. Second, the absence of a reliable proxy system model for pollen records (i.e., the most abundant terrestrial proxy for the Holocene) precludes their direct assimilation based on raw proxy values. We chose not to assimilate pollen records based on their reconstructed temperature values, as done by Erb et al. [14], because this would introduce an additional layer of uncertainty. Nevertheless, the assimilated seasonal temperatures exhibit strong agreement with independent terrestrial temperature records, providing validation for the PDA approach and our findings. Furthermore, the limited availability of winter temperature proxy records in Asia prevents direct validation of the assimilated winter temperature anomalies, underscoring the critical need for expanded winter temperature reconstructions in the region. Future research should prioritize the development of a multi-model prior framework for PDA [55], the development of proxy system models for terrestrial proxies, and the inclusion of additional proxy records from under-sampled regions. These efforts will be essential for generating a more comprehensive and globally representative seasonal temperature reanalysis.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Basic Science Center for Tibetan Plateau Earth System (BSCTPES, 41988101), the National Key Research and Development Program of China (2023YFF0805200), and the National Natural Science Foundation of China (42101148).

Author contributions

Fahu Chen and Xu Zhang conceived the idea. Shuo Hao conducted data assimilation and prepared the figures. Xu Zhang, Shuo Hao, and Yanwu Duan performed data analysis. Xu Zhang and Shuo Hao led the manuscript writing. All authors contributed to the interpretation of results and the final version of the manuscript.

Data availability

The sea surface temperature proxy dataset used for data assimilation is publicly accessible via the National Oceanic and Atmospheric Administration (NOAA) Paleoclimatology Data Archive at https://www.ncdc.noaa.gov/paleo/study/33112. The independent seasonal temperature proxy dataset used for validation is available from Zhang et al. [23] (https://doi.org/10.1016/j.scib.2021.09.004). Source data supporting this study are provided with the paper. The MATLAB code for paleoclimate data assimilation and analysis is publicly available at https://github.com/ShuoHao1996/DA_6ka. The DASH MATLAB package used for paleoclimatic data assimilation is accessible at https://github.com/JonKing93/DASH. The MATLAB code of Bayesian forward models used in this study can be found at https://github.com/jesstierney.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scib.2025.03.039.

