Climate reconstructions based on GDGT and pollen surface datasets from Mongolia and Siberia: calibrations and applicability to extremely cold-dry environments over the Late Holocene.

Lucas Dugerdil[1,2], Sébastien Joannin[2], Odile Peyron[2], Isabelle Jouffroy-Bapicot[3], Boris Vannière[3], Boldgiv Bazartseren[4], Julia Unkelbach[5], Hermann Behling[5], and Guillemette Ménot[1]

[1]Univ. Lyon, ENS de Lyon, Université Lyon 1, CNRS, UMR 5276 LGL-TPE, F-69364, Lyon, France
[2]Université de Montpellier, CNRS, IRD, EPHE, UMR 5554 ISEM, Montpellier, France
[3]Université Bourgogne Franche Comté, CNRS UMR 6249 Laboratoire Chrono-environnement, F-25030, Besançon, France
[4]Ecology Group, Department of Biology, School of Arts and Sciences, National University of Mongolia, Ulaanbaatar 14201, Mongolia
[5]Department of Palynology and Climate Dynamics, Albrecht-von-Haller-Institute for Plant Sciences, University of Goettingen, Germany

Correspondence: Lucas Dugerdil, lucas.dugerdil@ens-lyon.fr

Abstract. Our understanding of climate and vegetation changes throughout the Holocene is hampered by representativeness in sedimentary archives. Potential biases such as production and preservation of the markers are identified by comparing these proxies with modern environments. It is important to conduct multi-proxy studies and robust calibrations on each terrestrial biome. These calibrations use large data base dominated by forest samples. Therefore, including data from taiga and steppe

5　sites becomes mandatory to better calibrate arid environments. The Mongolian plateau, ranging from the Baikal basin to the Gobi desert, is especially characterized by low annual precipitation and continental annual air temperature. The characterization of the climate system of this area is crucial for the understanding of Holocene Monsoon Oscillations. This study focuses on the calibration of proxy-climate relationships for pollen and glycerol dialkyl glycerol tetraethers (GDGTs) by comparing large published Eurasian calibrations with a set of 49 new surface samples (moss polster, soil and mud from temporary dry pond).

10　These calibrations are then cross-validated by an independent dataset of top-core samples and applied to two Late Holocene paleosequences in the Altai mountains and the Qaidam basin. We show that: (1) preserved pollen assemblages are clearly imprinted on the extremities of the ecosystem range but mitigated and unclear on the ecotones; (2) for both proxies, inferred relationships depend on the geographical range covered by the calibration database as well as on the nature of samples; (3) even if local calibrations suffer from reduced amplitude of climatic parameter due to local homogeneity, they better reflect

15　actual climate than the global ones by reducing the limits for saturation impact, (4) a bias in climatic reconstructions is induced by the over-parameterization of the models by addition of artificial correlation and (5) paleoclimate values reconstructed here are consistent with Mongolia-China Late Holocene climate trends, and validate the application of local calibrations for both pollen and GDGTs (closest fit to actual values and realistic paleoclimate amplitude). We encourage the application of this surface calibration method to reconstruct palaeoclimate and especially consolidate our understanding of the Holocene climate

20　and environment variations in Arid Central Asia.

1 Introduction

Since the understanding of the interactions between the palaeoclimate proxies, such as pollen or biomarker abundances, and General Circulation Model outputs became a major issue in future climate change modelling, resolving the issue of climate
25 proxy calibration is crucial (Braconnot et al., 2012). Current climate changes in extremely cold environments (Masson-Delmotte, 2018), such as Mongolia and Siberia (Fig. 1), are amplified compared with other places around the world (Tian et al., 2014) and the drivers of the current degradation of Mongolian environments in diversity and biomass production still need to be understood. From a climatic point of view, Mongolia is on a junction between the westerlies which are driven by the North Atlantic Oscillation (NAO) and the East Asian Summer Monsoon which is linked to the El Niño-Southern Oscillations
30 (ENSO) and the Inter-tropical Convergence Zone (ITCZ, An et al., 2008). The Mongolian plateau is a hinge area: the high altitude of the Altai range to the west and the Sayan range to the north-west of the country partially block both the westerlies arriving from the northern Atlantic ocean through the Siberian basin and the East Asian Summer Monsoon (EASM, Fig. 2, Chen et al., 2009). The Mongolian system is thus driven by a mix of the distant drag of these two main climatic cells. The understanding of the complex interaction of these cells is necessary and continuously improved by palaeoclimatic studies (**?**).

35

Lake sediment archives are commonly used to infer past variations of these climate and environmental systems associated with vegetation and human land use (Lehmkuhl et al., 2011; Felauer et al., 2012; Wang and Feng, 2013). Among the proxies available, pollen and geochemical biomarkers are used as past temperature indicators (Ter Braak and Juggins, 1993; Weijers et al., 2007b) and the combination of these proxies helps to polish lake sediment shift interpretations (Atahan et al., 2015;
40 Watson et al., 2018; Martin et al., 2020; Kaufman et al., 2020). From decades the pollen signal is used to display shifts in vegetation composition and structure (Bennett and Willis, 2002) and allows quantitative reconstructions of climate parameters such as precipitation regime and temperatures (Birks et al., 2010; Ohlwein and Wahl, 2012; Wen et al., 2013; Cao et al., 2014; Marsicek et al., 2018). Since vegetation structure and pollen production are mainly influenced by climatic parameters (Zheng et al., 2008) in the absence of human influences, the paleo-pollen signal is very often interpreted as a response to the climate variations
45 through time (Kröpelin et al., 2008; Wagner et al., 2019). Even if human activities influence pollen rain as well (Hjelle, 1997; Hellman et al., 2009a), these empirical observations of the pollen-climate relation leads to the development of semi-quantitative (Ma et al., 2008) and quantitative calibrations (Brewer et al., 2008; Salonen et al., 2019) of the signal. Different methods have been developed to reconstruct past climates: Probability Density Functions, Assemblages approaches, Transfer Functions (TF) and methods based on vegetation models (Guiot et al., 2000; Birks et al., 2010; Bartlein et al., 2011; Ohlwein and Wahl, 2012).
50 More precisely, these methods are: the Inverse Modelling Method (IMM, Guiot et al., 2000), the Weighted Averaging Partial Least Squares regression (WAPLS, Ter Braak and Juggins, 1993; Ter Braak et al., 1993), the Artificial Neural Networks (ANN, Peyron et al., 1998), the Modern Analogue Technique (MAT, Overpeck et al., 1985; Guiot, 1990; Jackson and Williams, 2004), the Response Surface Technique (RST, Bartlein et al., 1986), Probability Density Functions (PDF, Kühl et al., 2002; Chevalier,

2019), Modified Mutual Climate Range Method (MMCRM, Klotz et al., 2003, 2004), Bayesian Hierarchical Models (BHM,
55 Ohlwein and Wahl, 2012), the Boosted Regression Trees (BRT, Salonen et al., 2014), etc. For northern Europe and despite some
problems and pitfalls, Seppä et al. (2004) demonstrated that pollen-inferred climate reconstructions are generally consistent
and agree well with other independent climatic reconstructions. This study encourages us to lead multi-proxy studies to refine
climate reconstruction understanding, and especially in tricky and dry context such as Mongolian plateau (Rudaya et al., 2009).

60   Among new promising proxies and from the three last decades, biomarkers such as the glycerol dialkyl glycerol tetraethers
(GDGTs) have provided new perspectives on continental temperature reconstructions (Naafs et al., 2017a, b). Among the
GDGTs, we will focus on two major groups: the isoprenoid-GDGTs (isoGDGTs, Hopmans et al., 2000) and the branched-
GDGTs (brGDGTs, Damsté et al., 2000; Weijers et al., 2007a, b; Dearing Crampton-Flood et al., 2020). BrGDGT assemblages
reflect bacterial activity in rivers (De Jonge et al., 2014b), soil (De Jonge et al., 2014a) or lake water column (Dang et al., 2018).
65 Environmental drivers are linked to climate parameters (Weijers et al., 2007b), soil typology and vegetation cover (Davtian
et al., 2016), which in turn imply land cover and land use. Accurate determinations of the relationships between brGDGT
assemblages and climate still need some improvements (Naafs et al., 2018; Wang et al., 2019, 2020) and especially on local
to regional scales and in extreme environments. It has been shown empirically (Weijers et al., 2004; Huguet et al., 2013) on
cultures of pure strains (Salvador-Castel et al., 2019 in press) as well as on meso- and micro-cosm experiments (Chen et al.,
70 2018; Martínez-Sosa et al., 2020), that organisms adjust their membrane plasticity by the degree of methylation and cyclisation
of the compounds. Moreover, some studies have focused on variations in the bacterial community structure (Xie et al., 2015),
the bacterial group responses to environmental changes (Knappy et al., 2011) and the GDGT occurrences in different bacte-
rial communities (Liu et al., 2012b) to determine the potential effects of community structure on GDGT relative abundances.
To evaluate the provenance and the climatic information brGDGTs bear, several indexes have been proposed in the literature
75 (Supplementary Table S1). To monitor these changes, CBT and MBT indexes linked to environmental factors such as climate
and soil parameters have been proposed (Weijers et al., 2007b; Huguet et al., 2013). In particular, the methylation degree, ratio
of 5, 6 (De Jonge et al., 2013) and 7-methyl isomers (Ding et al., 2016) reacts to environment forcing : the 5-methyl correlates
mainly with temperature (Naafs et al., 2017a), while 6 and 7-methyl seem to react to moisture and pH (Yang et al., 2015; Ding
et al., 2016). More specific indexes have been proposed by De Jonge et al. (2014a) to limit the multi-correlation systems such
80 as MBT'_{5Me} which is independent of the pH and CBT_{5Me} which is more representative of the soil pH than the former version
of the index. The statistical relevance of these indexes is a major issue in brGDGT calibration (Dearing Crampton-Flood et al.,
2020). Some regional indexes for soil temperature such as Index_1 (De Jonge et al., 2014a) and Index_2 for Chinese soils (Wang
et al., 2016) have been explored too, in a context of a strong local calibration demand (Ding et al., 2015; Yang et al., 2015).
For moisture variations, the $\mathrm{R}_{i/b}$ index has been proposed as a reliable aridity proxy (Yang et al., 2014; Xie et al., 2012). It has
85 been shown that a linear relation exists between these GDGT indexes and some climatic features at large regional scales (in
the wide Chinese biome range, from tropical forest to central arid plateau, for instance, Yang et al., 2014; Lei et al., 2016).

Since multi-proxy studies become more and more accurate in both temperature and precipitation reconstruction, local to regional calibrations have been proposed for dry areas such as the Arid Central Asian (ACA) area: pollen semi-quantitative

90   climate reconstruction (Ma et al., 2008), pollen transfer functions (Herzschuh et al., 2003, 2004; Cao et al., 2014; Zheng et al., 2014), and brGDGT regression models (Sun et al., 2011; Yang et al., 2014; Ding et al., 2015; Wang et al., 2016; Thomas et al., 2017). Even if all of these studies focus on areas surrounding the EASM line (Fig. 2, Chen et al., 2010; Li et al., 2018), the understanding of the climate cells interaction and oscillation over time is still lacunary, and especially on the ACA upper edge. In this context, our study took place on the northernmost part of this climatic system (Haoran and Weihong,

95   2007). Moreover, we propose the first multi-proxy calibration exercise in ACA area based on pollen and brGDGT fractional abundances retrieved from modern samples (soil, moss litter, pond mud) in semi-arid to temperate conditions. The aim of this study is to take advantage of new, modern surface sample datasets in Siberia and Mongolia to propose an adapted calibration of pollen and bacterial biomarker proxies for cold and dry environments. For that purpose, local calibrations are compared with global calibrations to infer the influence of calibration scale and proxy types on derived climatic parameters. Our approach is

100   summarized in the following steps:

1. Collection of a new set of modern surface samples for Mongolia with homogeneous characterisation of their bioclimate environment followed by pollen and GDGT pattern characterization.

