Global biome changes over the last 21,000 years inferred from model-data comparisons

Chenzhi Li[1,2], Anne Dallmeyer[3], Jian Ni[4], Manuel Chevalier[5], Matteo Willeit[6], Andrei A. Andreev[1], Xianyong Cao[7], Laura Schild[1,2], Birgit Heim[1], Ulrike Herzschuh[1,2,8]

5   [1] Polar Terrestrial Environmental Systems, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Telegrafenberg A45, 14473 Potsdam, Germany

[2] Institute of Environmental Science and Geography, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

[3] Max Planck Institute for Meteorology, Bundesstraße 53, 20146 Hamburg, Germany

10   [4] College of Life Sciences, Zhejiang Normal University, 321004 Jinhua, China

[5] Institute of Geosciences, Sect. Meteorology, Rheinische Friedrich-Wilhelms-Universität Bonn, Auf dem Hügel 20, 53121 Bonn, Germany

[6] Department of Earth System Analysis, Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany

15   [7] Alpine Paleoecology and Human Adaptation Group (ALPHA), State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources (TPESER), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, 100101 Beijing, China

[8] Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

20   *Correspondence to:* Ulrike Herzschuh (Ulrike.Herzschuh@awi.de)

Abstract. We present a global megabiome reconstruction for 43 timeslices at 500-year intervals throughout the last 21,000 years based on an updated and thus currently most extensive global taxonomically and temporally standardized fossil pollen dataset of 3,691 records. The evaluation with modern potential natural vegetation

25   distributions yields an agreement of ~80%, suggesting a high degree of reliability of the pollen-based megabiome reconstruction. With its high temporal and spatial resolution, this reconstruction is ideally suited for the evaluation of paleo-simulations from Earth System Models (ESMs). As an example, we compare the reconstruction with an ensemble of six different biomized simulations based on transient vegetation simulations performed by ESMs.

The global spatiotemporal patterns of megabiomes estimated by the simulation ensemble and reconstructions are

30   generally consistent, i.e., from glacial non-forest megabiomes to Holocene forest megabiomes, in line with the

general climate warming trend and continental ice-sheet retreat. The shift to a global spatial megabiome distribution similar to today's took place during the early Holocene.

At a global scale over the last 21,000 years, the deviations between the reconstruction and the simulation ensemble are (a) largest during the Last Glacial Maximum and early deglaciation periods, mainly due to different estimates

35   of tundra in the circum-Arctic areas and the Tibetan Plateau; and (b) moderate during the Holocene, mainly due to different estimates of non-forest megabiomes in relatively semi-arid zones such as North Africa and the Mediterranean that increases over time. To some extent, these mismatches could be attributed to systematic model biases in the simulated climate, as well as to the different plant representations and low taxonomic resolution of pollen in the reconstructions.

40

1 Introduction

Earth system models (ESMs) that incorporate vegetation dynamics are useful tools for understanding historical simulations and future projections of changes in the composition, structure, and distribution of vegetation ecosystems, as well as their responses and feedbacks to climate change (Song et al., 2021; Brierley et al., 2020).

45   However, to assess model biases and further improve these models for obtaining more reliable and reduced uncertainty in future projections, global and long-lasting paleo-vegetation reconstructions are needed for the evaluation of the vegetation response to climate change (Cao et al., 2019; Dallmeyer et al., 2022). Pollen records, as the most widespread terrestrial paleoecological archives, and their conversion into paleo-vegetation are most suitable for this purpose (Prentice et al., 1996). To date, however, the synthesis of global-scale pollen-based

50   vegetation reconstructions has been limited to selected timeslices (i.e., mid-Holocene and Last Glacial Maximum (LGM); Harrison, 2017; Hoogakker et al., 2016), while at long temporal scales has it been limited to specific regions (such as northern and eastern Asia, extratropical Northern Hemisphere; Tian et al., 2018; Cao et al., 2019). A global view of reconstructed vegetation dynamics and distributions since the LGM with high temporal resolution is still missing.

55   In a recent effort, we synthesized LegacyPollen 2.0 (Li et al., 2024), a taxonomically and temporally standardized global Late Quaternary fossil pollen dataset of 3,691 records, that covers the main global ecoregions (Herzschuh et al., 2022). In this study, we biomize the LegacyPollen 2.0 dataset for 43 timeslices at 500-year intervals throughout the last 21,000 years with a biomization method (Prentice et al., 1996; Prentice and Webb, 1998) that incorporates updated and harmonized pollen taxa-plant functional types-megabiome assignment schemes. For a

60   direct comparison with ESM-simulated vegetation, we assign the reconstructions into the same megabiomes used in the biomization tool for ESM output by Dallmeyer et al. (2019). This paper aims (a) to present megabiome dynamics at the global scale since the Last Glacial Maximum, (b) to compare the reconstruction with megabiome simulations from an ensemble of six different transient ESM simulations using the Earth mover's distance (EMD; Chevalier et al., 2023b) while taking into account the uncertainties of the biomized data and case-specific weighted

65   distances, and (c) to identify regions and periods with strong data-model mismatches to provide clues for improving systematic model biases.

2. Data and methods

2.1 Pollen dataset

70 We expanded the LegacyPollen 1.0 dataset (Herzschuh et al., 2022), a taxonomically and temporally standardized global late Quaternary fossil pollen dataset, to LegacyPollen 2.0. The updated dataset comprises 3,691 palynological records (Appendix Fig. A1), approximately 900 more than the previous dataset. Of these new records, 654 were derived from the Neotoma Paleoecology Database (Neotoma hereafter; https://www.neotomadb.org/, last access: August 31, 2022; Williams et al., 2018) and its constituent databases

75 (e.g., African Pollen Database (APD; Lézine et al., 2021), European Pollen Database and Alpine Pollen Database (EPD and ALPADABA; Fyfe et al., 2009; Giesecke et al., 2014), and Latin American Pollen Database (LAPD; Flantua et al., 2015)). Also, 61 records from the ACER 1.0 database (https://doi.org/10.1594/PANGAEA.870867, last access: September 22, 2022; Sánchez Goñi et al., 2017), 178 records from the Chinese fossil pollen dataset (Zhou et al., 2023; Cao et al., 2022), and 8 of our own new records (AWI, Alfred Wegener Institute) were included.

80 A total of 1124 records originate from North America, 1448 from Europe, 690 from Asia, 187 from South America, 160 from Africa, and 82 from the Indo-Pacific. While there are geographical gaps in pollen record coverage, particularly in the Southern Hemisphere, the dataset LegacyPollen 2.0 covers the world's main vegetation and climate zones.

Taxonomic harmonization (i.e., woody taxa and major herbaceous taxa have been harmonized to genus level and
85 other herbaceous taxa to family level) and temporal standardization (i.e., re-estimation of age-depth models) follow the previously established frameworks LegacyPollen 1.0 (Herzschuh et al., 2022) and LegacyAge 1.0 (Li et al., 2022), respectively. In compiling the dataset, we also followed the practices recommended by Flantua et al. (2023) for large-scale paleoecological data synthesis, such as how to select data sources and filter the dataset.

Compared to the LegacyPollen 1.0 dataset, we now include the Neotoma DOI (if Neotoma source) in the overview
90 table of site metadata to eliminate the broken chain of static LegacyPollen 2.0 dataset with living (such as updating discovered metadata errors and chronologies) Neotoma and associated risk of data staleness. The Neotoma DOIs were generated with the *doi* function from the *neotoma2* package in R (version 1.0.3; https://github.com/asardaes/dtwclust, last access: June 10, 2024; Socorro and Goring, 2024). Furthermore, we also added the PANGAEA Event (PANGAEA dataset identifier) for each new record to ensure that our dataset meets
95 PANGAEA's high standards for quality, usability, and compliance. The LegacyPollen 2.0 dataset is archived in the PANGAEA repository (https://doi.pangaea.de/10.1594/PANGAEA.965907; Li et al., 2024) and is open-access.

2.2 Pollen-based megabiome reconstruction

100 We converted pollen data from LegacyPollen 2.0 into megabiomes for 56,053 samples in 43 timeslices at 500-year intervals throughout the last 21,000 years using the biomization method of Prentice et al. (1996). The assignment of pollen taxa to plant functional types (PFTs), the first step required by the biomization procedure, referenced previous biomization schemes on each continent, with some updates and harmonizations based on a globally applicable standardized classification of PFTs (Harrison et al. 2010; Harrison, 2017). The PFTs were
105 then assigned to megabiome categories (Dallmeyer et al., 2019), namely tropical forest (TRFO), warm-temperate

(subtropical) forest (WTFO), temperate forest (TEFO), boreal forest (BOFO), (warm) savanna and dry woodland (SAVA), grassland and dry shrubland (STEP), (warm) desert (DESE), tundra and polar desert (TUND). These categories are also applied to biomize Earth System Model results, allowing for direct data-model comparisons (Dallmeyer et al., 2019).

110 We assigned the 1447 harmonized pollen taxa from the 3,691 records to 98 PFTs and then to 8 megabiomes (Table 1; Supplement). The pollen abundances of *Larix* and *Pinus* were multiplied by factors of 15 and 0.5 (following Bigelow et al., 2003 and Cao et al., 2019), respectively, to compensate to some extent for pollen productivity-related representativeness issues, prior to calculating affinity scores in the applied biomization routine. When the affinity scores for each megabiome were calculated (cf. Prentice et al., 1996) for every pollen sample, pollen taxa

115 with less than 0.5% abundance were excluded to reduce noise resulting from occasional pollen grains derived from long-distance transport or contamination (Prentice et al., 1996; Chen et al., 2010). Finally, the megabiome with the highest affinity score was allocated to each pollen sample, subject to a criterion that the least PFT-rich megabiome takes precedence when the affinity values for two or more megabiomes are identical (following Chen et al., 2010). The biomization affinity scores were calculated using a biomization algorithm implemented in R

120 (Cao and Tian, 2021; R version 4.2.3, https://www.r-project.org/, last access: May 10, 2023; R Core Team, 2020). Furthermore, the assignment of pollen taxa to megabiomes and biomization routines were performed independently for each continent. The pollen sample at each target timeslice was selected from the time-nearest sample within a time window (time window= target timeslice ± 250 years).

125 **Table 1 Overview of the number of pollen records, pollen taxa, plant functional types (PFTs), and megabiomes used in the biomization procedures, along with references to biomization schemes by continent.** The lists of taxa-PFTs and PFTs-megabiome assignments are available at MPG.PuRe repository.