References

[1] Fischer H, Meissner KJ, Mix AC, et al. Palaeoclimate constraints on the impact of 2℃ anthropogenic warming and beyond. Nat Geosci 2018;11:474–85.
[2] Marcott SA, Shakun JD, Clark PU, et al. A reconstruction of regional and global temperature for the past 11,300 years. Science 2013;339:1198–201.
[3] Kaufman D, McKay N, Routson C, et al. Holocene global mean surface temperature, a multi-method reconstruction approach. Sci Data 2020;7:201.
[4] Liu ZY, Zhu J, Rosenthal Y, et al. The Holocene temperature conundrum. Proc Natl Acad Sci USA 2014;111:E3501–5.
[5] Kaufman DS, Broadman E. Revisiting the Holocene global temperature conundrum. Nature 2023;614:425–35.
[6] Chen FH, Duan YW, Hao S, et al. Holocene thermal maximum mode versus the continuous warming mode: problems of data-model comparisons and future research prospects. Sci China Earth Sci 2023;66:1683–701.
[7] Bova S, Rosenthal Y, Liu ZY, et al. Seasonal origin of the thermal maxima at the Holocene and the last interglacial. Nature 2021;589:548–53.
[8] Zhang WC, Wu HB, Cheng J, et al. Holocene seasonal temperature evolution and spatial variability over the Northern Hemisphere landmass. Nat Commun 2022;13:5334.
[9] Marsicek J, Shuman BN, Bartlein PJ, et al. Reconciling divergent trends and millennial variations in Holocene temperatures. Nature 2018;554:92–6.
[10] Park HS, Kim SJ, Stewart AL, et al. Mid-Holocene Northern Hemisphere warming driven by Arctic amplification. Sci Adv 2019;5:eaax8203.
[11] Brierley CM, Zhao AN, Harrison SP, et al. Large-scale features and evaluation of the PMIP4-CMIP6 midHolocene simulations. Clim Past 2020;16:1847–72.
[12] Thompson AJ, Zhu J, Poulsen CJ, et al. Northern Hemisphere vegetation change drives a Holocene thermal maximum. Sci Adv 2022;8:eabj6535.
[13] Osman MB, Tierney JE, Zhu J, et al. Globally resolved surface temperatures since the Last Glacial Maximum. Nature 2021;599:239–44.
[14] Erb MP, McKay NP, Steiger N, et al. Reconstructing Holocene temperatures in time and space using paleoclimate data assimilation. Clim Past 2022;18:2599–629.
[15] Li X, Liu F, Fang M. Harmonizing models and observations: data assimilation in Earth system science. Sci China Earth Sci 2020;63:1059–68.
[16] Tierney JE, Zhu J, King J, et al. Glacial cooling and climate sensitivity revisited. Nature 2020;584:569–73.
[17] Hakim GJ, Emile-Geay J, Steig EJ, et al. The last millennium climate reanalysis project: framework and first results. J Geophys Res Atmos 2016;121:6745–64.
[18] Marcott SA, Shakun JD. Global temperature changes mapped across the past 24,000 years. Nature 2021;599:208–9.
[19] Tierney JE, Tingley MP. A Bayesian, spatially-varying calibration model for the TEX_{86} proxy. Geochim Cosmochim Acta 2014;127:83–106.
[20] Tierney JE, Tingley MP. BAYSPLINE: a new calibration for the alkenone paleothermometer. Paleoceanogr Paleoclimatol 2018;33:281–301.
[21] Malevich SB, Vetter L, Tierney JE. Global core top calibration of $\delta^{18}O$ in planktic foraminifera to sea surface temperature. Paleoceanogr Paleoclimatol 2019;34:1292–315.
[22] Tierney JE, Malevich SB, Gray W, et al. Bayesian calibration of the Mg/Ca paleothermometer in planktic foraminifera. Paleoceanogr Paleoclimatol 2019;34:2005–30.
[23] Zhang WC, Wu HB, Geng JY, et al. Model-data divergence in global seasonal temperature response to astronomical insolation during the Holocene. Sci Bull 2022;67:25–8.
[24] Cauquoin A, Werner M, Lohmann G. Water isotopes - climate relationships for the mid-Holocene and preindustrial period simulated with an isotope-enabled version of MPI-ESM. Clim Past 2019;15:1913–37.
[25] Tardif R, Hakim GJ, Perkins WA, et al. Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. Clim Past 2019;15:1251–73.
[26] King J, Tierney J, Osman M, et al. DASH: a MATLAB toolbox for paleoclimate data assimilation. Geosci Model Dev 2023;16:5653–83.
[27] Inman HF, Bradley EL. The overlapping coefficient as a measure of agreement between probability-distributions and point estimation of the overlap of 2 normal densities. Commun Stat-Theory Methods 1989;18:3851–74.