2. Evaluation of the match between actual bioclimate environments and the associated pollen rain and biomarker assemblages based on mathematical criterion without eco-physiological considerations.

105   3. Creation of local Siberia-Mongolia climate calibrations for pollen and GDGTs and comparison of local and global calibrations on the Mongolian study case.

4. Posteriori validation of the inferred relationships between proxies and ecological likelihood based on the currently developed evidences of brGDGT and pollen rain ecological significance.

5. Discussion of the implications of the calibration mismatches in terms of climatic reconstructions in arid and cold envi-

110   ronments.

6. Testing of the new calibrations (pollen and brGDGTs) through their application on two surrounding Late Holocene paleosequences.

2   Mongolian and Siberian Study Area

2.1   Coring, Sampling Area and Sample Types

115   The study area lies from $52°29'$N to $43°34'$N in latitude and from $101°00'$E to $107°06'$E in longitude (Fig. 1.A). The sample sites (n = 49) are listed in the Supplementary Tables S2 with a description of the sample type, the applied analyses, the coordinates and the associated ecosystem. For each site, the *Garmin eTreX10* was used to record GPS coordinates to five-meter
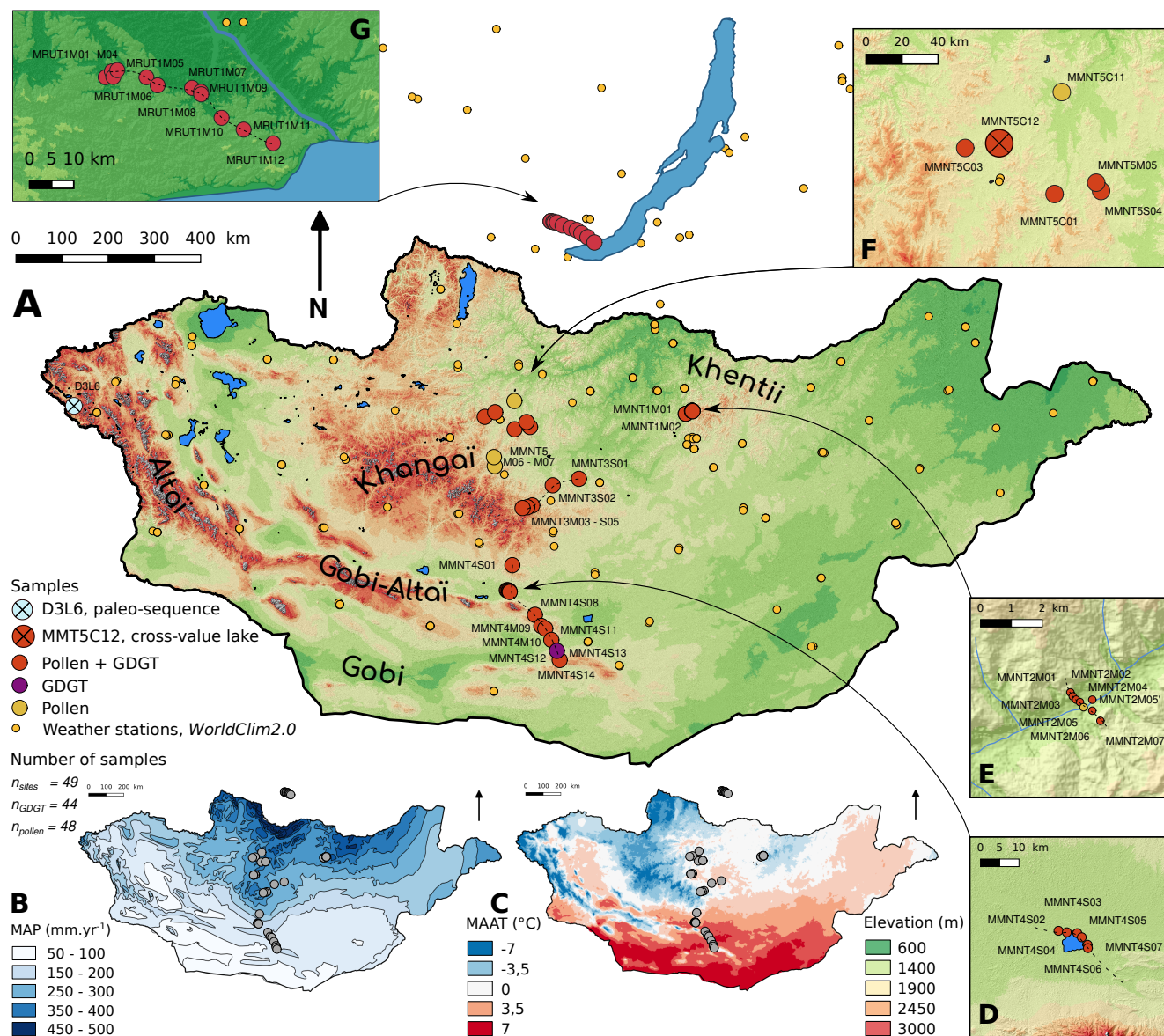
Figure 1. **A**: Topographic map of Mongolia (from ASTER data) with the location of surface samples and weather stations considered in the present study; **B**: Mean Annual Precipitations; **C**: Mean Annual Air Temperatures; **D**: Focus on the samples surrounding the Taatsiin Tsagaan Lake, Gobi desert; **E**: Focus on the samples along a valley in the Khentii range; **F**: Localisation of Khangai surface samples; **G**: Focus on the Baikal Lake transect. The Mongolian GIS Data is issued from a dataset ASTER (https://biosurvey.ku.edu/directory/nicholas-kotlinski), the meteorological dataset from WorldClim2 and infrastructures from public dataset (ALAGaC) (https://marine.rutgers.edu/ cfree/gis-data/mongolia-gis-data/)

accuracy. The surface samples were collected throughout Mongolia in 2016 following four transects *(n = 29)*: in the Khentii

mountain range (*MMNT1* and *MMNT2*, Fig. 1.E), in the Orkhon valley (*MMNT3*), in the Gobi desert and the Gobi-Altai range

120 (*MMNT4*, Fig. 1.D). During the same field trip, a fifth transect has been done in the Sayan range from the Siberian Oblast of

Irkoustsk, Russia (*MRUT1, n = 12*, on Fig. 1.G). A Khangai mountains field trip from spring 2009 enlarge this set of data

with a sixth transect of surface samples (*MMNT5, n = 6*, on Fig. 1.F) and two lake coring from Ayrag Nuur (*MMNT5C12*) and

Shargyl Nuur (*MMNT5C11*) both on Fig. 1.F. Both of these top-cores were added to the surface pollen database, while only

the *MMNT5C12* core has been used as cross-value to check the accuracy of the brGDGTs climate models. This core has been

125 clipped into 62 samples of which the top-core has been replicated 6 times (samples *MMNT5C12-1 to MMNT5C12-6)*. Into the

MMNT5 transect, mud from two temporary dry ponds has been sampled. These surface muds are referred in following figures

as *mud*. Into the other transects and depending on aridity and vegetation at each site, a soil or a moss polster was sampled.

In figures, *soil* refers to the 3 to 5 first centimetres of the ground in dry ecosystems, while *moss* is a mix between soil, litter

and a bryophyte (or cyperaceae) layer in wetter environments. Moss acts as a pollen trap recording a three to five-year mean

130 pollen signal (Räsänen et al., 2004). In drier areas, the soil surface samples have the same function, in spite of a lower pollen

conservation and over-representation of some *taxa* (Lebreton et al., 2010). In parallel for the GDGT analysis and following

the calibration approaches presented in De Jonge et al. (2014a), Davtian et al. (2016) and Naafs et al. (2017a, 2018) mud from

temporary ponds, soil samples as well as soil part of moss litter were also used for actual GDGT analysis. To summarize, this

study is based on 49 sites, 48 samples in the pollen dataset, 44 in the brGDGT dataset and 6 cross-validation samples to test

135 the brGDGT models. In terms of sample types, the dataset consists of 30 mosses, 15 soils, 2 pond muds and 2 top-cores.

To test the reliability of our modern calibrations, we have finally selected two paleosequences close to the Mongolian plateau

used as test-benches of the calibrations. For the pollen analysis, the core D3L6 from Unkelbach et al. (2019) and located in

the Mongolian Altai range is compared to the Xiangride section (XRD) used for brGDGT sequence from Sun et al. (2019),

140 sampled in the Chinese Qaidam Basin (Fig. 2). These two cores have recorded the paleoenvironmental changes of the Late

Holocene period.

2.2 Vegetation and Biomes

The central part of the Mongolian plateau is characterized by a dry and cold flat desert with a 1220 m.a.s.l. median elevation

(Fig. 1.A, Wesche et al., 2016) and is intersected in its northern part by the Khentii range and in its southern part by the

145 Gobi-Altai range aligned along a NW-SE direction. A wet and cold highland in the Khangai ranges culminates at 4000 m.a.s.l

and a flatter and wetter Mongolian area, the Darkhad basin, is located in the north, close to the Russian border on the edge of

the Siberian Sayan range. In the northernmost area, the geography is characterized by the Baikal lake basin at a lower altitude

(around 600 m a.s.l, Fig. 1.G, Demske et al., 2005).

150 The distribution of vegetation and biomes follows a latitudinal belt organization: in the North the boreal forest presents

a mosaic of light-taiga dominated by *Pinus sylvestris* mixed with birches (Demske et al., 2005). On the Mongolian plateau,

the dark-taiga dominated by larches (*Larix sibirica*) and Siberian pines (*Pinus sibirica*) also presents some spruces and fewer birches (*Betula* spp.). The Mongolian taiga is constrained to a region spanning from the Darkhad Basin to the Khentii range (Fig. 1.A). On the north face of the Khangai piedmont, the vegetation is dominated by a mosaic of forest-steppe ecosystems:

155 the steppe is dominated by the *Artemisia* spp. associated with Poaceae, Amaranthaceae, Liliaceae, Fabaceae and Apiaceae. On these open-lands there are some patches of taiga forest, following roughly the broadside and the northern face of the crest letting on to the grasslands in the valley. The two last vegetation layers in Mongolia through the elevation gradient is an alpine meadow dominated by Cyperaceae and Poaceae with a huge floristic biodiversity and an alpine shrubland with pioneer vegetation on the summits. On the southern slope of the range, the ecotone between the steppe and the desert vegetation extends

160 hundreds of kilometers from the northern part of the Gobi desert (with water supplied by the Gobi lake area in between) to the Gobi-Altai range in the south (Klinge and Sauer, 2019). In the southernmost part of the country, the warm and dry climate conditions favour desert vegetation dominated by Amaranthaceae, Nitrariaceae and Zygophyllaceae. The vegetation cover is lower than 25% and is mainly composed of short herbs, succulent plants and a few crawling shrubs.

2.3 Bioclimate Systems

In the central steppe-forest biome, the vegetation is marked by an ecotone with short grassland controlled by grazing in the valley and larches on the slopes. The forest is gathered in patches constituting between 10% and 20% of the total vegetation cover. There are also some patches of *Salix* and *Betula* riparian forests among the sub-alpine meadows on the upper part of the range. This vegetation is characteristic of the northern border of the Palaearctic steppe biome (Wesche et al., 2016). This

170 biome is characterised by a range of 800 to 1600 m a.s.l, a Mean Annual Air Temperature (MAAT, Fig. 1.C) between -2 and 2°C and a Mean Annual Precipitations (MAP, Fig. 1.B) from 180 to 400 $\mathrm{mm.yr}^{-1}$ (Wesche et al. (2016) based on Hijmans et al., 2005). In Mongolia, even if the MAP are very low ($\mathrm{MAP}_{\mathrm{Mongolia}} \in [50; 500]\mathrm{mm.yr}^{-1}$), the major part of the water available for plants is delivered during late spring and early summer, in contrast to Mediterranean and European steppes (Bone et al., 2015; Wesche et al., 2016). These seasons are the optimal plant growth periods. Mongolian summer precipitations are

175 controlled by the East Asian Summer Monsoon system (EASM) instead of the Westerlies' winter precipitation stocked onto the Sayan and Altai range (Fig. 2; An et al., 2008).