Continent	Pollen records	Taxa	PFTs	Megabiomes	References
Europe	1448	243	41	7	Ni et al. (2014)
					Binney et al. (2017)
					Marinova et al. (2018)
					Cao et al. (2019)
Asia	692	424	49	8	Chen et al. (2010)
					Ni et al. (2014)
					Binney et al. (2017)
					Tian et al. (2018)
					Cao et al. (2019)
North America	1124	393	47	8	Thompson and Anderson (2000)
					Ortega-Rosas et al. (2008)
					Bigelow et al. (2003)
					Ni et al. (2014)
					Cao et al. (2019)
Africa	160	556	8	6	Vincens et al. (2006)
					Lézine et al. (2009)
Indo-Pacific	82	429	22	8	Pickett et al. (2004)
South America	187	576	19	8	Marchant et al. (2001 & 2009)
Total	3693	1447	98	8	

2.3 Transient ESM-based simulations with dynamic vegetation

130   We use six transient simulations for the last 21,000 years performed with Earth System Models with fully coupled dynamic vegetation. Among these are two simulations conducted with the Max-Planck-Institute Earth-System-Model (MPI-ESM; Mauritsen et al., 2019), further referred to as MPI-ESM_GLAC1D (Dallmeyer et al., 2022) and MPI-ESM_ICE6G (Ice6G_P2 in Kapsch et al., 2022; Mikolajewicz et al., 2023). Besides differences in the model version and tuning, these simulations differ in particular with respect to the prescribed ice-sheet history,

135   using either the GLAC-1D (Tarasov et al., 2012) or ICE-6G (Peltier et al., 2015) reconstruction. Both simulations ran at the spatial resolution T31 (~3.75°x3.75° on a Gaussian grid) for the atmosphere and land model. Orbital forcing has been prescribed from Berger (1978) and greenhouse gas (GHG) forcings from Köhler et al. (2017). Bathymetry, topography, and river routing were continuously updated in ten-year intervals throughout the deglaciation. The meltwater flux from the Laurentide ice sheet has been modified in the period of 15.2–11.8 cal.

140   ka BP (calibrated thousand years before present, where "present" is 1950 CE) in the simulation MPI-ESM-GLAC1D, mimicking the meltwater storage and release from proglacial lakes and thus more realistically simulate the Younger Dryas event (cf. Dallmeyer et al., 2022).

    In addition, the set of simulations includes the full-forcing TRACE-21k-I (cf. Liu et al., 2009) and TRACE-21k-II (cf. He and Clark, 2022) simulations performed with the Community Climate Model version 3 (CCSM3, Collins

145   et al., 2006) forced with variations in insolation (Berger, 1978), GHG concentration (Joos and Spahni, 2008), and continental ice sheets from the ICE5G reconstructions (Peltier, 2004). TRACE-21k-II was based on the protocol of prescribing the reconstructed Atlantic meridional overturning circulation (AMOC) for the Bølling-Allerød interstadial (~14.7–12.9 cal. ka BP) and the Holocene instead of the reconstructed freshwater forcing, while in TRACE-21k-I, the AMOC has been forced by the meltwater flux to the North Atlantic and the Gulf of Mexico

150   during the entire simulation. Similar to the MPI-ESM simulations, the TRACE-21k simulations ran at a spatial resolution of T31 (~3.75°x3.75° on a Gaussian grid).

    The set of simulations contains two simulations performed with the fast Earth System model CLIMBER-X (Willeit et al., 2022 & 2023; Willeit and Ganopolski, 2016) at a spatial resolution of ~5°x5°. These simulations were both performed in an identical setup (similar to Masoum et al., 2024) but with different ice-sheet and surface

155   topography forcings (GLAC-1D or ICE-6G reconstructions; Peltier et al., 2015; Tarasov et al., 2012). GHG and insolation have been prescribed from Köhler et al. (2017) and Laskar et al. (2004), respectively.

    All simulations have been aggregated to time series of 100-year climatological means. The dynamic vegetation in all models is represented by different sets of plant functional types that can coexist in the grid-cells. Land use is not included in any of these simulations. The PFTs have been converted into megabiomes using the tool by

160   Dallmeyer et al. (2019) that converts simulated PFT distributions into megabiomes based on the relative PFT cover fractions and temperature constraints. We only consider grid-cells and timeslices for which reconstructions are available. As representatives of the simulation ensemble, we choose the megabiome that occurs most frequently in the set of simulations for each site and timeslice, further referred to as the ESM-representative megabiome. When the highest frequency megabiomes were not unique, 999 random sampling of these

165   megabiomes was performed to reselect the most representative megabiome.

2.4 Evaluation with modern potential vegetation

Information on modern vegetation distribution is required to validate the performance of pollen-based reconstructions and model ESM-based simulations. However, the simulations used here only determine potential natural vegetation in a quasi-equilibrium with climate, whereas the pollen-based reconstruction of modern

170   vegetation also incorporates information on anthropogenic disturbances. Therefore, the modern potential natural vegetation distributions are used for validation, allowing us to evaluate not only the level of modern anthropogenic disturbance to natural vegetation in the pollen-based reconstructions, but also simulation biases. For this purpose, we employed the modern potential natural vegetation distribution (spatial resolution: 5 arc minutes) provided by Ramankutty et al. (2010). It represents the world's vegetation cover that would have most likely existed for 1986–

175   1995 C.E. in equilibrium with present-day climate and natural disturbances in the absence of human activities (Ramankutty and Foley, 1999). To allow direct comparisons between reconstructions and simulations, as well as among simulations at the hemispheric or continental scales, we aggregated the modern potential natural vegetation types into modern potential megabiomes (Fig. 1) following Dallmeyer et al. (2019).

To assess the accuracy of the pollen-based reconstructions and ESM-based simulations, we calculated the

180   proportion of sites where reconstructed or simulated megabiomes at timeslice 0 cal. ka BP match these modern potential megabiomes. For each site, the simulated megabiome at timeslice 0 cal. ka BP and the modern potential megabiome were extracted from the grid-cells in which the site is located.

2.5 Methods for comparison of the simulated and reconstructed megabiome datasets

The Earth mover's distance (EMD), which takes into account the uncertainties of the biomized data and case-

185   specific weighted distances (Chevalier et al., 2023b), was applied to quantify the mismatch between the pollen-based reconstructions and ESM-based simulation ensemble at each site. Specifically, the EMD calculates the distance between the reconstruction and simulation ensemble by considering the entire range of megabiome affinity scores. This means that the details of the underlying vegetation structure are part of the comparison, in contrast to commonly used methods that solely compare the megabiome with the highest affinity score estimated

190   from the reconstructions or simulations. To match the distribution of megabiome scores obtained from biomization algorithms, we translated the frequencies of the six simulated megabiomes into a simulated megabiome affinity score set. For example, for an ensemble of simulations with two boreal forests and four temperate forests in its six simulations, the affinity scores for the boreal and temperate forests would be 2/6 and 4/6, respectively, while the affinity scores for the remaining megabiomes are zero. In addition, we developed an ecological and climatic

195   distance-based (Allen et al., 2020; Sato et al., 2021) EMD weighting scheme (Table 2) to penalize mismatches between the reconstructions and simulation ensemble in terms of differences in vegetation structure (i.e., forest megabiomes, non-forest megabiomes, and deserts) and climate zone preferences (i.e., tropical, warm-temperate (subtropical), temperate, boreal, and polar regions). In this study, we assume that the basal distance between two different megabiomes with the same vegetation structure and climate zone is set to 1. Then, each difference in

200   vegetation structure or climate zone adds an extra weight of 1. For example, the reconstructed tropical forest has a distance weight of two from the simulated temperate forest and three from the simulated boreal forest. The EMD routines were implemented by using the *paleotools* package in R (version 0.1.0; https://github.com/mchevalier2/paleotools, last access: March 13, 2024; Chevalier, 2023a).

205 **Table 2 Earth mover's distance (EMD) weighting scheme for ecological and climatic distances between the pollen-based reconstructed and simulated megabiomes used in this study.** Higher values in the table indicate a greater ecological or climatic distance between the reconstructed and simulated megabiomes. Megabiome code: TRFO- tropical forest, WTFO- warm-temperate (subtropical) forest, TEFO- temperate forest, BOFO- boreal forest, SAVA- (warm) savanna and dry woodland, STEP- grassland and dry shrubland, DESE- (warm) desert, TUND- tundra and polar desert. Of these, TRFO,

210 WTFO, TEFO, and BOFO are forest megabiomes, whereas the others are non-forest megabiomes.

Reconstruction vs. Simulation	TRFO	WTFO	TEFO	BOFO	SAVA	STEP	DESE	TUND
TRFO	0	1	2	3	1	2	3	4
WTFO	1	0	1	2	2	2	3	3
TEFO	2	1	0	1	3	2	3	2
BOFO	3	2	1	0	4	3	2	1
SAVA	1	2	3	4	0	1	2	4
STEP	2	2	2	3	1	0	1	1
DESE	3	3	3	2	2	1	0	1
TUND	4	3	2	1	4	1	1	0

We aggregated the sites into regular longitude-latitude grid-cells of size 3.75°x3.75° to reduce the sampling bias from the non-uniform spatial distribution of sites and to facilitate a more direct model-data comparison. That is, the reconstructed or simulated megabiomes of a grid-cell at each timeslice were derived from the reconstructed

215 or simulated megabiomes with the highest frequencies among the available sites in that grid-cell. In addition, the data-model EMDs of a grid-cell were derived from the median EMDs of available sites in that grid-cell. We also created an ice-sheet ensemble set with a spatial resolution of 3.75°, synthesized from the maximum extent of ICE-5G, ICE-6G, and GLAC-1D reconstructions, for fair comparisons among simulations.

To cluster the regions, we performed the dynamic time warping with the time series of the data-model EMDs of

220 all grid-cells on each continent, which allows time series to be grouped based on their patterns or shapes (Müller et al., 2007). The number of clusters was determined using the elbow method (Syakur et al., 2018) and adjusted based on the sample availability. The global data-model EMDs time series, representing the global mean dynamics, was then synthesized from the median EMDs for each clustered region. The dynamic time warping algorithm was implemented by using the *dtwclust* package in R (version 5.5.12; https://github.com/asardaes/dtwclust, last access:

225 March 17, 2024; Sarda-Espinosa, 2023).

3. RESULTS AND DISCUSSION

3.1 Evaluation of megabiome reconstructions and simulations for the present-day

3.1.1 Pollen-based reconstructions

230 We consider global-scale, pollen-based megabiome reconstructions reliable because site-by-site comparisons of reconstructed megabiomes at timeslice 0 cal. ka BP from 2,232 available sites with modern potential megabiomes indicate an 80.2% agreement (Table 3). Incorrectly assigned megabiomes are distributed all over the world, and no systematic mismatch is evident (Fig. 1a). We assume that the high agreement not only originates from the high quality of the pollen data set used with respect to taxonomical and temporal harmonization, but also relates to the

235 fact that the biomization method employs updated and harmonized schemes assigning pollen taxa to plant

functional types to megabiomes. As a result, we argue that the data quality as well as the higher spatial and temporal coverage compared to previous biomization studies (Bigelow et al., 2003; Marinova et al., 2018) make our pollen-based megabiome reconstruction a robust dataset for various applications.

240 **Table 3 Agreement of modern potential megabiomes, aggregated from modern potential natural vegetation, with (a) pollen-based reconstructions and (b-h) simulations at 0 cal. ka BP.** We use a set of six transient simulations that have been run in an Earth System Model: **(c-d)** MPI-ESM (MPI-ESM_GLAC1D, MPI-ESM_ICE6G), **(e-f)** CLIMBER-X (CLIMBER-X_GLAC1D, CLIMBER-X_ICE6G), and **(g-h)** CCSM3 (TRACE-21K-I_ICE5G, TRACE-21K-II_ICE5G), as well as **(b)** ESM-representative megabiome that occur most frequently in the set of simulations. The megabiome codes are given in Table 2.