[28] Comas-Bru L, Rehfeld K, Roesch C, et al. SISALv2: a comprehensive speleothem isotope database with multiple age-depth models. Earth Syst Sci Data 2020;12:2579–606.
[29] Fang M, Li X. Paleoclimate data assimilation: its motivation, progress and prospects. Sci China Earth Sci 2016;59:1817–26.
[30] Amrhein DE, Hakim GJ, Parsons LA. Quantifying structural uncertainty in paleoclimate data assimilation with an application to the Last Millennium. Geophys Res Lett 2020;47:e2020GL090485.
[31] Hoelzmann P, Jolly D, Harrison SP, et al. Mid-Holocene land-surface conditions in northern Africa and the Arabian Peninsula: a data set for the analysis of biogeophysical feedbacks in the climate system. Glob Biogeochem Cycle 1998;12:35–51.
[32] Tierney JE, Pausata FSR, deMenocal PB. Rainfall regimes of the Green Sahara. Sci Adv 2017;3:e1601503.
[33] Bigelow NH, Brubaker LB, Edwards ME, et al. Climate change and Arctic ecosystems: 1. vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present. J Geophys Res 2003;108:8170.
[34] Muschitiello F, Zhang Q, Sundqvist HS, et al. Arctic climate response to the termination of the African Humid Period. Quat Sci Rev 2015;125:91–7.
[35] Pausata FSR, Zhang Q, Muschitiello F, et al. Greening of the Sahara suppressed ENSO activity during the mid-Holocene. Nat Commun 2017;8:16020.
[36] Pausata FSR, Gaetani M, Messori G, et al. The greening of the Sahara: past changes and future implications. One Earth 2020;2:235–50.
[37] Houtekamer PL, Mitchell HL. Data assimilation using an ensemble Kalman filter technique. Mon Weather Rev 1998;126:796–811.
[38] Hamill TM, Whitaker JS, Snyder C. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon Weather Rev 2001;129:2776–90.
[39] Guarino MV, Sime LC, Schröeder D, et al. Sea-ice-free Arctic during the Last Interglacial supports fast future loss. Nat Clim Chang 2020;10:928–32.
[40] Chen J, Zhang Q, Lu Z, et al. Reconciling East Asia's mid-Holocene temperature discrepancy through vegetation-climate feedback. Sci Bull 2024;69:2420–9.
[41] Ellis EC, Gauthier N, Goldewijk KK, et al. People have shaped most of terrestrial nature for at least 12,000 years. Proc Natl Acad Sci USA 2021;118:e2023483118.
[42] Ruddiman WF. The early anthropogenic hypothesis: challenges and responses. Rev Geophys 2007;45:RG4001.
[43] Goldewijk KK, Beusen A, Doelman J, et al. Anthropogenic land use estimates for the Holocene - HYDE 3.2. Earth Syst Sci Data 2017;9:927–53.
[44] Zhang X, Chen FH. Non-trivial role of internal climate feedback on interglacial temperature evolution. Nature 2021;600:E1–6.
[45] Laepple T, Shakun J, He F, et al. Concerns of assuming linearity in the reconstruction of thermal maxima. Nature 2022;607:E12–4.
[46] Dong YJ, Wu NQ, Li FJ, et al. The Holocene temperature conundrum answered by mollusk records from East Asia. Nat Commun 2022;13:5153.
[47] England MR, Polvani LM, Sun LT, et al. Tropical climate responses to projected Arctic and Antarctic sea-ice loss. Nat Geosci 2020;13:275–81.
[48] Affolter S, Hauselmann A, Fleitmann D, et al. Central Europe temperature constrained by speleothem fluid inclusion water isotopes over the past 14,000 years. Sci Adv 2019;5:eaav3809.
[49] Bader J, Jungclaus J, Krivova N, et al. Global temperature modes shed light on the Holocene temperature conundrum. Nat Commun 2020;11:4726.
[50] Cartapanis O, Jonkers L, Moffa-Sanchez P, et al. Complex spatio-temporal structure of the Holocene thermal maximum. Nat Commun 2022;13:5662.
[51] Jiang J, Meng B, Wang H, et al. Spatial patterns of Holocene temperature changes over mid-latitude Eurasia. Nat Commun 2024;15:1507.
[52] Kaufman D, McKay N, Routson C, et al. A global database of Holocene paleotemperature records. Sci Data 2020;7:115.
[53] Wu ZP, Yin QZ, Guo ZT, et al. Hemisphere differences in response of sea surface temperature and sea ice to precession and obliquity. Glob Planet Change 2020;192:103223.
[54] Lambert FH, Webb MJ, Joshi MM. The relationship between land-ocean surface temperature contrast and radiative forcing. J Clim 2011;24:3239–56.
[55] Parsons LA, Amrhein DE, Sanchez SC, et al. Do multi-model ensembles improve reconstruction skill in paleoclimate data assimilation? Earth Space Sci 2021;8:e2020EA001467.