3 Methods

3.1 Pollen Analysis, Modern Pollen Datasets and Transfer Functions

180 Different chemical processes were performed on the samples: bryophytic part of the moss samples were deflocculated by KOH and filtered by $250\mu m$ and $10\mu m$ sieves to eliminate the vegetation pieces and the clay particles. Then, acetolysis was performed to destroy biological cells and highlight the pollen grains. For the soil and pond mud samples, 2 steps of HCl and

HF acid attacks were added to the previous protocol to remove all the carbonate and silicate components. All the residuals were finally concentrated in glycerol and mounted between slide and lamella. The pollen counts were carried out with a *Leica*
185 *DM1000 LED* microscope on a 40× magnification lens. The total pollen count size was determined by the asymptotic behaviour of the rarefaction curve. This diagram was plotted during the pollen count using *PolSais 2.0*, software developed in *Python 2.7* for this study. The rarefaction curve was fitted with a logarithmic regression analysis. The counter was suspended whenever the regression curve reached a flatter step (Birks et al., 1992). A threshold for the local derivation at $dx/dy = 0.05$ was set. For each sample, the total pollen count is usually around $n \in [350; 500]$ grains for steppe or forest and $n \in [250; 300]$
190 for drier environment such as desert and desert-steppe.

Among all of the pollen-inferred climate methods, the MAT and the WAPLS were applied in this study on 4 different modern pollen datasets, and on the D3L6 fossil pollen sequence to test the accuracy of these calibrations (Unkelbach et al., 2019, Figs. 1.A and 2). The MAT consists of the selection of a limited number of analogue surface pollen assemblages with their associated
195 climatic values (Guiot, 1990); while the WAPLS uses a Weighted Average correlation method on a limited number of Principal Components connecting the surface pollen fractional abundance to the climate parameters associated (Ter Braak and Juggins, 1993; Ter Braak et al., 1993). The first dataset, called New Mongolian-Siberian Database (NMSDB), is composed of pollen surface samples analysed in this study ($N = 49$, Fig. 2). The second one is the same subset aggregated to the larger Eurasian Pollen Dataset (EAPDB) compiled by Peyron et al. (2013, 2017). From this dataset of 3191 pollen sample sites, a pollen–
200 *Plant Functional Type* method was applied to determine the biome for each sample according to the actual pollen rain (Fig. 2, Prentice et al., 1996; Peyron et al., 1998). Then, only the Cold Steppe (COST) dominant samples were extracted from the main dataset and aggregated with the NMSDB to produce the COSTDB ($N = 430$ sites, figured by lozenge in Fig. 2). Finally, a scale-intermediate dataset of samples located within the Mongolian border merged with the Mongolian New dataset is presented as MDB ($N = 151$ sites). The relation between each taxon and climate parameter was checked and then the MAT
205 and WAPLS methods were applied with the *Rioja* package from the R environment (Juggins and Juggins, 2019).

3.2 SIG Bioclimatic Data

Because Mongolia and Siberia have relatively few weather stations (Fig. 1.A), climate parameters were extracted with R from the interpolated climatic database *WorldClim2* (Fick and Hijmans, 2017). We used Mean Annual Precipitation (MAP, Fig. 1.B), Mean Annual Air Temperature (MAAT, Fig. 1.C), as well as temperatures and precipitations for spring, summer and winter
210 (T_{spr}, P_{spr}, T_{sum}, P_{sum}, T_{win} and P_{win}), Mean Temperature of the Coldest Month (MTCO) and the Mean Temperature of the Warmest Month (MTWA) in this study to characterize the actual climate. Because the Mongolian plateau is not rich in weather stations, the *WorldClim2* database suffers of interpolation errors. The surface sites presenting inconsistent climate parameters (MAP < 0 or MAP < Season Precipitation) were removed from the global database. The elevation data and the topographic map originate from the *ASTER* imagery (Fig. 1.A). The biome type for each site derives from the *LandCover* database (Olson
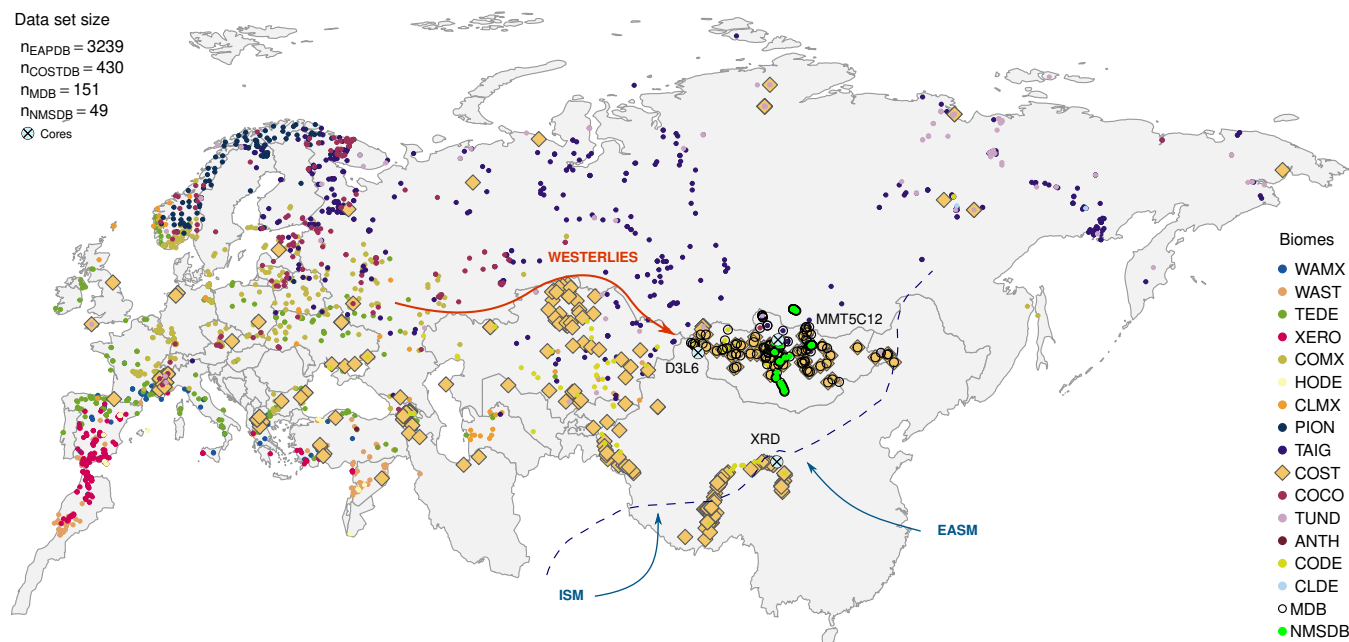215 et al., 2001), classification and field trip observations.

Figure 2. Eurasian map of all the pollen surface samples included in the database. The color code refers to the biome pollen inferred for each site. The biomes are **WAMX**, Warm Mixed Forest; **WAST** Warm Steppe; **TEDE**, Temperate Deciduous Forest; **XERO**, Xerophytic Shrubland; **COMX**, Cool Mixed Forest; **HODE**, Hot Desert; **CLMX** Cold Mixed Forest; **PION** Pioneer forest; **TAIG**, Taiga forest; **COST**, Cold Steppe; **COCO**, Cold Conifer Forest; **TUND**, Tundra; **ANTH**, Anthropic environment; **CLDE**, Cold Deciduous Forest; **CODE**, Cold Desert. The thickest points underline the COST samples selected for this study to operate the transfer function method among the whole COST sites (figured by lozenge on map). The arrows indicate the main climatic system driving the Mongolian climate: in orange the Westerlies arriving from the North Atlantic ocean and in blue the Indian Summer Monsoon (ISM) and the East Asian Summer Monsoon (EASM). The dashed line represents the EASM limit following Chen et al. (2010), (Li et al., 2018) and Haoran and Weihong (2007) for the northernmost part of the boundary. The light blue crossed circles localize the 3 cores used as test-benches for the calibrations (top-core and paleo-sequences).

3.3 GDGT Analysis

For consistency with the sampling process and the modelling methodologies developed for pollen analysis; soil part of the moss polsters, soil samples and pond mud were treated for GDGT analysis. After freeze drying, about 0.6 grams of material were sub-sampled. The Total Lipid Extract (TLE) was microwave extracted (*MARS 6 CEM*) with dichloromethane (DCM):MeOH

220 (3:1) and filtered on empty SPE cartridges. The extraction step was processed twice. Following Huguet et al. (2006), C_{46} GDGT (GDGT with two glycerol head groups linked by C_{20} alkyl chain and two C_{10} alkyl chains) was added as internal standard for GDGT quantification. Then, apolar and polar fractions were separated on an alumina SPE cartridge using hexane: DCM (1:1) and DCM/MeOH (1:1), respectively. Analyses were performed in hexane: iso-propanol (99.8:0.2) by High Performance Liquid Chromatography Mass Spectrometery (HPLC-APCI-MS, Agilent 1200) proceeded in LGLTPE-ENS de Lyon,

225 Lyon, following Hopmans et al. (2016) and Davtian et al. (2018).

Each compound was identified and manually integrated according to its m/z and relative retention time (Liu et al., 2012a; De Jonge et al., 2014a). Statistical treatments on isoGDGT (Fig. 4.A) and brGDGT (Fig. 4.B) abundances were treated following two methods presented in Deng et al. (2016), Wang et al. (2016) and Yang et al. (2019): compounds were gathered by

230 chemical structures as cycles (CBT) or methyl groups (MBT, De Jonge et al., 2014a). brGDGTs were expressed as fractional abundance $[x_i]$ (Fig. 4.B, Sinninghe Damsté, 2016), as follows:

$$f[x]_i = \frac{n_i}{\sum\limits_{j=1}^{N_{brGDGT}} x_j} \tag{1}$$

To infer temperatures from brGDGT abundances, two types of model were applied: linear relationships between temperature and MBT–CBT indexes, and Multiple Regression (mr) models between one climate parameter and a proportion of multiple

235 brGDGT fractional abundances. For the simple linear regression model, a correlation matrix between climate parameters and indexes was calculated using the *corrplot* Rcran library. For mr models, we developed in the R environment a *Stepwise Selection Model* (SSM, Yang et al., 2014) to determine the best fitting model connecting climate parameters with brGDGTs fractional abundances. Then we gathered some of the climate–GDGT linear relations established in previous studies (De Jonge et al., 2014a; Naafs et al., 2017a, b, 2018; Sinninghe Damsté, 2016; Yang et al., 2014, 2019) focusing on a single climatic

240 parameter, MAAT (Supplementary Table S1). These models were clustered into 3 categories, by sample types (mosses, soils or pond muds), geographical area (regional or worldwide scale) and the statistical model (MBT-CBT based of multiple regression models). According to the type of environment from which the samples originated, there was peat, soil and lake-inferred modelling. All these models were applied to the Siberian–Mongolian surface samples, compared with the actual MAAT value at each site and applied to the brGDGT XRD section (Fig. 2, Sun et al., 2019).