	Record number at 0 ka	(a) Pollen-based reconstruction	(b) ESM-representative Simulation	MPI-ESM		CLIMBER-X		CCSM3	
				(c) MPI-ESM_GLAC1D	(d) MPI-ESM_ICE6G	(e) CLIMBER-X_GLAC1D	(f) CLIMBER-X_ICE6G	(g) TRACE-21K-I_ICE5G	(h) TRACE-21K-II_ICE5G
TRFO	112	81.2 %	86.6 %	71.4 %	62.5 %	81.2 %	69.6 %	62.5 %	56.2 %
WTFO	59	78.0 %	49.2 %	35.6 %	10.2 %	50.8 %	49.2 %	42.4 %	42.4 %
TEFO	1249	86.9 %	74.0 %	77.0 %	75.0 %	66.0 %	49.3 %	11.2 %	15.5 %
BOFO	464	79.1 %	52.4 %	40.5 %	35.3 %	49.6 %	38.4 %	14.7 %	36.0 %
SAVA	57	77.2 %	7.0 %	3.5 %	29.8 %	1.8 %	0.0 %	0.0 %	1.8 %
STEP	163	52.8 %	45.4 %	20.9 %	33.7 %	38.0 %	40.5 %	43.6 %	38.0 %
DESE	22	72.7 %	50.0 %	59.1 %	45.5 %	18.2 %	40.9 %	45.5 %	50.0 %
TUND	106	50.0 %	46.2 %	40.6 %	29.2 %	28.3 %	33.3 %	59.4 %	55.7 %
Overall	2232	80.2 %	64.1 %	60.2 %	57.8 %	57.0 %	45.3 %	20.0 %	26.1 %

245

Megabiome

- Tropical forest (TRFO)
- (Warm) savanna and dry woodland (SAVA)
- Warm-temperate (subtropical) forest (WTFO)
- Grassland and dry shrubland (STEP)
- Temperate forest (TEFO)
- (Warm) desert (DESE)
- Boreal forest (BOFO)
- Tundra and polar desert (TUND)
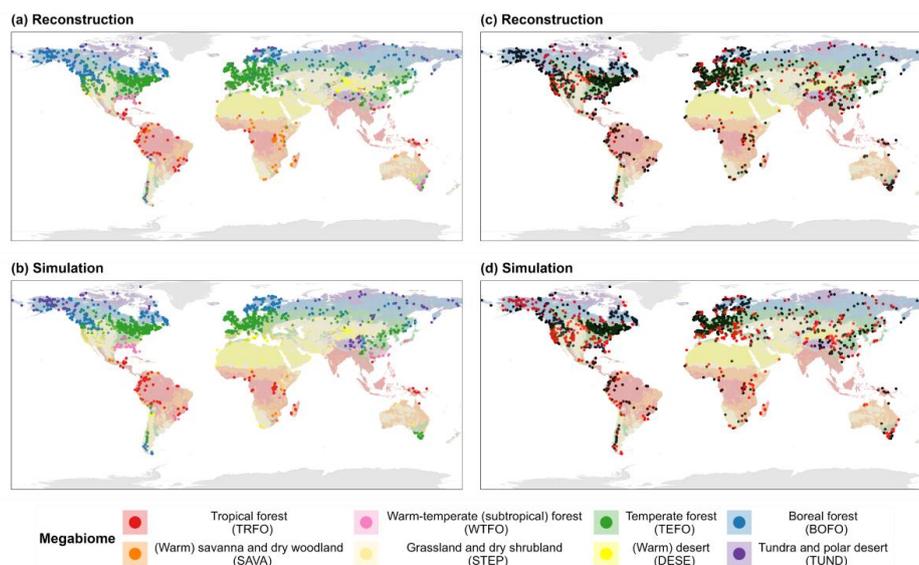
Figure 1 (a-b) Spatial distributions of megabiomes at 0 cal. ka BP, and (c-d) their agreement (matches in black, mismatches in red) with modern potential megabiomes, for each site derived from the pollen-based reconstruction and the ESM-based simulation ensemble. Shown here are the ESM-representative megabiomes that occur most frequently in the

250 set of simulations. The background depicts modern potential megabiomes (Dallmeyer et al., 2019) aggregated from modern potential natural vegetation (Ramankutty and Foley, 1999; Ramankutty et al., 2010), representing the world's vegetation cover that had most likely existed for 1986–1995 C.E. in equilibrium with present-day climate and natural disturbance in the absence of human activities.

8

Several factors may contribute to the incorrect reconstruction of modern potential megabiomes in our study (Fig.
1a). (a) The different pollen representation (including production, dispersion, and preservation) of plant taxa is
the principal reason for inadequate separation of forest and open landscape ecotones. For example, the high pollen
productivity of key taxa (such as *Artemisia*; Xu et al., 2014) has resulted in an overestimation of grasslands and
dry shrublands (STEP) in the East Asian summer monsoon northern marginal zone and the Great Plains of North
America. However, woody PFTs are generally not assigned to non-forest megabiomes in the PFTs-megabiome
assignment scheme, leading to a potentially incorrect prioritization of forest megabiomes in cases of woody pollen
grain occurrences (from long-distance transportation or local existence) in the samples (Marinova et al., 2018;
Chen et al., 2010). (b) The low taxonomic resolution could also cause a mismatch between neighboring forest
megabiomes. The woody taxa at the genus level rather than at species level in pollen identification reduces the
amount of ecological information available for PFT assignment (Chen et al., 2010). For example, different species
from *Pinus*, *Alnus*, *Fagus*, and *Betula* (Tian et al., 2018) have different bioclimatic controls, phenology, and life
forms, but identification at the genus level results in them being shared by key PFTs in different forest megabiomes
when assigning taxa to PFTs. A typical area in which this problem occurs is southern Scandinavia. Pollen grains
from *Betula pendula* in temperate forests and *Betula pubescens* in boreal forests (Beck et al., 2016) in this region
can only be identified to genus level, resulting in these two key species not being able to serve as indicators to
distinguish between temperate and boreal forests. (c) Anthropogenic modification of pollen assemblages has, to
some extent, contributed to mismatches in forested areas. For example, incorrectly reconstructed grasslands and
dry shrublands (STEP) in North China and Southern Europe may reflect intensive land use (e.g., deforestation).
However, the modern anthropogenic megabiomes are not well reconstructed at a broad spatial scale here, as with
previous studies (Ni et al., 2014; Cao et al., 2022). We suggest that this may be related to the absence of
anthropogenic PFTs and megabiomes in our taxa-PFT-megabiome assignment schemes, as well as the difficulty
of distinguishing between anthropogenic and non-anthropogenic pollen when using genus or family levels (e.g.,
Poaceae, Rosaceae), and pollen samples generally being collected from sites with less human disturbance.

3.1.2 ESM-based simulations

The agreement between modern potential megabiomes and simulated megabiomes at timeslice 0 cal. ka BP is
higher for the ESM-representative megabiome (cf. Sect. 2.3) than for individual ESM-based simulation (64.1%
vs. 20.0–60.2%; Table 3). As a result, the ESM-representative megabiome depicts more reliable patterns of
megabiome dynamics and distribution than individual simulations, with higher agreement especially in Alaska,
the Iberian Peninsula, the Atlantic Coastal Plain of North America, the southeastern United States, and the Alps
(Fig. 1b and Fig. A2). However, there are still certain regions with low agreement due to climatic biases. These
include nearly all highlands (such as the Rocky Mountains in North America, East African plateau, and Great
Dividing Range in Australia) for which an overestimation of the temperature can be expected in the models due
to a much lower mean orography than reality caused by the smoothing in the coarse spatial resolution (3.75°x3.75°
and 5°x5°) of the model grids. All models simulate non-forest megabiomes instead of forest in the Mediterranean
region indicating that the models simulate a climate that is too dry, at least seasonally. The TRACE-21K
simulation as well as the MPI-ESM simulations fail to reproduce the boreal forest (BOFO) in Alaska, which is
then also reflected in the ESM-representative megabiomes. Similar to the reconstructions, the transition zones
between temperate forest (TEFO) and non-forest megabiomes, such as the East Asian summer monsoon margin,

are regions with lower simulated megabiome agreement to the modern potential megabiome distribution. In North Africa, the models also tend to underestimate the northern extension of the grassland and dry shrubland (STEP)

295 and incorrectly assign (warm) savanna and dry woodland (SAVA) sites to tropical forest (TRFO). This is related to the biomization procedure for the model results that only relies on simulated vegetation cover fractions and simulated climate, whereas savannas are additionally determined by other ecological processes such as fire intensity and frequency (Dallmeyer et al., 2019) or grazing (van Langevelde et al., 2019).

300 **3.2 Global megabiome dynamics and distributions over the last 21,000 years**