245 ### 3.4 Statistical Analyses

GDGTs and pollen data were analyzed with a Principal Components Analysis (PCA) to determine the axes explaining the variance within the samples. The biotic values (pollen and GDGTs) were also compared to abiotic parameters (climate, elevation, location and soil features) by the way of a Redundancy Analysis (RDA). The regression models were run with the $p-value < 0.05$ (model relevance), the R^2 (correlation level between the variables), the Root Mean Square Deviation (RMSE,

250 error on parameter reconstruction) and Akaike's information criterion (AIC, effect of over-parameterization on multiple regression models ; Arnold, 2010; Symonds and Moussalli, 2011). A cross-validation test was performed for all the brGDGT calibrations (from this study and from the literature) using an independent set of six lacustrine samples from the lake *MMNT5C12*. Statistical analyses were performed with the *Rcran* project, using the *ade4* package (Dray and Dufour, 2007) for multivariate analysis. All the plots were made with the *ggplot2* package (Wickham, 2016) or the *Rioja* package (Juggins and Juggins, 2019)

255 for the stratigraphic plot and the pollen clustering using the CONISS analysis method (Grimm, 1987).
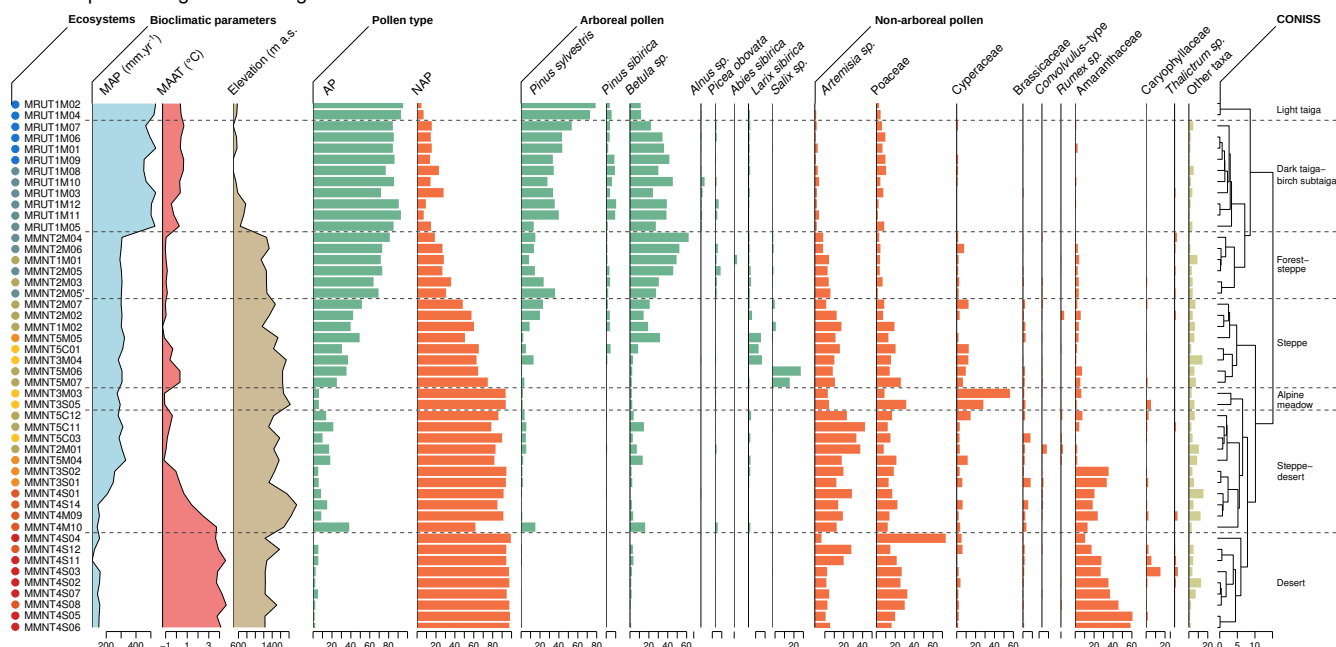
Figure 3. Simplified surface pollen diagram, bio-climatically sorted, of the Siberian–Mongolian transect. The pollen taxa are expressed in %TP. The Ecosystem Units were determined with a CONISS analysis. The left hand coloured dots represent the ecosystem for each sample from dark taiga (deep blue), light taiga, steppe-forest, alpine meadow, steppe, steppe-desert and desert (deep red). The color scale is presented in Fig. 5. The MAP and MAAT are extracted from Fick and Hijmans (2017).

4 Results

4.1 Pollen, Climate and Ecosystems Relations

4.1.1 Modern Pollen Rain and Vegetation Representation

The pollen rain (Fig. 3) is dominated by six main pollen taxa: *Pinus sylvestris*, *Betula* spp., *Artemisia* spp., Poaceae, Cyper-
260 aceae and Amaranthaceae. The pollen diagram, sorted by bio-climate from the wet and relatively warm Siberian basin on the
upper part to the dry-warm Gobi desert on the bottom, presents a net AP decrease from 85% to 5%. 34.26 % of the variance
is explained by PC_1 extending from positive values associated with NAP (Amaranthaceae, Poaceae and *Artemisia* spp.) to
negative values associated with AP (*Pinus* undet., *Betula* spp. *Picea obovata* and *Larix sibirica*, on Fig. 5.C). This trend shows
the transition between ecosystems, marked by the seven main *CONISS* clusters (Fig. 3) and PC_1 and PC_2 variations (Fig. 5.C).
265 Below are the over-representative main *taxa* for each of the Siberian–Mongolian transect ecosystem:

1. **Light taiga** dominated by *Pinus sylvestris* ($> 70\%$), *Pinus sibirica* and very low NAP ($< 5\%$).

11

2. **Mixed dark taiga–birches sub-taiga** with an assemblage of *Larix sibirica*, *Picea obovata*, *Pinus sylvestris* and *P. sibirica*.

3. **Forest-steppe** ecotone same AP assemblages that the dark taiga ecosystem with 20% of *Artemisia* spp., plus occurence of Poaceae, Cyperaceae, *Thalictrum* spp. and *Convolvulus* spp.

4. **Steppe** still dominated by *Artemisia* spp. (30%) and rising Poaceae (25%), Brassicaceae

5. **Alpine meadow** overpowered by Cyperaceae up to 50 %, Poaceae, Brassicaceae, Amaranthaceae and *Convolvulus* spp.

6. **Steppe-desert** ecotone highlighting by the transition between Amaranthaceae–Caryophyllaceae community and Poaceae–*Artemisia* spp. assemblages.

7. **Desert** dominated by Amaranthaceae (from 25 % to 65 %) and by rare pollen-type Caryophyllaceae, *Thalictrum* spp., *Nitraria* spp. and *Tribulus* spp.

4.1.2 Pollen – Climate Interaction

The pollen rain trends follow similar variations than bio-climate parameters in MAP, MAAT and elevation (Fig. 3). Highest AP values are correlated to large MAP (up to 500 $\mathrm{mm.yr}^{-1}$) and relatively high MAAT (around 1 °C), in the low range Siberian basin. Then the transition between AP and NAP dominance is marked by decreases in both MAAT (–1 °C) and MAP (300 $\mathrm{mm.yr}^{-1}$) connected to the high-elevation Khangai range. Finally, the dominance of NAP in the Gobi desert area is linked to very arid conditions (MAP < 150 $\mathrm{mm.yr}^{-1}$) and relatively warm MAAT (up to 4 °C). The correlation between the taxa themselves and climate parameters is $R^2 = 0.38$ (RDA, Fig. 5.D). Rise in MAAT is associated with that of Amaranthaceae, Poaceae, *Sedum*-type and Caryophyllaceae percentages. On the contrary, decrease in MAAT is associated with a rise in the AP and Cyperaceae, *Artemisia* spp. and Brassicaceae percentages. MAP, fairly related to RDA_1, rises with AP and decreases with NAP (Fig. 5.D). Finally, the elevation gradient favors *Artemisia* spp. and Cyperaceae for NAP and *Salix* spp. and *Larix sibirica* for AP (Fig. 5.D).

4.1.3 Pollen-inferred Climate Reconstructions: MAT and WAPLS Results

To reconstruct climate parameters from pollen data, MAT and WAPLS methods were applied on the four scales, modern pollen datasets and the ten climate parameters (Table 1). All these methods can be run with $n \in [1; 10]$ parameters: the number of analogues for MAT and the number of components for WAPLS. The best transfer functions among all of them were selected by the following approach: in a first step, for each climate parameter the methods fitting with the higher R^2 and the lower RMSE were selected. Then, in case the highest R^2 and the lowest RMSE were not applied for the same number of analogues or components, we chose the method presenting the lower number of parameters. Despite the small number of parameters relative to the number of observations, the method fits well (Arnold, 2010, Table 1). MAT method gives better R^2 in bigger DB than in smaller ones. Fitting increases with the diversity and the size of DB, since MAT is looking for the closest value between

Table 1. Statistical results of the MAT and WAPLS methods applied to four surface pollen datasets and ten climate parameters[a].

Database	Climate parameter	WAPLS					MAT				
		k best R^2[b]	k best $RMSE$[b]	k selected[c]	R^2 selected[c]	$RMSE$ selected[c]	k best R^2[b]	k best $RMSE$[b]	k selected[c]	R^2 selected[c]	$RMSE$ selected[c]
NMSDB	MAAT	2	2	2	0.65	1.18	**3**	13	3	0.58	1.45
	MTWAQ	2	2	2	0.62	1.69	**2**	9	2	0.68	1.8
	Tspr	2	2	2	0.71	1.26	**2**	13	2	0.63	1.63
	MAP	2	**1**	1	0.79	61.22	**2**	6	2	0.88	55.73
	Pspr	2	**1**	1	0.67	11.39	**2**	4	2	0.89	8.28
	Psum	2	2	2	0.8	34.75	**2**	10	2	0.82	38.49
MDB	MAAT	2	**1**	1	0.35	1.9	**5**	8	5	0.6	1.65
	MTWA	2	**1**	1	0.24	2.14	**5**	9	5	0.53	1.84
	MTCO	2	**1**	1	0.27	2.75	**5**	7	5	0.66	2.05
	MAP	3	**1**	1	0.23	95.05	**9**	11	9	0.38	88.73
	Psum	1	1	1	0.47	47.02	**8**	12	8	0.54	46.17
COSTDB	MAAT	2	2	2	0.54	4.09	**7**	10	7	0.73	3.34
	MTWA	3	**2**	2	0.48	3.55	**8**	10	8	0.67	3.01
	MTCO	2	2	2	0.56	6.34	**6**	9	6	0.77	4.86
	MAP	4	**2**	2	0.55	224.43	**6**	9	6	0.77	161.86
	Psum	3	**2**	2	0.34	70.89	**5**	9	5	0.65	55.08
EAPDB	MAAT	3	3	3	0.72	4.08	**5**	8	5	0.88	2.9
	MTWA	3	3	3	0.55	3.31	**5**	9	5	0.79	2.5
	MTCO	3	3	3	0.72	6.49	**4**	8	4	0.89	4.46
	MAP	3	3	3	0.43	239.6	**4**	10	4	0.74	181.21
	Psum	2	2	2	0.52	62.33	**4**	8	4	0.8	44.66

[a] Only the five better fitting regression models for each climate parameter are shown .

[b] Corresponding to the number of parameters used in the model inferring the best R^2 and $RMSE$.

[c] Number of parameters, R^2 and $RMSE$ of the finally selected model.

climate and pollen abundance. By contrast, WAPLS fits better on the local scale and especially with a smaller number of sites. In this case, the pull of data is largest and the variance is largest (Ter Braak and Juggins, 1993). WAPLS also leads to better value of RMSE than R^2, in contrast to MAT. For temperature, pollen fits better with T_{spr} or MTWA in Mongolia. Temperatures

300   of the warmest months indeed control both vegetation extension and pollen production (Ge et al., 2017; Li et al., 2011) and especially in very cold areas such as Mongolia. For precipitation, the significant season is the one associated with the Summer Monsoon System in Mongolia (Wesche et al., 2016). Almost all the Mongolian precipitation falls during the spring and the summer (Wang et al., 2010), and the amount of precipitation controls, among other parameters, the openness of the landscape in Mongolia (Klinge and Sauer, 2019). To simplify the confrontation of the diverse models, the MAAT and MAP were isolated

305   from the rest of the climate parameters. Even if these two climate parameters are not the best fitting pollen methods, they are the easiest to interpret and are comparable with the GDGT regression models commonly based on MAAT and MAP. In any case, these models are mitigated by the spatial autocorrelation affecting any models made on ecological database (Legendre, 1993) and especially the MAT method more than the WAPLS (Telford and Birks, 2005, 2011).

4.2   GDGT – Climate Calibration

4.2.1   GDGT Variance in the dataset

IsoGDGTs are dominated by $GDGT_0$ and crenarcheol (74,6% and 9.8% in relative abundances, respectively, in Fig. 4.A). These compounds were first attributed to lake water column production (Schouten et al., 2012), but they were also described in significant but lower proportions in soils (Coffinet et al., 2014). The variations of their fractional abundance in our soil and

315   moss polster dataset are discrete and poorly linked to climate parameters (Fig. 4.A). IsoGDGT patterns in lake sediments do not really diverge from surface samples which leads to postulate that the *in-situ* production of isoGDGTs in shallow and temporary lakes like MMNT5C12 is reduced (Fig. 4.A). Then, it appears that the isoGDGT soil-produced are dominated by crenarcheol in accordance with studies on aridity impact (Zheng et al., 2015). However, no relationship exists between [crenarcheol] and MAP ($R^2 = 0.14, p - value > 0.005$). The putative regio-isomer reaction linked to MAP (Buckles et al., 2016) is not evidenced in

320   NMSDB.

The average $[brGDGT]_{tot}$ concentrations differ depending on the sample type:

$$[brGDGT_{tot}]_{sed} = 25.8 \pm 8.1 \text{ ng.g}^{-1}_{sed} \tag{2}$$

$$[brGDGT_{tot}]_{moss} = 23.2 \pm 26.8 \text{ ng.g}^{-1}_{moss}$$

$$[brGDGT_{tot}]_{soil} = 0.3 \pm 0.14 \text{ ng.g}^{-1}_{soil}$$

325   $[brGDGT_{tot}]_{all} = 16.7 \pm 23.6 \text{ ng.g}^{-1}_{sample}$

brGDGT fractional abundances are consistent with each sample type: the predominant compounds are the I_a, II'_a, II_a and III_a (Fig. 4.B). These compounds explain dominantly the total variance (Fig. 5.A). Particularly, the PC_1 represents 22.77 % of the total variance and distinguishes two contrasted poles: the 5-methyl group (mostly with $PC_1 > -0.3$) associated with steppe-

forest and forest sites and the 6, 7-methyl group on the far negative PC_1 values associated with steppe and desert sites.
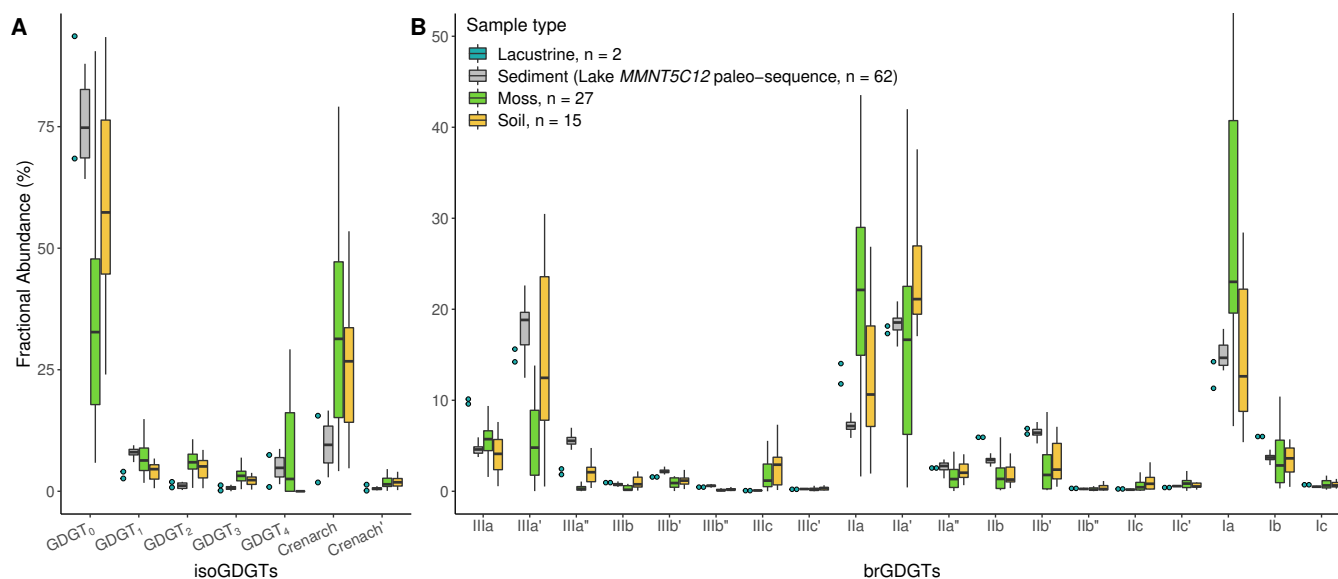
330

Figure 4. Fractional abundances of (**A**) isoGDGTs and (**B**) brGDGTs for moss polsters (green), soil surface samples (orange) mud from temporary dry pond (blue) and the full sequence of the Lake *MMNT5C12* as palaeo brGDGTs comparison (grey). The punctuation marks ' and " refer to 6 and 7 methyl, respectively.

The sediment samples from the lake MMNT5C12, used as past sequence comparison, are more homogeneous than the surface samples, especially when compared with the moss polsters that present a wide variability (Fig. 4.B). On this figure it appears that, globally, soil samples are more relevant analogues to sediments than moss polsters (especially the [IIIa'], [IIa] and [Ia] fractions in Fig. 4). This variability shows an influence of the sample type on brGDGT responses. On the other hand,

335   sample type also bears in first order climate and environment information, since soil and moss polsters originate mainly from steppe to desert environments and forest/alpine meadows, respectively.

4.2.2   Climate Influence on brGDGT Indexes

The brGDGT/climate RDA shows that the brGDGT variance is dominated by the MAP as first component (Fig. 5.B: $RDA_1 = 10.01\%$). The negative values show higher precipitations and uncyclized 5-Me GDGTs, such as I_a, II_a and III_a, while the lower MAP

340   match with 6 or 7-Me GDGTs, such as III'_a, II'_a, II''_a in accordance with De Jonge et al. (2014a). The RDA_2 is slightly more connected to MAAT opposed to elevation, also clustering the methyled and cyclized GDGTs to the higher MAAT. Such as in the pollen-climate response, the elevation is a second driving factor not to neglect. The correlation between relative abundance of methylated and cyclisated brGDGTs with climate parameters was not strong (Weijers et al., 2004; Huguet et al., 2013). All the MBT, MBT', MBT'_{5Me} and CBT, CBT' , CBT_{5Me} relations with climate parameters were tested and showed very
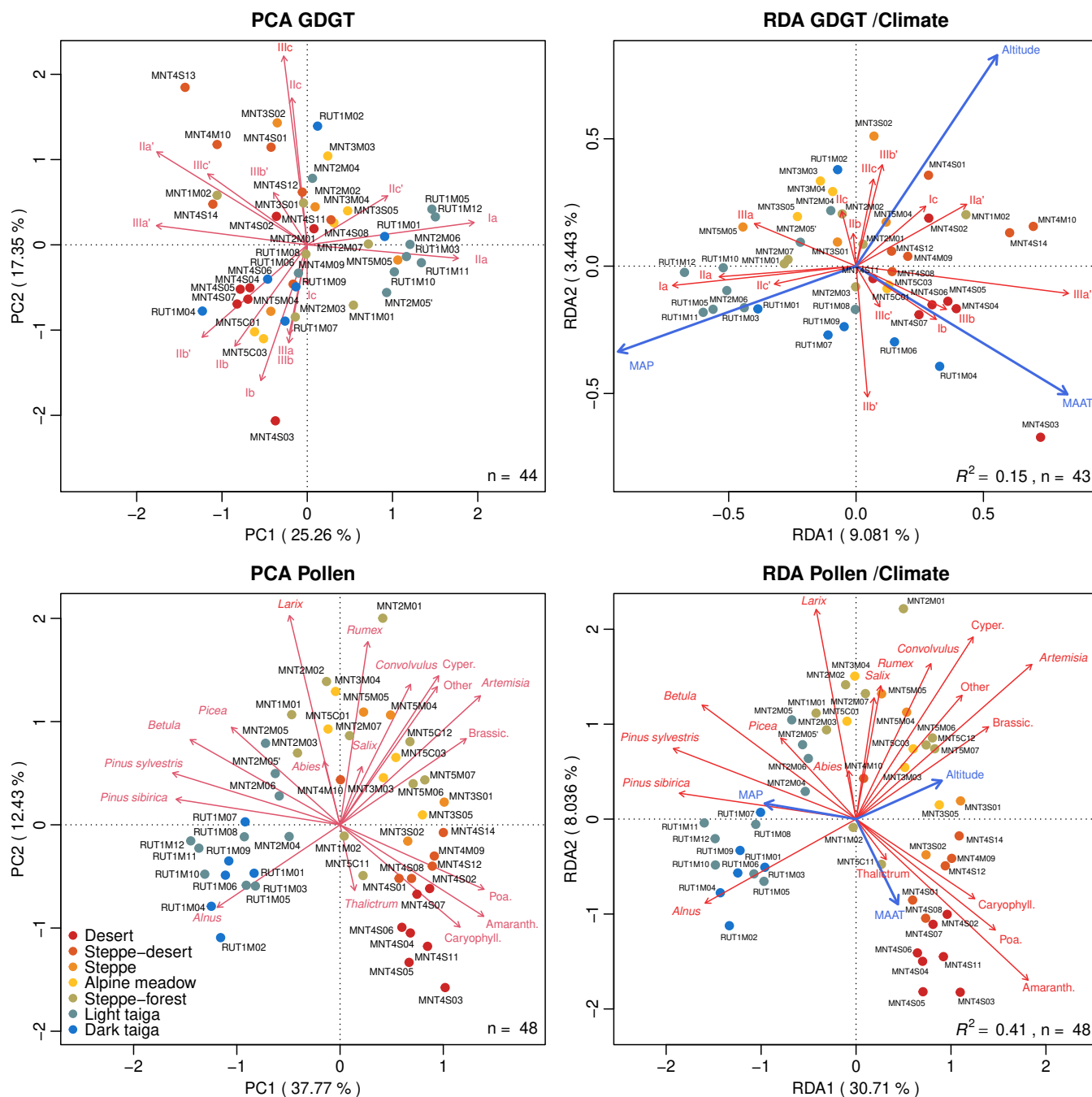
Figure 5. Multivariate statistics for the proxies clustered by ecosystems. **A**: Principal Components Analysis (PCA) and (**B**) Redundancy Analysis (RDA) for brGDGT fractional abundances; **C**: PCA and (**D**) RDA for pollen fractional abundances. The variance percentage explained is displayed on the axis label; the size of the dataset (n) and the RDA linear regression (R^2) are inserted in each plot area.