We present a global assessment of megabiome dynamics and distributions derived from pollen-based reconstructions and ESM-based simulations over the last 21,000 years, with a temporal resolution of 500 years. Overall, there has been a global shift from open glacial non-forest megabiomes to Holocene forest megabiomes since the LGM (Fig. 2), in line with the general climate warming trend and continental ice-sheet retreat (Fig. 3):
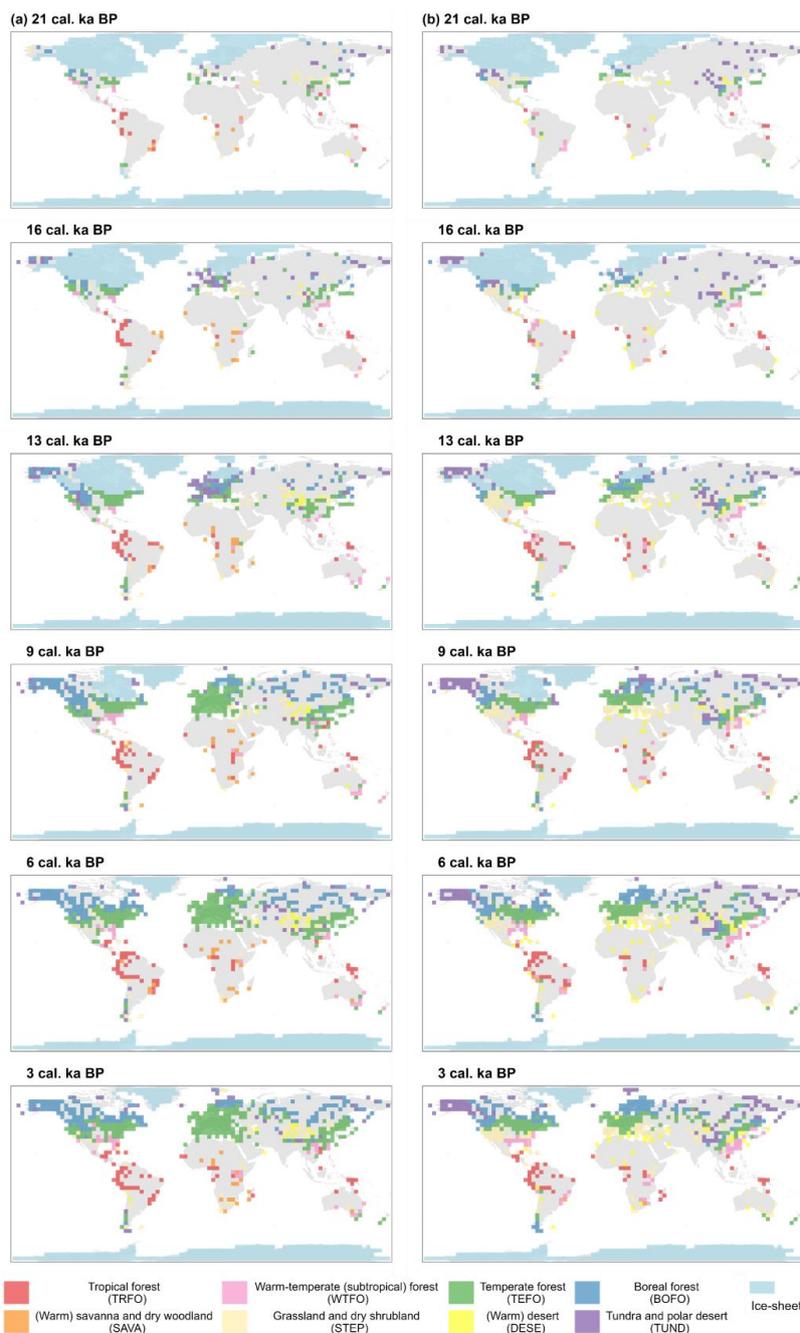
Figure 2 Spatial distributions of megabiomes, derived from the (a) pollen-based reconstruction and (b) the ESM-based simulation ensemble, as well as the ice-sheet ensemble, at 21, 16, 13, 9, 6, and 3 cal. ka BP based on grid-cells of 3.75°x3.75°. Shown here are the ESM-representative megabiomes that occur most frequently in the set of simulations. The ice sheets are shown at their maximum extent at timeslices synthesized for the ICE-5G (Peltier, 2004), ICE-6G (Peltier et al., 2015), and GLAC-1D (Tarasov et al., 2012) reconstructions.
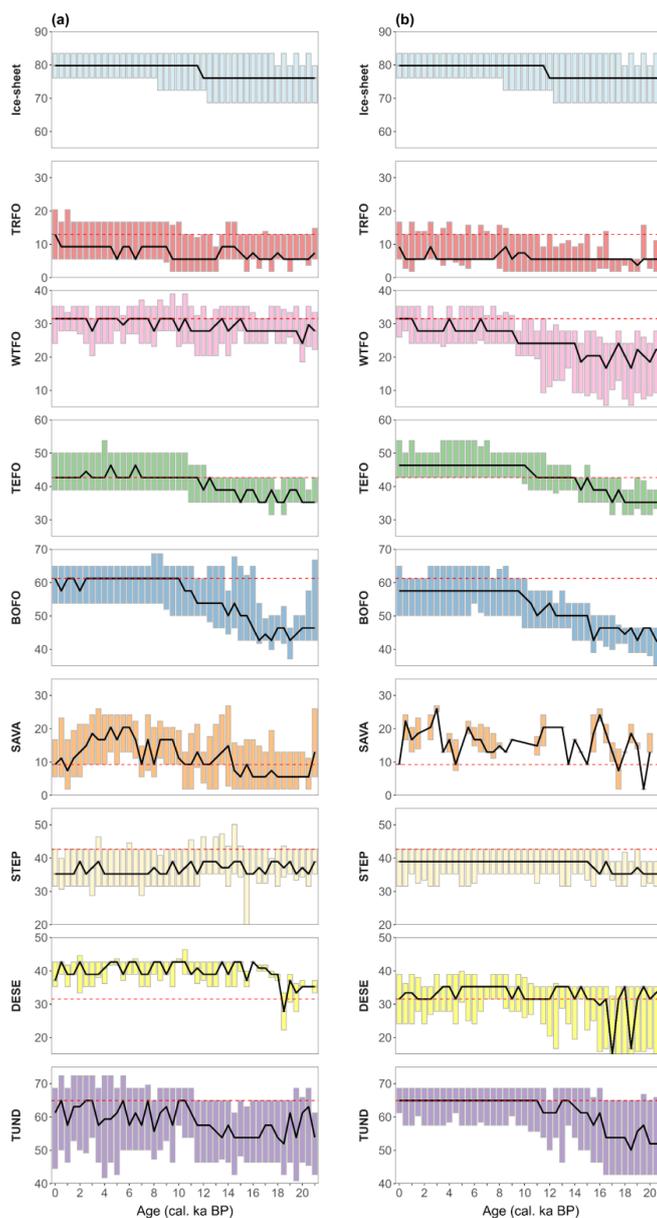
Figure 3 Temporal changes in the latitudinal location (°) of each megabiome, derived from the **(a)** pollen-based reconstruction and **(b)** ESM-based simulation ensemble, as well as the ice-sheet ensemble, based on grid-cells of 3.75°x3.75° over the last 21,000 years globally. The red dashed lines are the median latitudinal location of the corresponding modern potential megabiomes, derived from grid-cells including pollen samples at 0 cal. ka BP. The black solid line represents the median latitude for each timeslice, while the top and lower boundaries of each box represent the upper and lower quartiles of latitude distribution for that timeslice. Megabiome code: TRFO - tropical forest, WTFO - warm-temperate (subtropical) forest, TEFO - temperate forest, BOFO - boreal forest, SAVA - (warm) savanna and dry woodland, STEP - grassland and dry shrubland, DESE - (warm) desert, TUND - tundra and polar desert.

320 LGM (represented by the timeslice 21 cal. ka BP): TUND and BOFO dominate the high latitudes and periglacial areas, whereas the relatively warm forest megabiomes (e.g., WTFO and TEFO) are distributed at lower latitudes than present-day, in response to cold and dry climates (Nolan et al., 2018). However, the ESM-representative megabiome (simulations hereafter in this Sect.) reveals more non-forest megabiomes (such as TUND and STEP) in periglacial areas of North America (e.g., Alaska and the Rocky Mountains) and northern Asia (e.g., northeastern 325 Siberia), as well as in the Mediterranean regions, as compared to the reconstructions. Although previous pollen-based biomization studies with different biomization schemes have reported similar results (such as Binney et al., 2017 and Cao et al., 2019 in periglacial areas; Elenga et al., 2000 in the Mediterranean regions), assessments of modern megabiome distributions suggest that these studies overestimated the occurrence of non-forest megabiomes in these regions. Furthermore, STEP occurred in central Asia in the reconstructions rather than 330 TUND in the simulations, and TRFO and SAVA appeared in tropical South America and Africa in the reconstructions rather than WTFO in the simulations.

Deglaciation (represented by the timeslices 16 and 13 cal. ka BP): Compared with the LGM, the extratropical megabiomes experienced a remarkable expansion to higher latitudes that coincided with the retreat of the continental ice sheets (Fig. 3). In particular, BOFO, TUND, and TEFO underwent a more extensive expansion 335 compared to the other megabiomes in both our reconstructions and simulations; a result similar to previous biomization studies (such as Binney et al., 2017 and Cao et al., 2019 in north of 30°N). However, in contrast to the expansion of forest megabiomes (mostly TEFO and BOFO) in the reconstructions of the Rocky Mountains, northeastern Siberia, and the Mediterranean regions, more non-forest megabiomes (mostly STEP and TUND) occurred in the simulations. TRFO and SAVA expanded in the reconstructions of tropical South America and 340 Africa, whereas the simulations show a shift from WTFO to TRFO since the LGM. In Australia, the Great Dividing Range region was dominated by WTFO in the reconstructions and STEP in the simulations.

Early Holocene (represented by the timeslice 9 cal. ka BP): During this period, the global spatial patterns of megabiome distributions have shifted to closely resemble those of the present-day. That is, forest megabiomes replaced the glacial non-forest megabiomes during the early Holocene and expanded to similar distributional 345 positions as those of today. For example, as the ice sheets receded in the Northern Hemisphere, BOFO continued to move northward and dominated the northern Rockies during the early Holocene, with distributions comparable to today, inferred from both reconstructions and simulations. Due to the extended and homogenized dataset used here, our study also challenges the previous regional-based views that similar distribution patterns of modern megabiomes (Binney et al., 2017) and maximum forest expansion occurred in the mid-Holocene (Ni et al., 2014; 350 Tian et al., 2018). However, mismatches persist between our reconstructions and simulations. For example, Scandinavia was dominated by TEFO and BOFO in the reconstructions but BOFO and TUND in the simulations; Alaska and the Mediterranean regions shifted to BOFO and TEFO, respectively, in the reconstructions, while TUND and STEP remained dominant in the simulations.

Mid-Holocene to Late Holocene (represented by the timeslices 6 and 3 cal. ka BP): The spatial patterns of 355 megabiome distributions during this period are only slightly different from those of the early Holocene. TRFO, for example, expanded in Mesoamerican reconstructions and simulations. It is also worth noting that forest megabiomes in certain areas (represented by North China and southern Europe) have gradually been partially

replaced by STEP since the Late Holocene in the reconstructions, in contrast to the dominance of forests in the

simulation. Given that the simulated vegetation was in a quasi-equilibrium with climate and unaffected by humans,

360 enhanced human disturbance is proposed as the most plausible driver of forest degradation over this time period

(Stephens et al., 2019; Cao et al., 2022).

3.3 Comparison of pollen-based and ESM-based simulated megabiome reconstructions

To identify regions and periods with the largest deviations between pollen- and model-derived megabiome

365 distributions, as well as to infer regional contributions to such deviations, we calculated their EMDs at each

available timeslice and grid-cell (Fig. 4a). Following that, we aggregated the EMD time-series over all grid-cells

into 15 regional clusters (Fig. 4b) and synthesized the median EMDs over these regional clusters as representative

of the global mean dynamic.
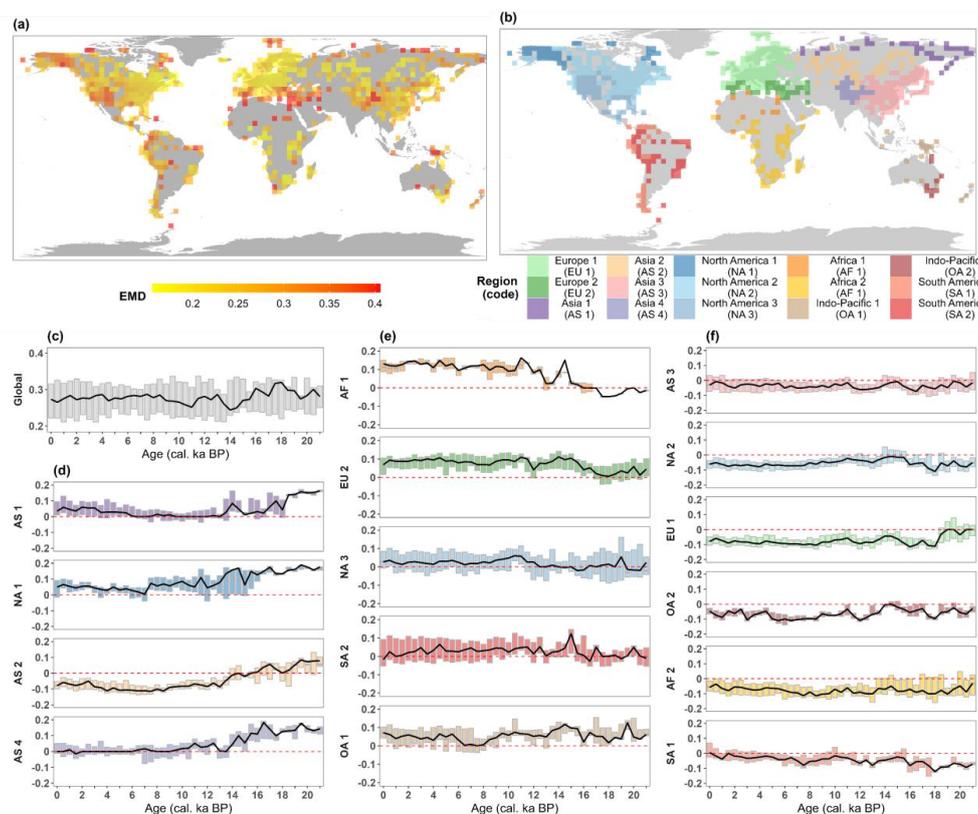
370 **Figure 4 Spatiotemporal patterns of Earth mover's distance (EMD) between the pollen-based reconstructions and
ESM-based simulation ensemble over the last 21,000 years, based on grid-cells of 3.75°x3.75°. (a)** Spatial pattern of the
median data-model EMD of available timeslices over the last 21,000 years. Highest EMD values and consequently largest
data-model deviations occur especially in the Mediterranean, North Africa, highlands (such as the Rocky Mountains), and