345   low correlation with $R^2 \in [0.1; 0.35]$ (Supplementary Fig. S3.B). Once the MBT (Supplementary Fig. S2.A) and the $\mathrm{MBT}'_{5\mathrm{Me}}$ indexes (Fig. 6) compared with the world database (Yang et al., 2014; Naafs, 2017; Dearing Crampton-Flood et al., 2019) it appears that the NMSDB set is consistent with known values instead of a while sample dispersion.
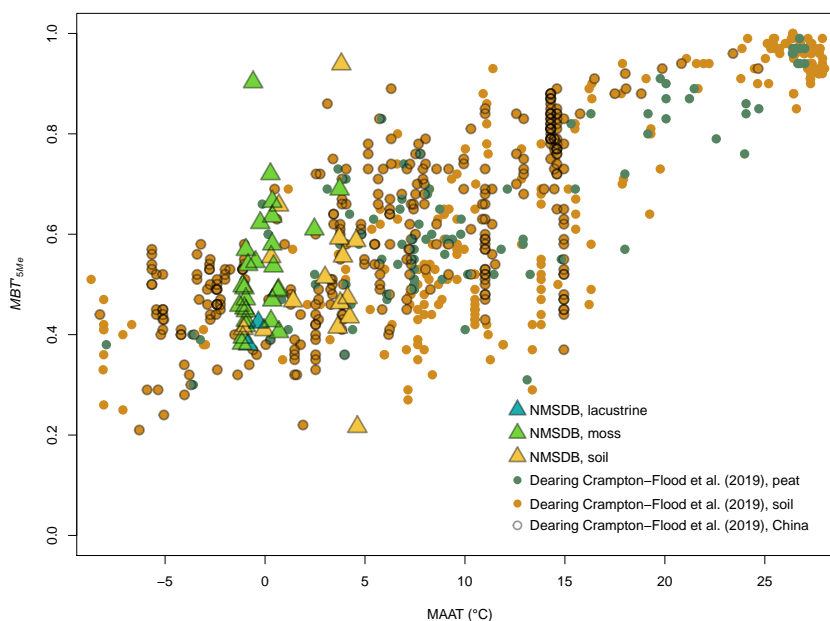
Figure 6. $\mathrm{MBT}'_{5\mathrm{Me}}$–MAAT relation comparison between the NMSDB surface samples and the world peat and soil database Dearing Crampton-Flood et al. (2019).

4.2.3 Multi-regression Models

The Stepwise Selection Model for climate – brGDGT modelling was applied only on the 5 and 6 Methyls, because 7-methyled

350   brGDGTs show weak significance in the variance explanation (PCA, Fig. 5.A). The $\mathrm{N_{SSM}}$ different combinations of the 15 brGDGT compounds result in $\mathrm{N_{SSM}} = 2^{15} = 32768$ models possible for each climate parameter. Even the models including some minor compounds $([\mathrm{br}]_i < 5\%)$ have been considered since, in the NMSDB, brGDGT fractional abundances are more fairly distributed than in the global database, in which few compounds overlap the majority of the compound (Supplementary Fig. S2.B). Indeed, the cumulative fractional abundance curve (Supplementary Fig. S2.C) is fairly faster to rich the asymptote

355   line for the world peat (blue curve) and the world soil (in brown) than in the NMSDB. The world peat database needs only three brGDGTs to explain more than 85 % of the fractional abundance against more than ten compounds in the NDMSDB soils. Then, the better fitting equations (with low RMSE and AIC and high R-squared) were selected for each number of parameters (number of brGDGT issued in the linear regression) for both MAAT and MAP. Within the 15 models (one model for each

17

parameter addition), the 9 more contrasted ones were selected for discussion (Supplementary Table S3). The models with the

360 best statistical results were comprised of between 5 to 12 parameters and present a $R^2 \in [0.60; 0.76]$, a RMSE around 1.1 °C or

70 $mm.yr^{-1}$ and a $\text{AIC}_{\text{MAAT}} \in [152.6; 166.2]$ or $\text{AIC}_{\text{MAP}} \in [152.6; 166.2]$. The earlier a parameter is used in the mr models,

the greater is its influence. For both MAAT_{mr} and MAP_{mr} models, III_{a}, III'_{a}, III_{b} and III'_{b} are the most relevant compounds

for the climate reconstruction (Table 2) which is consistent with the PCA and RDA observations displayed (Fig. 5.A and

B). Both the MAAT_{mr} models infer on a positive contribution of III'_{a} and a negative contribution of III_{a}, which confirms

365 these models as eco-physiologically consistent with the RDA results. Moreover, except for II'_{b}, all the compounds positively

correlate with MAAT and negatively with MAP, in accordance with the MAP–MAAT anti-correlation. The ΔT values closest

to 0 reveal the best fitting model on each point (Fig. 8, panel 1). Then, the box-plot (Fig. 8, panel 2) summarises the best fitting

model at a regional scale.

Table 2. Statistical values and equations of the best brGDGT MAAT_{mr} and MAP_{mr} models.

Model	Formula	R^2	RMSE	AIC
MAAT_{mr3}	$= 0.6 \times 1 - 25.1 \times [\text{IIIa}] + 12.3$ $\times [\text{IIIa}'] + 7.2 \times [\text{Ib}]$	0.57	1.3	156.4
MAAT_{mr5}	$= 4.8 \times 1 - 38.5 \times [\text{IIIa}] + 7.9 \times [\text{IIIa}']$ $- 27.3 \times [\text{IIIc}] - 3.3 \times [\text{IIa}'] - 26.3 \times [\text{IIb}]$ $+ 8.5 \times [\text{IIb}'] - 5.6 \times [\text{Ia}]$	0.66	1.1	153.9
MAP_{mr6}	$= -511.3 \times 1 + 1205.9 \times [\text{IIIa}] + 1387.2 \times [\text{IIIb}]$ $+ 738.8 \times [\text{IIa}] + 969.8 \times [\text{IIa}'] + 1957.1 \times [\text{IIb}]$ $+ 3006.3 \times [\text{IIb}'] + 2406.4 \times [\text{IIc}] + 2003.1 \times [\text{IIc}']$ $+ 1081.7 \times [\text{Ia}] - 2406.3 \times [\text{Ic}]$	0.73	72.5	525.8
MAP_{mr7}	$= -502.6 \times 1 + 1359.5 \times [\text{IIIa}] + 2462.5 \times [\text{IIIb}]$ $- 2178.3 \times [\text{IIIb}'] + 657.7 \times [\text{IIa}] + 986.8 \times [\text{IIa}']$ $+ 2440.5 \times [\text{IIb}] + 3423.5 \times [\text{IIb}'] + 2831.2 \times [\text{IIc}]$ $+ 1967.2 \times [\text{IIc}'] + 1150.2 \times [\text{Ia}] - 955.6 \times [\text{Ib}]$ $- 2103.4 \times [\text{Ic}]$	0.75	68.5	524.8

Climate
of the Past
Discussions

5 Discussion

5.1 Issues in Modelling Mongolian Extreme Bioclimate

5.1.1 Appraisal Modelling in Arid Environments

According to Dirghangi et al. (2013) and Menges et al. (2014) ; the commonly used brGDGT indexes (MBT and CBT) are not relevant for arid areas with $MAP < 500mm.yr^{-1}$ because of the relation between low soil water content and soil brGDGT preservation and conservation interferes in the brGDGT / climate parameters (Dang et al., 2016). The MAAT models based on MBT and MBT' indexes provide colder reconstructions (Fig. 8.C2) as shown by De Jonge et al. (2014a), because arid soils favors 6-Methyl (by pH raising due to low weathering effect of the weak precipitation, Dregne, 1976; Haynes and Swift, 1989) and drives the MBT to decrease towards zero. This explains the colder $MAAT_{Ding}$ and $MAAT_{MBT'_{DJ}}$ reconstructed values compared to the modern ones. Moreover, the main issue in climate modelling in Mongolia is the narrow relationship between MAAT and MAP. Because of the climatic gradient from dry deserts in the southern latitudes to wet taiga forests in the northern ones, MAAT and MAP maps are strongly anti-correlated (Fig. 1 .B, C and Supplementary Fig. S1). If this correlation is not statistically determined on the range of the global database ($R^2 = 0.35$, $p < 0.005$), the impact is significant on the range of the Mongolian sites ($R^2 = 0.91$, $p < 0.005$). This correlation could be a bias resulting from the interpolation method of the *WorldClim2* database. In fact, there are very few weather stations (Fig. 1.A, Fick and Hijmans, 2017) and their distribution on the large Mongolian plateau is interrupted by mountain ranges. According to Fick and Hijmans (2017) the interpolation model used in the Central Arid Area (which includes our study area) gives a strong correlation ($R^2 = 0.99$) and a little error ($RMSE = 1.3°C$) for MAAT and $R^2 = 0.89$ and $RMSE = 23mm.yr^{-1}$ for MAP. Whenever the Siberian-Mongolian calibrations are used for palaeoclimatic reconstructions, the RMSE of the climate parameters has to be added to the RMSE model. Moreover, the relevance of the interpolation models suffers from the transition threshold made in Mongolia between the EASM and the Siberian Westerlies (Fig. 2, An et al., 2008) and reinforced by the topographic break (Fig. 1.A). Because the mr–GDGT models have been compiled with the group of Siberian sites which are out of the MAAT–MAP strong auto-correlation range (Supplementary Fig. S1), the reliability of the independence of the MAAT and MAP models seems to be guaranteed.

The topographic fence in Mongolia also affects the pollen and brGDGT distributions by it-self, as seen in both RDA analyses (Fig. 5.B and D) where elevation appears to be a main ecophysiological parameter. Elevation affects vegetation and pollen rain not just because of its influence on local MAAT and MAP but also because it drives other ecophysiological parameters such as 0_2 concentration, wind intensity, slopes and creeping soils, snow cover and exposure (Stevens and Fox, 1991; Hilbig, 1995; Klinge et al., 2018). Elevation as one of the main brGDGT drivers could also be explained by the bacterial community responses to pH, moisture and soil compound variations along the altitude gradient (Laldinthar and Dkhar, 2015; Shen et al., 2013; Wang et al., 2015) and the vegetation shifts (Lin et al., 2015; Davtian et al., 2016; Liang et al., 2019).

5.1.2 Particularity of the Siberian–Mongolian climate system

Both GDGT and pollen calibrations show that the precipitation calibrations are more reliable than temperature ones (Tables 1, 2, Figs. 3, 7 and 8), reflecting that the Siberian-Mongolian system seems to be mainly controlled by precipitation. This dominance of precipitation could be due to seasonality. Even if the brGDGT production is considered to be mainly linked to annual temperature means (Weijers et al., 2007a, b; Peterse et al., 2012), the high pressure Mongolian climate system (Zheng et al., 2004; An et al., 2008) favors a strong seasonal contrast: almost all the precipitation and the positive temperature values happen during the summer (Wesche et al., 2016). Consequently, for the NMSDB pollen transfer functions, the seasonal parameters such as MTWA, T_{sum} and P_{sum} better describe the pollen variability than MAAT and MAP climate parameters (better R^2 and RMSE in Table 1). While the opposite is found on EAPDB and COSTDB models, the calibration made on large scale databases. The Mongolian permafrost persists half the year in the northern part of the country (Sharkhuu, 2003) and acts on vegetation cover and pollen production (Klinge et al., 2018). Furthermore, the effects of frozen soils on soil bacterial communities and GDGT production are thought to be important (Kusch et al., 2019). This seasonality leads to a quasi equivalence between MAP and P_{Sum} (if $P_{win} \approx 0$ then $MAP \approx P_{sum}$) while MAAT is torn apart by the large $T_{Sum} - T_{win}$ contrast (because the MAAT is an average value and not a sum as for MAP). The mathematical consequences of the seasonality on these two climate parameters are not the same. Finally, the MAP appears to be the most reliable climate parameter for Siberian-Mongolian climate studies according to the NMSDB sites (with $MAAT < 5°C$). Even if the brGDGTs seem to react to summer temperature (Wang et al., 2016; Kusch et al., 2019), the summer mr-models are not significantly improving the calibration compared to the $MAAT_{mr}$ ones. For instance, the best $Tsum_{mr}$ is selected by its AIC, $Tsum_{mr6}$ is inferred using 6 brGDGTs fractional abundance ($R^2 = 0.63$ and $RMSE = 1.53°C$). This lack of seasonality effect, expected in such cold areas, is consistent with temperate Chinese sites (brGDGT reconstructions, Lei et al., 2016).