circum-Arctic areas. Note that the map legend shows EMD values from the 5th to 95th percentile, with values above the 95th percentile shown in the 95th percentile color and values below the 5th percentile in the 5th percentile color. **(b)** Regional clustering of the data-model EMD time-series for grid-cells using Dynamic Time Warping by continent. **(c)** The global data-model EMD at each timeslice, synthesized from the median EMDs of the clustered regions at that timeslice. The largest data-model deviations occur during the LGM and early deglaciation periods. The solid black line represents the median EMD for each timeslice, while the top and lower boundaries of each box represent the upper and lower quartiles of EMD distribution for that timeslice. **(d-f)** The data-model EMD as an anomaly to the global median in clustered regions at each timeslice. That is, regions with **(d)** the highest data-model EMD during the LGM and the early deglaciation, **(e)** the data-model EMD that increases with time during the Holocene, and **(f)** the lower data-model EMD than the global level. Colors and region codes in the boxplots correspond to the colors and region codes of the clusters displayed on the map. The red dashed line is the zero value of EMD.

The largest EMD-assessed deviations between pollen- and model-derived megabiome distributions on a global scale occur during the LGM and early deglaciation (~21–16 cal. ka BP; Fig. 4c). In contrast, the best data-model agreement occurs during the Bølling-Allerød interstadial (represented by the timeslice 14 cal. ka BP) and Early Holocene periods (represented by the timeslice 11 cal. ka BP). Furthermore, the global median EMD stays relatively constant at moderate values over the last 9,000 years.

A closer look at the data-model EMD dynamics of the 15 regions (Fig. 4b) identified by the dynamic time warping reveals three sub-clusters. First, regions in which the data-model EMD is particularly high during the LGM and the early deglaciation (Fig. 4d), driving the strong global data-model mismatch during this period. Second, regions in which the data-model EMD rather increases with time (Fig. 4e), contributing to the moderate global EMD values during the Holocene. Third, regions in which the data-model EMD are predominantly lower than the global median EMD (Fig. 4f), i.e., high data-model agreement. However, the reasons for the regional data-model mismatch are very different.

Different estimates of tundra in the circum-Arctic areas and the Tibetan Plateau are the primary sources of the strong global data-model deviations during the LGM and early deglaciation periods (Fig. 4d) at 21 and 16 cal. ka BP (Fig. 3). We observe inconsistent estimates of tundra (TUND) and boreal forest (BOFO) from the pollen-based reconstructions and the ESM-based simulations in northern Siberia (AS1), Alaska (NA1), and the East Siberian Highlands (AS2). To some extent, this mismatch could be attributed to systematic model biases in the simulated climate, as climate models tend to underestimate moisture and summer temperature in the periglacial areas compared to proxy-based reconstructions, as previously indicated in studies with different models (Deplazes et al., 2013; Alley, 2000) for that period. We assume that the simulations used in this study share this rather common problem of a cold bias in boreal latitudes, resulting in the overestimation of tundra in simulations.

The large data-model deviations on the Tibetan Plateau (AS4) result from different estimates of tundra and grasslands (STEP) in the simulations and reconstructions. Given that the simulated megabiome in that area at timeslice 0 cal. ka BP closely resembles modern potential natural vegetation distributions when compared to the reconstructions (Fig. 1 and Fig. A2), we assume that tundra may have been misrepresented as grassland in the reconstructions. This misrepresentation can be attributed to the alpine tundra sharing dominant characteristic species of Poaceae and Cyperaceae with grasslands, which are defined by fewer PFTs and thus are preferentially allocated.

Different estimates of non-forest megabiomes in relatively semi-arid zones such as North Africa and the Mediterranean have moderate but increasing data-model deviations since the early deglaciation (Fig. 4e). As

415 shown in Figure 3, with the transition from the glacial to the Holocene, the Mediterranean-Black Sea-Caspian Corridor (EU2) and the Mediterranean coast of northern Africa have gradually been dominated by temperate forests (TEFO) in the reconstructions, rather than grasslands and dry shrublands (STEP) in the simulations. The Mediterranean region has warm to hot dry summers and mild wet winters. Modeling studies report systematic model biases of too-warm summers and too-dry winters in this region (García-Herrera and Barriopedro, 2018).

420 Similar climate biases could also be the reason for the underrepresentation of the cover fraction of woody PFTs in the simulations. In addition, the data-model deviations for the Sahara (AF1) are mainly in the Holocene, resulting from a mismatch between desert (DESE) in the simulations and savanna (SAVA) in the reconstructions. The North African monsoon system has weakened and thus the desert expanded within the Holocene in the simulations due to the seasonal changes in insolation, as evidenced both by proxy-based reconstructions

425 (deMenocal et al., 2000; Shanahan et al., 2015) and simulations (Dallmeyer et al., 2021). In our reconstructions, the overrepresentation of woody taxa (e.g., *Acacia* and Arecaceae) results in the classification of deserts as savanna and dry woodlands (SAVA), which may have contributed to the increasing data-model deviations in the Sahara during the Holocene.

430 **4. Summary and Conclusions**

We present a global assessment of megabiome dynamics and distributions derived from pollen-based reconstructions and ESM-based simulations over the last 21,000 years, with a temporal resolution of 500 years. The reconstructions and simulations both reveal a global shift from open glacial non-forest megabiomes to Holocene forest megabiomes, with a megabiome distribution similar to today's establishing during the early

435 Holocene. The largest global-scale deviations between pollen- and model-derived megabiome distributions occurred during the LGM and early deglaciation, mainly in the circum-Arctic areas and Tibetan Plateau. In addition, North Africa and the Mediterranean regions contribute to moderate global data-model deviations during the Holocene. On the whole, our results are suitable for the evaluation of ESM-based paleo-megabiome simulations, as well as providing clues for improving systematic model biases.

440

Code and data availability

The LegacyPollen 2.0 dataset is open access at PANGAEA (https://doi.pangaea.de/10.1594/PANGAEA.965907; Li et al., 2024) and provides both count and percentage pollen data. The dataset files are stored in machine-readable data format (.csv) and are already separated into western North America (west of 105°W; Williams et

445 al., 2000), eastern North America, Europe, Asia, South America, Africa, and the Indo-Pacific for easy access and use. We have provided an overview table of site metadata and the taxa harmonization table at PANGAEA, as in the LegacyPollen 1.0 dataset (Herzschuh et al., 2022).

The simulation MPI-ESM_ICE6G and an equivalent simulation to MPI-ESM_GLAC1D for the biomization tool are available from the Word Data Centre of Climate at https://doi.org/10.26050/WDCC/PMMXMCRTDIP122

450 (last access: May 16, 2024; Mikolajewicz et al., 2023) and https://doi.org/10.26050/WDCC/PMMXMCHTD (last access: May 16, 2024; Kleinen et al., 2023), respectively. The input data of TRACE-21k-I and TRACE-21k-II for

the biomization tool can be downloaded from https://www.earthsystemgrid.org/project/trace.html (last access: May 16, 2024) and https://trace-21k.nelson.wisc.edu/portal.html (last access: May 16, 2024), respectively. The CLIMBER-X simulation is not published, but the input data for the biomization tool can be provided upon request.

455 The data of modern potential natural vegetation distributions estimated by Ramankutty et al. (2010) can be downloaded from https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=961 (last access: May 16, 2024). The ice-sheet data for ICE-5G (Peltier, 2004) and ICE-6G (Peltier et al., 2015) reconstructions can be downloaded from http://www.atmosp.physics.utoronto.ca/~peltier/data.php (last access: May 16, 2024), and for GLAC-1D (Tarasov et al., 2012) reconstructions can be downloaded from

460 https://pmip4.lsce.ipsl.fr/doku.php/data:ice_glac_1d#download (last access: May 16, 2024).

The pollen-based biomization algorithm in R and the tool for the biomization of simulated PFT cover fractions are available from Zenodo (https://doi.org/10.5281/zenodo.7523423, last access: May 10, 2023; Cao and Tian, 2021) and MPG.PuRe repository (https://hdl.handle.net/21.11116/0000-0001-B800-F, last access: May 16, 2024; Dallmeyer et al., 2019), respectively. The R packages *paleotools* (version 0.1.0; Chevalier, 2023a) and *dtwclust*

465 (version 5.5.12; Sarda-Espinosa, 2023) that execute the Earth mover's distance (EMD) routines and the dynamic time warping algorithm are available from https://github.com/mchevalier2/paleotools (last access: March 13, 2024) and https://github.com/asardaes/dtwclust (last access: March 17, 2024), respectively.

The megabiome data estimated by pollen-based reconstructions and ESM-based simulations are freely available from the MPG.PuRe repository in both Excel worksheet (.xlsx; includes an overview of site metadata,

470 reconstructed and simulated megabiome, and normalized megabiome affinity score set) and network common data form (.nc; T31 grid cell-based) formats. Furthermore, the lists of taxa-PFTs and PFTs-megabiome assignments per continent used in the biomization procedure are available from the MPG.PuRe repository.

Author contributions

475 UH, CL, and AD designed the study. CL and AD performed pollen-based reconstruction and model-based biomization, respectively. CL, JN, and A.A. revised and updated the taxa-PFTs-megabiome assignment schemes in the biomization procedures under the supervision of UH. CL implemented the analysis under the supervision of UH and AD. MW provided the CLIMBER-X simulation. MC and LS contributed to the analytical methods. XC contributed an initial R script for biomization procedures. BH checked the PANGAEA Event for the

480 LegacyPollen 2.0 dataset. CL wrote the first draft of the manuscript under the supervision of UH and AD. All co-authors discussed the results and contributed to the final manuscript.

Competing interests

UH, MC, and AD are guest members of the editorial board of Climate of the Past for the special issue "Past vegetation dynamics and their role in past climate changes". The authors have no other competing interests to

485 declare.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

510   **Appendix A**

Figure A1 Location and source of 3,691 fossil pollen records in the LegacyPollen 2.0 dataset. The background depicts modern potential megabiomes (Dallmeyer et al., 2019) aggregated from modern potential natural vegetation (Ramankutty and Foley, 1999; Ramankutty et al., 2010).

515
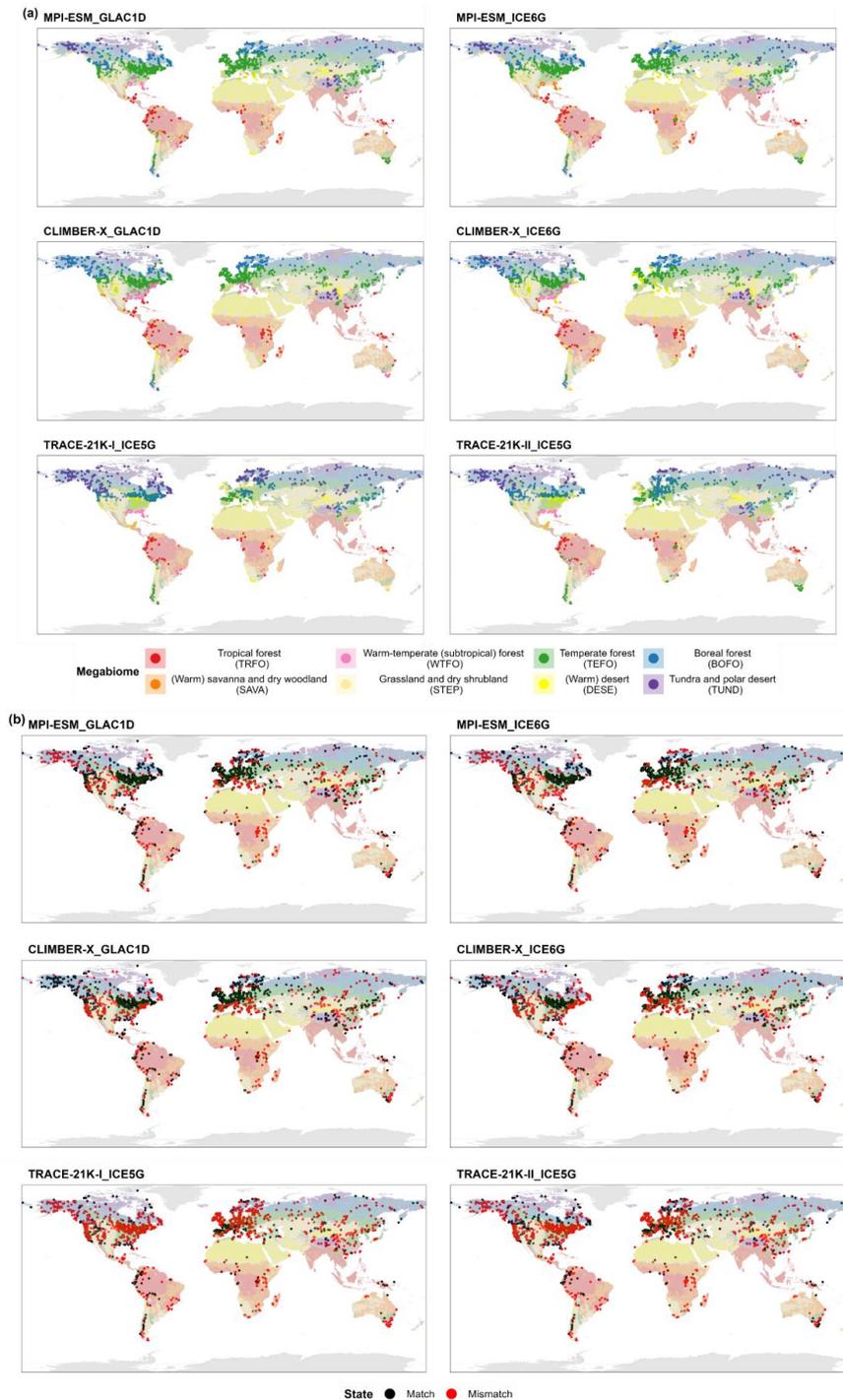
Figure A2 (a) Spatial distributions of ESM-based simulated megabiomes at 0 cal. ka BP, and (b) their agreement with

520   **modern potential megabiomes, for each site derived from MPI-ESM, CLIMBER-X, and TRACE-21K.**

References

Allen, J. R. M., Forrest, M., Hickler, T., Singarayer, J. S., Valdes, P. J., and Huntley, B.: Global vegetation patterns of the past 140,000 years, J. Biogeogr., 47, 2073–2090, https://doi.org/10.1111/jbi.13930, 2020.

Alley, R. B.: The Younger Dryas cold interval as viewed from central Greenland, Quat. Sci. Rev., 19, 213–226, https://doi.org/10.1016/S0277-3791(99)00062-1, 2000.

Beck, P., Caudullo, G., de Rigo, D., and Tinner, W.: Betula pendula, Betula pubescens and other birches in Europe: distribution, habitat, usage and threats, in: European Atlas of Forest Tree Species, edited by: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A., Publication Office of the European Union, Luxembourg, 70–73, https://doi.org/ 10.7892/boris.80789, 2016.

Berger, A.: Long-term variations of daily insolation and Quaternary climatic changes, J. Atmos. Sci., 35, 2362–2367, https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2, 1978.

Bigelow, N. H., Brubaker, L. B., Edwards, M. E., Harrison, S. P., Prentice, I. C., Anderson, P. M., Andreev, A. A., Bartlein, P. J., Christensen, T. R., Cramer, W., Kaplan, J. O., Lozhkin, A. V., Matveyeva, N. V., Murray, D. F., McGuire, A. D., Razzhivin, V. Y., Ritchie, J. C., Smith, B., Walker, D. A., Gajewski, K., Wolf, V., Holmqvist, B. H., Igarashi, Y., Kremenetskii, K., Paus, A., Pisaric, M. F. J., and Volkova, V. S.: Climate change and Arctic ecosystems: 1. Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present, J. Geophys. Res. Atmos., 108, https://doi.org/10.1029/2002JD002558, 2003.

Binney, H., Edwards, M., Macias-Fauria, M., Lozhkin, A., Anderson, P., Kaplan, J. O., Andreev, A., Bezrukova, E., Blyakharchuk, T., Jankovska, V., Khazina, I., Krivonogov, S., Kremenetski, K., Nield, J., Novenko, E., Ryabogina, N., Solovieva, N., Willis, K., and Zernitskaya, V.: Vegetation of Eurasia from the last glacial maximum to present: Key biogeographic patterns, Quat. Sci. Rev., 157, 80–97, https://doi.org/10.1016/j.quascirev.2016.11.022, 2017.

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.

Cao, X. and Tian, F.: Pollen-based biome reconstruction in R, Zenodo [code], https://doi.org/10.5281/zenodo.7523423, 2021.

Cao, X., Tian, F., Dallmeyer, A., and Herzschuh, U.: Northern Hemisphere biome changes (>30°N) since 40 cal ka BP and their driving factors inferred from model-data comparisons, Quat. Sci. Rev., 220, 291–309, https://doi.org/10.1016/j.quascirev.2019.07.034, 2019.

Cao, X., Tian, F., Herzschuh, U., Ni, J., Xu, Q., Li, W., Zhang, Y., Luo, M., and Chen, F.: Human activities have reduced plant diversity in eastern China over the last two millennia, Glob. Change Biol., 28, 4962–4976, https://doi.org/10.1111/gcb.16274, 2022.

Chen, Y., Ni, J., and Herzschuh, U.: Quantifying modern biomes based on surface pollen data in China, Global Planet. Change, 74, 114–131, https://doi.org/10.1016/j.gloplacha.2010.09.002, 2010.

Chevalier, M.: mchevalier2/paleotools: The version of the package used in the Climate of the Past EMD manuscript, Zenodo [code], https://doi.org/10.5281/zenodo.7889631, 2023a.

Chevalier, M., Dallmeyer, A., Weitzel, N., Li, C., Baudouin, J, P., Herzschuh, U., Cao, X., and Hense, A.: Refining data–data and data–model vegetation comparisons using the Earth mover's distance (EMD), Clim. Past, 19, 1043–1060, https://doi.org/10.5194/cp-19-1043-2023, 2023b.

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.

Dallmeyer, A., Claussen, M., and Brovkin, V.: Harmonising plant functional type distributions for evaluating Earth system models, Clim. Past, 15, 335–366, https://doi.org/10.5194/cp-15-335-2019, 2019.

525
530
535
540
545
550
555
560
565
570

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.

Dallmeyer, A., Kleinen, T., Claussen, M., Weitzel, N., Cao, X., and Herzschuh, U.: The deglacial forest conundrum, Nat. Commun., 13, 6035, https://doi.org/10.1038/s41467-022-33646-6, 2022.

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

Deplazes, G., Lückge, A., Peterson, L. C., Timmermann, A., Hamann, Y., Hughen, K. A., Röhl, U., Laj, C., Cane, M. A., Sigman, D. M., and Haug, G. H.: Links between tropical rainfall and North Atlantic climate during the last glacial period, Nat. Geosci., 6, 213–217, https://doi.org/10.1038/ngeo1712, 2013.

Elenga, H., Peyron, O., Bonnefille, R., Jolly, D., Cheddadi, R., Guiot, J., Andrieu, V., Bottema, S., Buchet, G., De Beaulieu, J, L., Hamilton, A. C., Maley, J., Marchant, R., Perez-Obiol, R., Reille, M., Riollet, G., Scott, L., Straka, H., Taylor, D., Van Campo, E., Vincens, A., Laarif, F., and Jonson, H.: Pollen-based biome reconstruction for southern Europe and Africa 18,000 yr bp, J. Biogeogr., 27, 621–634, https://doi.org/10.1046/j.1365-2699.2000.00430.x, 2000.

Flantua, S. G. A., Hooghiemstra, H., Grimm, E. C., Behling, H., Bush, M. B., González-Arango, C., Gosling, W. D., Ledru, M, P., Lozano-García, S., Maldonado, A., Prieto, A. R., Rull, V., and Van Boxel, J. H.: Updated site compilation of the Latin American Pollen Database, Rev. Palaeobot. Palynol., 223, 104–115, https://doi.org/10.1016/j.revpalbo.2015.09.008, 2015.

Flantua, S. G. A., Mottl, O., Felde, V. A., Bhatta, K. P., Birks, H. H., Grytnes, J, A., Seddon, A. W. R., and Birks, H. J. B.: A guide to the processing and standardization of global palaeoecological data for large-scale syntheses using fossil pollen, Global Ecol. Biogeogr., 32, 1377–1394, https://doi.org/10.1111/geb.13693, 2023.

Fyfe, R. M., de Beaulieu, J, L., Binney, H., Bradshaw, R. H. W., Brewer, S., Le Flao, A., Finsinger, W., Gaillard, M, J., Giesecke, T., Gil-Romera, G., Grimm, E. C., Huntley, B., Kunes, P., Kühl, N., Leydet, M., Lotter, A. F., Tarasov, P. E., and Tonkov, S.: The European Pollen Database: past efforts and current activities, Veg. Hist. Archaeobot., 18, 417–424, https://doi.org/10.1007/s00334-009-0215-9, 2009.

García-Herrera, R. and Barriopedro, D.: Climate of the Mediterranean Region, in: Oxford Research Encyclopedia of Climate Science, https://doi.org/10.1093/acrefore/9780190228620.013.509, 2018.

Giesecke, T., Davis, B., Brewer, S., Finsinger, W., Wolters, S., Blaauw, M., de Beaulieu, J, L., Binney, H., Fyfe, R. M., Gaillard, M, J., Gil-Romera, G., van der Knaap, W. O., Kuneš, P., Kühl, N., van Leeuwen, J. F. N., Leydet, M., Lotter, A. F., Ortu, E., Semmler, M., and Bradshaw, R. H. W.: Towards mapping the late Quaternary vegetation change of Europe, Veg. Hist. Archaeobot., 23, 75–86, https://doi.org/10.1007/s00334-012-0390-y, 2014.