5.1.3 Extreme Bioclimatic Condition Modelling Lead to a Better Global Climate Understanding

To reduce the signal/noise ratio, a wider diversity of sample sites should be added as initial inputs in the models. This raises the question of the availability of reliable samples in desert areas. The soil samples in the steppe to desert biomes are often very dry and these over-oxic soil conditions are the worst for both pollen preservation (Li et al., 2005; Xu et al., 2009) and GDGT production (Dang et al., 2016). brGDGT concentrations in moss polsters and temporary dry pond muds are thus higher than in soils in our database (equation 2 and Fig. 4). The explanation of the signal difference between the three types of samples could also originate from the *in-situ* production of brGDGTs inside the moss predominant over the wind-derived particles brought to the moss net. As well, it seems that the pool of moss polster is associated with a similar trend than the worldwide peat samples from Naafs et al. (2017b) (Supplementary Fig. S2.A and Fig. 6). Moreover, in the steppe or desert context of poor availability in archive sites, the edge clay samples or top-cores of shallow and temporary lakes could be a solution for palaeo-sequence studies. The two pond mud samples of the NMSDB are included within the soil-moss trend for all models (Fig. 5, Supplementary Fig. S2 and S3).Even if the brGDGT production and concentrations are different in soils than in lakes due to lake

in-situ production (Tierney and Russell, 2009; Buckles et al., 2014), this effect is function of the lake depth (Colcord et al.,
435   2015), consequently negligible for shallow lakes, and almost absent for lake edge samples as shown by Coffinet (2015) for
Lake Masoko in Tanzania.

The soils of the Gobi desert also have a high salinity level which is also a parameter of control on brGDGT fractional
abundances (Zang et al., 2018). This taphonomic bias (also climatically induced) could explain part of the histogram variance
440   of Fig. 4 related to the sample type as well as the shift of the soil–cluster from the regression line in the cross-value plot of
brGDGT MBT'/CBT models in Supplementary Fig. S3. Even if the impact of salinity on sporopollenin is not well understood,
salt properties may affects pollen conservation in soils (Reddy and Goss, 1971; Gul and Ahmad, 2006).

Finally, the saturation effect of the proxies when they reach the limits of their range of appliance is also to be taken into
445   considerations. Since both pollen and brGDGT signals are analysed in fractional abundance (i.e. % of the total count of con-
centration), these proxies evolve in a $[0; 1]$ space. The saturation effect appears when extreme climatic conditions are reached
(Naafs et al., 2017a, b). For instance, in a tropical context, temperature values are too high to be linearly linked to fractional
abundances (Pérez-Angel et al., 2019). Considering pollen–climate relationships, the inferior limit of pollen percentage is crit-
ical: for the majority of pollen types, whenever MAAT or MAP reaches a very high or low threshold, the pollen fractional
450   abundance approaches zero (Fig. 9). These limit areas need to be closely investigated, which legitimises the local calibration
methods.

5.2 Statistical Tools for Best Model Selection

5.2.1 Over-Parameterization and Best Models Selection

Among the possible methods, statistical values help to select the most reliable ones for palaeoclimate reconstruction. However,
455   the correlation (R^2) and errors (RMSE) are not informative enough to discriminate between methods and to point to the
most suitable ones for palaeoclimate modelling. This is especially true for the multi-parameter methods (such as brGDGT
multi-regression models and pollen transfer functions). Indeed, the more input parameters in the method, the more accurate
it is (Tables 1, Supplementary Table S3 and Fig. 7.A and 7.B). All the regression models improve with parameter additions,
and especially the less fitting methods improve exponentially (lower limit of the R^2 area, Fig. 7.B). The best R-squared-
460   models for each parameter number (Fig. 7.A) correspond to the upper limit of the R^2 area (Fig. 7.B). This figure shows
that the R^2 vs. *parameter number* trend follows a logistic regression (both for $MAAT_{mr}$ and MAP_{mr} models). However,
and especially for $MAAT_{mr}$ regression models, this logistic curve becomes asymptotic early, similar to the RMSE decrease.
The over-parametrization of the models has proven to produce artefacts in ecological modelling (Arnold, 2010; Symonds and
Moussalli, 2011). The issue is thus to identify the threshold in the parameter numbers selected. We used Akaike's Information
465   Criterion (AIC) to determine the better model without over-parameterization for brGDGT regression models: the lower the
AIC, the better the model (Supplementary Table S3 and Table 2). The trend of AIC versus the parameter number is however

more complex (Fig. 7.C). For $MAAT_{mr}$, the regression model becomes more accurate from one to five parameters rapidly, but then slowly decreases. The AIC curve takes an asymmetrical hollow shape around five parameters with a steeper slope on the left side (Fig. 7.A). The AIC values for $MAAT_{mr6}$ and $MAAT_{mr7}$ are almost identical (Fig. 7.A) . The $MAP_{mr6,7,8}$

470 have almost equivalent AIC values, while the AIC curve shapes differ for the other MAP_{mr} models (asymmetrical hollow shape around five with a steeper slope on the left side, Fig. 7.A). To summarize, the most universal models are $MAAT_{mr5}$ and MAP_{mr7} (Table 2) but the closed models are also valuable in some local contexts, and especially in similar dry-cold regions. We need to determine the cross-values of these models to select the appropriate ones for the Siberian-Mongolian context.
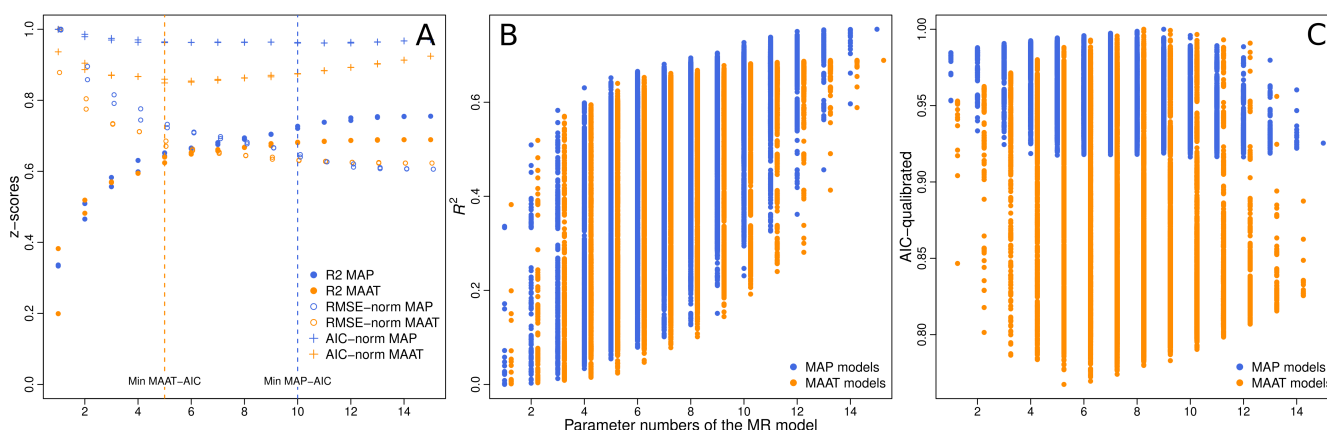
Figure 7. Statistical values plotted against the number of parameters of the different mr–GDGT models: the R^2, the RMSE normalized on the highest RMSE value and the AIC also normalized. **A**: selection of the two best multi-regression models for each number of parameters; **B**: combination of the R^2 (**B**) and the AIC (**C**) values for of all the mr models. The blue dots are for the MAP-models, orange dots for the MAAT one.

5.2.2   Assessment of the Calibration Feedback

475 The cross-values of the nine best $MAAT_{mr}$ regression models (Fig. 8.A1 and 8.A2) and the best MAP_{mr} regression models (Fig. 8.B1 and 8.B2) were tested. The MAAT reconstructions provide different responses to the three main bio-climate areas (parcel A1): if they properly estimate temperatures in the Siberian Baikal basin, they overestimate and underestimate them for the center of the northern Mongolian mountains and the Gobi desert, respectively. For precipitation (parcel B1), all the MAP_{mr} calculated with local to regional databases also misrepresent the extreme values: precipitation values are too high and too low

480 for the Gobi desert and the Baikal basin, respectively. To conclude, the wider the dataset extension, the more alleviated the extreme values.

Both on $MAAT_{mr}$ and MAP_{mr} models, the 95 % interval shrinks with parameter addition, but the mean values do not necessarily get closer to the measured value of the climate parameter (the dashed line in Fig. *8.A2 and B2*). Therefore, if the
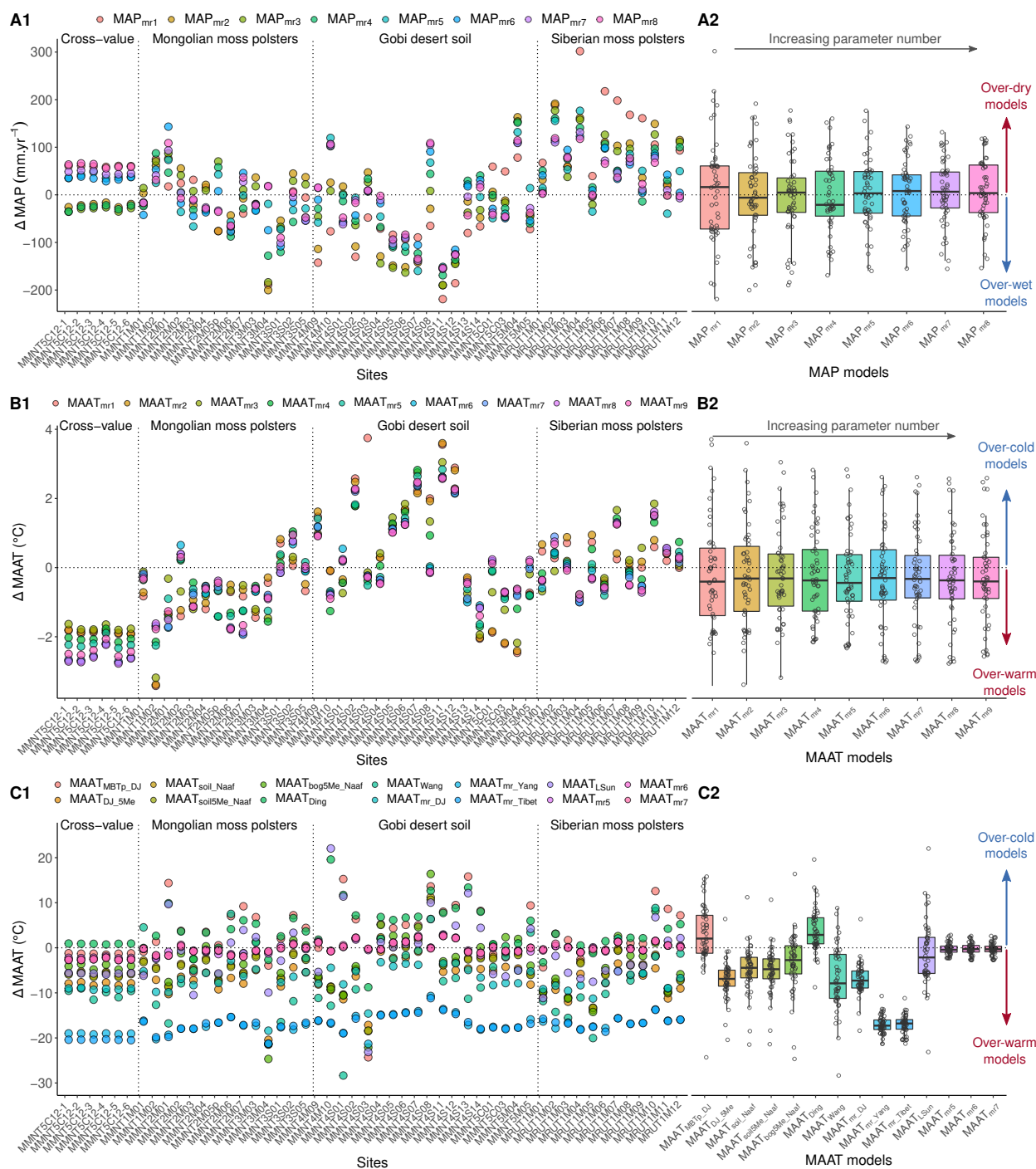
Figure 8. Validation of brGDGT-climate models on the study sites: reconstructed values for literature MAAT (**A**), NMSDB mr–MAAT (**B**) and MAP (**C**). Models are tested on the NMSDB sites (**1**) and the box-plot statistics (**2**) are provided. Sites are clustered in 4 groups: cross-value on the 6 first samples of the independent core MMNT5C12, Arkhangai; moss polsters from Mongolian steppe-forest; Gobi steppe-desert soil samples and moss polsters from Baikal basin. Values are plotted in anomaly.