Harrison, S.: BIOME 6000 DB classified plotfile version 1, University of Reading [data set], https://doi.org/10.17864/1947.99, 2017.

Harrison, S. P., Prentice, I. C., Barboni, D., Kohfeld, K. E., Ni, J., and Sutra, J. P.: Ecophysiological and bioclimatic foundations for a global plant functional classification, J. Veg. Sci., 21, 300–317, https://doi.org/10.1111/j.1654-1103.2009.01144.x, 2010.

He, F. and Clark, P. U.: Freshwater forcing of the Atlantic Meridional Overturning Circulation revisited, Nat. Clim. Change, 12, 449–454, https://doi.org/10.1038/s41558-022-01328-2, 2022.

Herzschuh, U., Li, C., Böhmer, T., Postl, A. K., Heim, B., Andreev, A. A., Cao, X., Wieczorek, M., and Ni, J.: LegacyPollen 1.0: a taxonomically harmonized global late Quaternary pollen dataset of 2831 records with standardized chronologies, Earth Syst. Sci. Data, 14, 3213–3227, https://doi.org/10.5194/essd-14-3213-2022, 2022.

Hoogakker, B. a. A., Smith, R. S., Singarayer, J. S., Marchant, R., Prentice, I. C., Allen, J. R. M., Anderson, R. S., Bhagwat, S. A., Behling, H., Borisova, O., Bush, M., Correa-Metrio, A., de Vernal, A., Finch, J. M., Fréchette, B., Lozano-Garcia, S., Gosling, W. D., Granoszewski, W., Grimm, E. C., Grüger, E., Hanselman, J., Harrison, S. P., Hill, T. R., Huntley, B., Jiménez-Moreno, G., Kershaw, P., Ledru, M. P., Magri, D., McKenzie, M., Müller, U., Nakagawa, T., Novenko, E., Penny, D., Sadori, L., Scott, L., Stevenson, J., Valdes, P. J., Vandergoes, M., Velichko, A., Whitlock, C., and Tzedakis, C.: Terrestrial biosphere changes over the last 120 kyr, Clim. Past, 12, 51–73, https://doi.org/10.5194/cp-12-51-2016, 2016.

Joos, F. and Spahni, R.: Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years, Proc. Natl. Acad. Sci. U.S.A., 105, 1425–1430, https://doi.org/10.1073/pnas.0707386105, 2008.

625 Kapsch, M, L., Mikolajewicz, U., Ziemen, F., and Schannwell, C.: Ocean response in transient simulations of the last deglaciation dominated by underlying ice-sheet reconstruction and method of meltwater distribution, Geophys. Res. Lett., 49, e2021GL096767, https://doi.org/10.1029/2021GL096767, 2022.

Kleinen, T., Gromov, S., Steil, B., and Brovkin, V.: PalMod2 MPI-M MPI-ESM1-2-CR-CH4 transient-deglaciation-prescribed-glac1d-methane, World Data Center for Climate (WDCC) at DKRZ [data set], 630 https://doi.org/10.26050/WDCC/PMMXMCHTD, 2023.

Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H.: A 156 kyr smoothed history of the atmospheric greenhouse gases CO_2, CH_4, and N_2O and their radiative forcing, Earth Syst. Sci. Data, 9, 363–387, https://doi.org/10.5194/essd-9-363-2017, 2017.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., and Levrard, B.: A long-term numerical 635 solution for the insolation quantities of the Earth, Astron. Astrophys., 428, 261–285, https://doi.org/10.1051/0004-6361:20041335, 2004.

Lézine, A. M., Watrin, J., Vincens, A., and Hély, C.: Are modern pollen data representative of west African vegetation?, Rev. Palaeobot. Palynol., 156, 265–276, https://doi.org/10.1016/j.revpalbo.2009.02.001, 2009.

Lézine, A. M., Ivory, S. J., Gosling, W. D., and Scott, L.: The African Pollen Database (APD) and tracing 640 environmental change: State of the Art, in: Quaternary Vegetation Dynamics, edited by: Runge, J., Gosling, W., Lézine, A. M., Scott, L., CRC Press, London, United Kingdom, 8, https://doi.org/10.1201/9781003162766-2, 2021.

Li, C., Postl, A. K., Böhmer, T., Cao, X., Dolman, A. M., and Herzschuh, U.: Harmonized chronologies of a global late Quaternary pollen dataset (LegacyAge 1.0), Earth Syst. Sci. Data, 14, 1331–1343, 645 https://doi.org/10.5194/essd-14-1331-2022, 2022.

Li, C., Ni, J., Böhmer, T., Cao, X., Zhou, B., Liao, M., Li, K., Schild, L., Heim, B., and Herzschuh, U.: LegacyPollen2.0: an updated global taxonomically and temporally standardized fossil pollen dataset of 3728 palynological records, PANGAEA [data set], https://doi.pangaea.de/10.1594/PANGAEA.965907, 2024.

Liu, Z., Otto-Bliesner, B. L., He, F., Brady, E. C., Tomas, R., Clark, P. U., Carlson, A. E., Lynch-Stieglitz, J., 650 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.

Marchant, R., Berrio, J. C., Cleef, A., Duivenvoorden, J., Helmens, K., Hooghiemstra, H., Kuhry, P., Melief, B., Schreve-Brinkman, E., Van Geel, B., Van Reenen, G., and Van der Hammen, T.: A reconstruction of Colombian 655 biomes derived from modern pollen data along an altitude gradient, Rev. Palaeobot. Palynol., 117, 79–92, https://doi.org/10.1016/S0034-6667(01)00078-1, 2001.

Marchant, R., Cleef, A., Harrison, S. P., Hooghiemstra, H., Markgraf, V., van Boxel, J., Ager, T., Almeida, L., Anderson, R., Baied, C., Behling, H., Berrio, J. C., Burbridge, R., Björck, S., Byrne, R., Bush, M., Duivenvoorden, J., Flenley, J., De Oliveira, P., van Geel, B., Graf, K., Gosling, W. D., Harbele, S., van der Hammen, T., Hansen, 660 B., Horn, S., Kuhry, P., Ledru, M, P., Mayle, F., Leyden, B., Lozano-García, S., Melief, A. M., Moreno, P., Moar, N. T., Prieto, A., van Reenen, G., Salgado-Labouriau, M., Schäbitz, F., Schreve-Brinkman, E. J., and Wille, M.: Pollen-based biome reconstructions for Latin America at 0, 6000 and 18 000 radiocarbon years ago, Clim. Past, 5, 725–767, https://doi.org/10.5194/cp-5-725-2009, 2009.

Marinova, E., Harrison, S. P., Bragg, F., Connor, S., de Laet, V., Leroy, S. A. G., Mudie, P., Atanassova, J., 665 Bozilova, E., Caner, H., Cordova, C., Djamali, M., Filipova-Marinova, M., Gerasimenko, N., Jahns, S., Kouli, K., Kotthoff, U., Kvavadze, E., Lazarova, M., Novenko, E., Ramezani, E., Röpke, A., Shumilovskikh, L., Tanţău, I., and Tonkov, S.: Pollen-derived biomes in the Eastern Mediterranean–Black Sea–Caspian-Corridor, J. Biogeogr., 45, 484–499, https://doi.org/10.1111/jbi.13128, 2018.

Masoum, A., Nerger, L., Willeit, M., Ganopolski, A., and Lohmann, G.: Paleoclimate data assimilation with 670 CLIMBER-X: An ensemble Kalman filter for the last deglaciation, PLoS One, 19, e0300138, https://doi.org/10.1371/journal.pone.0300138, 2024.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster,

675   S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S, S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J, S., Tian, F., Voigt, A., Vrese, P., Wieners, K, H., Wilkenskjeld, S.,

680   Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing CO_2, J. Adv. Model. Earth Syst., 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

Mikolajewicz, U., Kapsch, M, L., Gayler, V., Meccia, V. L., Riddick, T., Ziemen, F. A., and Schannwell, C.: PalMod2 MPI-M MPI-ESM1-2-CR Transient Simulations of the Last Deglaciation with prescribed ice sheets

685   from ICE-6G reconstructions (r1i1p2f2), World Data Center for Climate (WDCC) at DKRZ [data set], https://doi.org/10.26050/WDCC/PMMXMCRTDIP122, 2023.

Müller, M.: Dynamic Time Warping, in: Information Retrieval for Music and Motion, edited by: Müller, M., Springer, Berlin, Heidelberg, Germany, 69–84, https://doi.org/10.1007/978-3-540-74048-3_4, 2007.

Ni, J., Cao, X., Jeltsch, F., and Herzschuh, U.: Biome distribution over the last 22,000 yr in China, Palaeogeogr.

690   Palaeoclimatol. Palaeoecol., 409, 33–47, https://doi.org/10.1016/j.palaeo.2014.04.023, 2014.

Nolan, C., Overpeck, J. T., Allen, J. R. M., Anderson, P. M., Betancourt, J. L., Binney, H. A., Brewer, S., Bush, M. B., Chase, B. M., Cheddadi, R., Djamali, M., Dodson, J., Edwards, M. E., Gosling, W. D., Haberle, S., Hotchkiss, S. C., Huntley, B., Ivory, S. J., Kershaw, A. P., Kim, S, H., Latorre, C., Leydet, M., Lézine, A, M., Liu, K, B., Liu, Y., Lozhkin, A. V., McGlone, M. S., Marchant, R. A., Momohara, A., Moreno, P. I., Müller, S., Otto-

695   Bliesner, B. L., Shen, C., Stevenson, J., Takahara, H., Tarasov, P. E., Tipton, J., Vincens, A., Weng, C., Xu, Q., Zheng, Z., and Jackson, S. T.: Past and future global transformation of terrestrial ecosystems under climate change, Science, 361, 920–923, https://doi.org/10.1126/science.aan5360, 2018.

Ortega-Rosas, C. I., Guiot, J., Peñalba, M. C., and Ortiz-Acosta, M. E.: Biomization and quantitative climate reconstruction techniques in northwestern Mexico—With an application to four Holocene pollen sequences,

700   Global Planet. Change, 61, 242–266, https://doi.org/10.1016/j.gloplacha.2007.10.006, 2008.

Peltier, W. R.: Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G (VM2) model and GRACE, Annu. Rev. Earth Planet. Sci., 32, 111–149, https://doi.org/10.1146/annurev.earth.32.082503.144359, 2004.

Peltier, W. R., Argus, D. F., and Drummond, R.: Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model, Geophys. Res. Atmos., 120, 450–487, https://doi.org/10.1002/2014JB011176,

705   2015.

Pickett, E. J., Harrison, S. P., Hope, G., Harle, K., Dodson, J. R., Peter Kershaw, A., Colin Prentice, I., Backhouse, J., Colhoun, E. A., D'Costa, D., Flenley, J., Grindrod, J., Haberle, S., Hassell, C., Kenyon, C., Macphail, M., Martin, H., Martin, A. H., McKenzie, M., Newsome, J. C., Penny, D., Powell, J., Ian Raine, J., Southern, W., Stevenson, J., Sutra, J, P., Thomas, I., van der Kaars, S., and Ward, J.: Pollen-based reconstructions of biome

710   distributions for Australia, Southeast Asia and the Pacific (SEAPAC region) at 0, 6000 and 18,000 ^{14}C yr BP, J. Biogeogr., 31, 1381–1444, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2699.2004.01001.x, 2004.

Prentice, C., Guiot, J., Huntley, B., Jolly, D., and Cheddadi, R.: Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka, Clim. Dyn., 12, 185–194, https://doi.org/10.1007/BF00211617, 1996.

715   Prentice, I. C. and Webb III, T.: BIOME 6000: reconstructing global mid-Holocene vegetation patterns from palaeoecological records, J. Biogeogr., 25, 997–1005, https://doi.org/10.1046/j.1365-2699.1998.00235.x, 1998.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing [code], Vienna, Austria, https://www.r-project.org/ (last access: 15 June, 2024), 2020.

Ramankutty, N. and Foley, J. A.: Estimating historical changes in global land cover: Croplands from 1700 to 1992,

720   Global Biogeochem. Cycles, 13, 997–1027, https://doi.org/10.1029/1999GB900046, 1999.

Ramankutty, N., Foley, J. A., Hall, F. G., Collatz, G. J., Meeson, B. W., Los, S. O., Brown De Colstoun, E., and Landis, D. R.: ISLSCP II Potential Natural Vegetation Cover, ORNL DAAC [data set], https://doi.org/10.3334/ORNLDAAC/961, 2010.

Sánchez Goñi, M. F., Desprat, S., Daniau, A, L., Bassinot, F. C., Polanco-Martínez, J. M., Harrison, S. P., Allen,

725   J. R. M., Anderson, R. S., Behling, H., Bonnefille, R., Burjachs, F., Carrión, J. S., Cheddadi, R., Clark, J. S., Combourieu-Nebout, N., Mustaphi, C. J. C., Debusk, G. H., Dupont, L. M., Finch, J. M., Fletcher, W. J., Giardini,

M., González, C., Gosling, W. D., Grigg, L. D., Grimm, E. C., Hayashi, R., Helmens, K., Heusser, L. E., Hill, T., Hope, G., Huntley, B., Igarashi, Y., Irino, T., Jacobs, B., Jiménez-Moreno, G., Kawai, S., Kershaw, A. P., Kumon, F., Lawson, I. T., Ledru, M, P., Lézine, A, M., Liew, P. M., Magri, D., Marchant, R., Margari, V., Mayle, F. E., McKenzie, G. M., Moss, P., Müller, S., Müller, U. C., Naughton, F., Newnham, R. M., Oba, T., Pérez-Obiol, R., Pini, R., Ravazzi, C., Roucoux, K. H., Rucina, S. M., Scott, L., Takahara, H., Tzedakis, P. C., Urrego, D. H., van Geel, B., Valencia, B. G., Vandergoes, M. J., Vincens, A., Whitlock, C. L., Willard, D. A., and Yamamoto, M.: The ACER pollen and charcoal database: a global resource to document vegetation and fire response to abrupt climate changes during the last glacial period, Earth Syst. Sci. Data, 9, 679–695, https://doi.org/10.5194/essd-9-679-2017, 2017.

Sarda-Espinosa, A.: dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance, R package version 5.5.12 [code], https://github.com/asardaes/dtwclust (last access: 17 March, 2024), 2023.

Sato, H., Kelley, D. I., Mayor, S. J., Martin Calvo, M., Cowling, S. A., and Prentice, I. C.: Dry corridors opened by fire and low CO_2 in Amazonian rainforest during the Last Glacial Maximum, Nat. Geosci., 14, 578–585, https://doi.org/10.1038/s41561-021-00777-2, 2021.

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.

Socorro, D. V. and Goring, S.: neotoma2: Working with the Neotoma Paleoecology Database, R package version 1.0.3 [code], https://github.com/asardaes/dtwclust (last access: 10 June, 2024), 2024.

Song, X., Wang, D, Y., Li, F., and Zeng, X, D.: Evaluating the performance of CMIP6 Earth system models in simulating global vegetation structure and distribution, Adv. Clim. Change Res., 12, 584–595, https://doi.org/10.1016/j.accre.2021.06.008, 2021.

Stephens, L., Fuller, D., Boivin, N., Rick, T., Gauthier, N., Kay, A., Marwick, B., Armstrong, C. G., Barton, C. M., Denham, T., Douglass, K., Driver, J., Janz, L., Roberts, P., Rogers, J. D., Thakar, H., Altaweel, M., Johnson, A. L., Sampietro Vattuone, M. M., Aldenderfer, M., Archila, S., Artioli, G., Bale, M. T., Beach, T., Borrell, F., Braje, T., Buckland, P. I., Jiménez Cano, N. G., Capriles, J. M., Diez Castillo, A., Çilingiroğlu, Ç., Negus Cleary, M., Conolly, J., Coutros, P. R., Covey, R. A., Cremaschi, M., Crowther, A., Der, L., di Lernia, S., Doershuk, J. F., Doolittle, W. E., Edwards, K. J., Erlandson, J. M., Evans, D., Fairbairn, A., Faulkner, P., Feinman, G., Fernandes, R., Fitzpatrick, S. M., Fyfe, R., Garcea, E., Goldstein, S., Goodman, R. C., Dalpoim Guedes, J., Herrmann, J., Hiscock, P., Hommel, P., Horsburgh, K. A., Hritz, C., Ives, J. W., Junno, A., Kahn, J. G., Kaufman, B., Kearns, C., Kidder, T. R., Lanoë, F., Lawrence, D., Lee, G, A., Levin, M. J., Lindskoug, H. B., López-Sáez, J. A., Macrae, S., Marchant, R., Marston, J. M., McClure, S., McCoy, M. D., Miller, A. V., Morrison, M., Motuzaite Matuzeviciute, G., Müller, J., Nayak, A., Noerwidi, S., Peres, T. M., Peterson, C. E., Proctor, L., Randall, A. R., Renette, S., Robbins Schug, G., Ryzewski, K., Saini, R., Scheinsohn, V., Schmidt, P., Sebillaud, P., Seitsonen, O., Simpson, I. A., Sołtysiak, A., Speakman, R. J., Spengler, R. N., Steffen, M. L., et al.: Archaeological assessment reveals Earth's early transformation through land use, Science, 365, 897–902, https://doi.org/10.1126/science.aax1192, 2019.

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., and Satoto, B. D.: Integration k-means clustering method and elbow method for identification of the best customer profile cluster, IOP Conf. Ser. Mater. Sci. Eng., 336, 012017, https://doi.org/10.1088/1757-899X/336/1/012017, 2018.

Tarasov, L., Dyke, A. S., Neal, R. M., and Peltier, W. R.: A data-calibrated distribution of deglacial chronologies for the North American ice complex from glaciological modeling, Earth Planet. Sci. Lett., 315–316, 30–40, https://doi.org/10.1016/j.epsl.2011.09.010, 2012.

Thompson, R. S. and Anderson, K. H.: Biomes of western North America at 18,000, 6000 and 0 ^{14}C yr bp reconstructed from pollen and packrat midden data, J. Biogeogr., 27, 555–584, https://doi.org/10.1046/j.1365-2699.2000.00427.x, 2000.

Tian, F., Cao, X., Dallmeyer, A., Lohmann, G., Zhang, X., Ni, J., Andreev, A., Anderson, P. M., Lozhkin, A. V., Bezrukova, E., Rudaya, N., Xu, Q., and Herzschuh, U.: Biome changes and their inferred climatic drivers in northern and eastern continental Asia at selected times since 40 cal ka bp, Veg. Hist. Archaeobot., 27, 365–379, https://doi.org/10.1007/s00334-017-0653-8, 2018.

van Langevelde, F., van de Vijver, C. A. D. M., Prins, H. H. T., and Groen, T. A.: Effects of grazing and browsing on tropical savanna vegetation, in: The Ecology of Browsing and Grazing II, edited by: Gordon, I. J. and Prins, H. H. T., Springer International Publishing, Cham, Switzerland, 237–257, https://doi.org/10.1007/978-3-030-25865-8_10, 2019.

Vincens, A., Bremond, L., Brewer, S., Buchet, G., and Dussouillez, P.: Modern pollen-based biome reconstructions in East Africa expanded to southern Tanzania, Rev. Palaeobot. Palynol., 140, 187–212, https://doi.org/10.1016/j.revpalbo.2006.04.003, 2006.

Willeit, M. and Ganopolski, A.: PALADYN v1.0, a comprehensive land surface–vegetation–carbon cycle model of intermediate complexity, Geosci. Model Dev., 9, 3817–3857, https://doi.org/10.5194/gmd-9-3817-2016, 2016.

Willeit, M., Ganopolski, A., Robinson, A., and Edwards, N. R.: The Earth system model CLIMBER-X v1.0 – Part 1: Climate model description and validation, Geosci. Model Dev., 15, 5905–5948, https://doi.org/10.5194/gmd-15-5905-2022, 2022.

Willeit, M., Ilyina, T., Liu, B., Heinze, C., Perrette, M., Heinemann, M., Dalmonech, D., Brovkin, V., Munhoven, G., Börker, J., Hartmann, J., Romero-Mujalli, G., and Ganopolski, A.: The Earth system model CLIMBER-X v1.0 – Part 2: The global carbon cycle, Geosci. Model Dev., 16, 3501–3534, https://doi.org/10.5194/gmd-16-3501-2023, 2023.

Williams, J. W., Webb III, T., Richard, P. H., and Newby, P.: Late Quaternary biomes of Canada and the eastern United States, J. Biogeogr., 27, 585–607, https://doi.org/10.1046/j.1365-2699.2000.00428.x, 2000.

Williams, J. W., Grimm, E. C., Blois, J. L., Charles, D. F., Davis, E. B., Goring, S. J., Graham, R. W., Smith, A. J., Anderson, M., Arroyo-Cabrales, J., Ashworth, A. C., Betancourt, J. L., Bills, B. W., Booth, R. K., Buckland, P. I., Curry, B. B., Giesecke, T., Jackson, S. T., Latorre, C., Nichols, J., Purdum, T., Roth, R. E., Stryker, M., and Takahara, H.: The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource, Quat. Res., 89, 156–177, https://doi.org/10.1017/qua.2017.105, 2018.

Xu, Q., Cao, X., Tian, F., Zhang, S., Li, Y., Li, M., Li, J., Liu, Y., and Liang, J.: Relative pollen productivities of typical steppe species in northern China and their potential in past vegetation reconstruction, Sci. China Earth Sci., 57, 1254–1266, https://doi.org/10.1007/s11430-013-4738-7, 2014.

Zhou, B., Liao, M., Li, K., Xu, D., Chen, H., Ni, J., Cao, X., Kong, Z., Xu, Q., Zhang, Y., Herzschuh, U., Cai, Y., Chen, B., Chen, J., Chen, L., Cheng, B., Gao, Y., Huang, C., Huang, X., Li, S., Li, W., Liao, G., Liu, G., Liu, P., Liu, X., Ma, C., Song, C., Sun, X., Tang, L., Wang, M., Wang, Y., Xia, Y., Xu, J., Yan, S., Yang, X., Yao., Y., Ye, C., Zhang, Z., Zhao, Z., Zheng, Z., Zhu, C.: A fossil pollen dataset of China [In Chinese], Chin. J. Plant Ecol., 47, 1453–1463, https://doi.org/10.17521/cjpe.2022.0316, 2023.