485   tests on the AIC point toward the MAAT_{mr4} and the MAP_{mr7} regression models, the back-cross plots suggest the MAAT_{mr3} and the MAP_{mr6} regression models (Supplementary Table S3, coloured in blue and Table 2) provide the best estimates for climate reconstruction in lacustrine archives ($\Delta\mathrm{MAP} = 0$ and best fitting temperature for the mean value of all samples, Fig. 8.B2 and Fig. 8.B1).

5.2.3   Global vs. Local Calibration

490   Whatever proxy is used, when reconstructing temperatures and precipitation from past records in a given location, there is the issue of basing reconstructions on calibrations based on local or global datasets (among others, Tian et al., 2014; Cao et al., 2014; Ghosh et al., 2017; Dearing Crampton-Flood et al., 2019). We tested both approaches on our datasets with a cross-value run on the NMSDB-independent set of MMNT5C12 core samples. The global brGDGT - climate calibration artificially reaches higher R-squared than local ones due to the larger range of values of the involved climate parameters. Since
495   the world soil database in Naafs et al. (2017) covers a wide temperature range ($\mathrm{MAAT} \in [-5;30]$), counter to the NMSDB ($\mathrm{MAAT} \in [0;5]$), then its signal/noise ratio gets lower (Fig. 6). Despite the relatively lower R-squared of 0.62 scored by the MAAT_{mr5} compared with world calibrations (Pearson et al., 2011; De Jonge et al., 2014a; Naafs et al., 2017a, b), the boxplots for the all MAAT_{mr} calculated from the NSMDB are mostly centred on the MAAT average value with the shortest variance spreading for all the sites (Fig. 8.C1 and 8.C2). These local calibrations fit best with the MAAT_{Ding} from Ding et al. (2015)
500   which is also a local calibration made on the Tibet-Qinghai plateau database. The global databases made on worldwide sites (De Jonge et al., 2014b, a; Naafs et al., 2017a, b) provide $\mathrm{MAAT}_{model} > \mathrm{MAAT}_{real}$ and large standard deviation (SD). These global calibration also attenuate the exterme MAAT values: the very cold Siberian basin / Mongolian plateau sites are reconstructed with warmer temperatures up to +5 to + 10 °C, while the warm Gobi desert sites are down by up to -3 to - 5 °C. On the other hand, the local calibrations performed on subtropical to tropical Chinese transects (Yang et al., 2014; Thomas
505   et al., 2017) have smaller SD but largely overestimate MAAT values due to the warmer conditions of the initial database sites. In brief, the lake core sediment samples match the best to the modern MAAT and MAP value with the brGDGTs mr–models which invite us to consider that these local brGDGT calibrations present a robust way to approach past climate.

Similarly, for pollen transfer functions, the geographic range of the surface samples on which the calibration relies is a rel-
510   evant parameter to take into account for the reliability of the paleoclimate reconstructions. The choice of the maximum value of this geographic range has been discussed previously for vegetation modelling, for example, the Relevant Source Area of Pollen (RSAP, Prentice, 1985; Hellman et al., 2009a, b; Bunting and Hjelle, 2010). For MAT and WAPLS regression models, the same issue holds true. The responses of the eight over-represented *taxa* to climate parameters are different in the three geographic ranges (NMSDB, COSTDB and EAPDB). The linear tendency allows for checking the main trends between taxa
515   distribution and climate parameters, despite the weak linear regressions ($p-value > 0.005$ and $R^2 < 0.4$, in Fig. 9). For the majority of these *taxa*, the trend is the same, independent of the database size (*Larix* spp. and Cyperaceae percentages increasing with weaker MAAT, or Amaranthaceae and *Pinus sylvestris* percentages increasing with higher MAAT). However, due to the shift between pollen types and their associated vegetation (i.e. Poaceae-pollen signal similar for a wide diversity
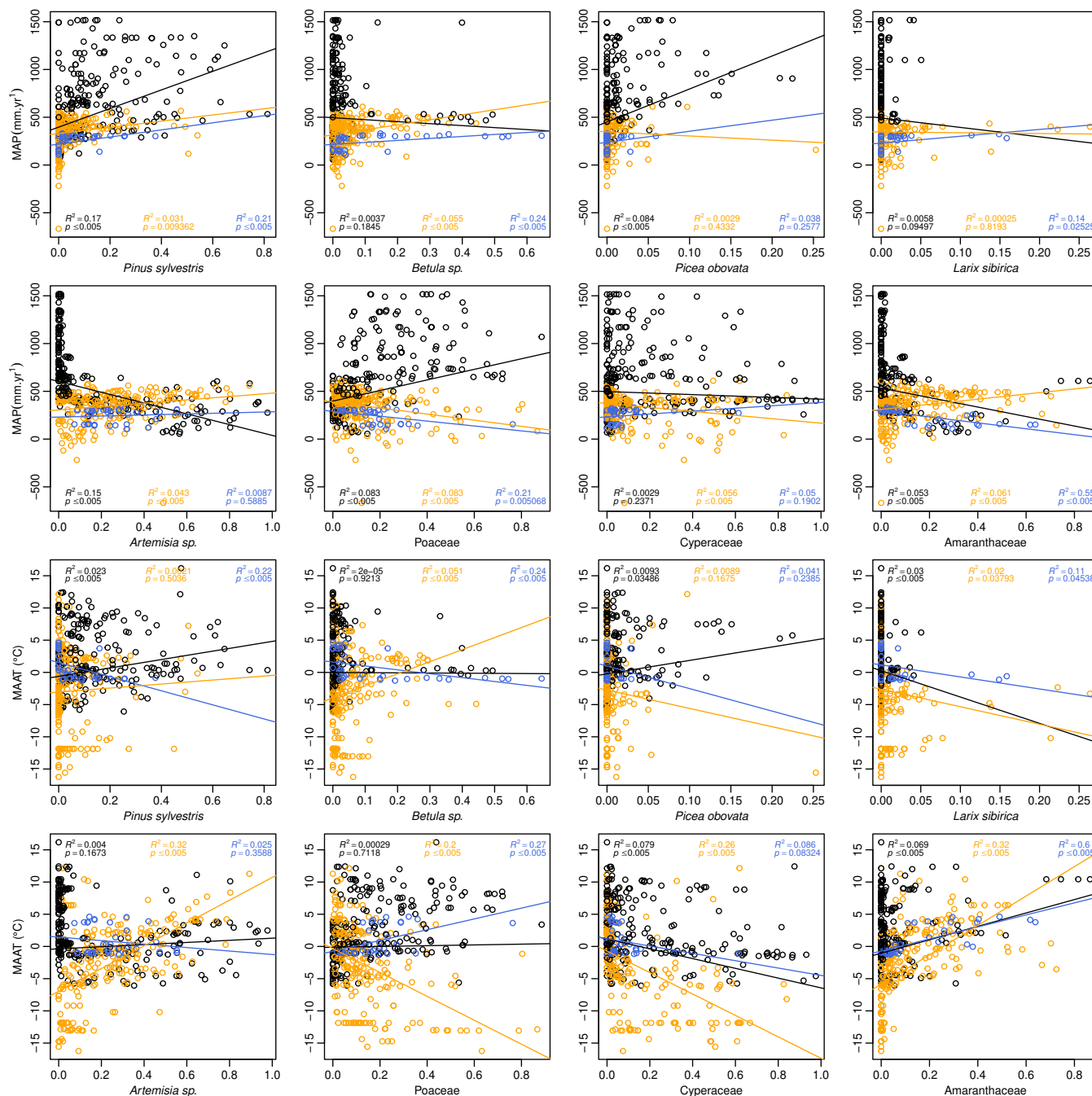
Figure 9. Relationships between the eight major pollen *taxa* (%TP) and MAP (mm.yr^{-1}, upper part of the facet plot) and MAAT (°C, lower part). The black line is the linear fitting for all samples (EAPDB), the orange for all the samples from steppe biome (COSTDB) and the blue only for the NSMDB samples presented in this article.

of Poaceae communities with very contrasted ecophysiological features), trends are controlled in some peculiar cases by the
520 geographical clipping of the DB. Thus, Poaceae have a positive response to MAP on the global scale but not inside the Mongolian area. The human influence on pollen rain is also dependent on the biogeographical context, thus, *Artemisia* spp. is not considered as much human influenced in the Asian steppe environment (Liu et al., 2006) than in the European one (Brun, 2011).

Concerning transfer functions, WAPLS performs better for the local database than for the COST and EAP database (Table
525 1). On these subsets, the WAPLS RMSE and R-square values are even higher than for the MAT transfer function. The major difficulty resides in the reconstructions of precipitation. Even if the RMSE and R^2 values are higher for all models of MAP than MAAT, the influence of precipitation on vegetation cover is not well understood. In Mongolia it is clear that the precipitation controls the treeline in mountainous areas (Klinge and Sauer, 2019) and the global openness in the steppe - forest ecotone (Wesche et al., 2016) as well as human land-use (Tian et al., 2014), but the risk of autocorrelation between MAAT and MAP
530 signals is important, even if the RMSE and R^2 values are higher for MAP regression models than for MAAT ones (Telford and Birks, 2009; Cao et al., 2014). Tangibly, for the two proxies, even if the global calibrations can operate on our study area, the local calibrations reach higher accuracy.

5.2.4 Test-bench of the Local Calibrations on Two Paleosequences

To test the reliability of our local calibrations, the pollen transfer function and the brGDGT mr-models have been applied on
535 paleo-sequence. Because there is still no available core analyzed both for pollen and brGDGTs in ACA, the D3L6 core (pollen, Altai, Unkelbach et al., 2019) and the XRD section (Qaidan, GDGTs, Sun et al., 2019) are used. The actual values of the climatic parameters are first compared to the top-core reconstructed climatic parameters (Fig. 10, dashed lines). The amplitude of the variations through time has then to be assessed with regards to the expected regional ranges (Zheng et al., 2004). Finally, reconstructions on known short term climate events are tested for the last 5000 years. They are namely the Little Ice Age,
540 Warm Medieval Period, Dark Ages Cold Period, Roman Warm Period and 4.2k year event (respectively LIA, WMP, DACP, RWP and 4.2 ; Zhang et al., 2008; Chen et al., 2015; Aichner et al., 2019).

For the pollen transfer function (Fig. 10.A), the inferred reconstructions display similar trends during the 4500 years with larger amplitude for MAT than for WAPLS. MAT is also more sensitive to the initial calibration dataset selected than WAPLS
545 (Fig. 10.A) which is in agreement with previous multi-methods studies (Brewer et al., 2008). Indeed, the COSTDB and EAPDB are over-reactive in front of local calibration, while the WAPLS display same amplitude for each calibration with a different offset. Particularly, the NMSDB consistently displays the values closest to the actual climate parameter values both for MAT and WAPLS. For the D3L6 precipitations, all models seem to drive away from the actual value. This shift highlights the *WorldClim2* interpolation issues in the Altai mountains (few weather stations and not accounting for snow melt).

550

On Fig. 10.B, about brGDGTs, the local calibrations (NMSDB applied to XRD section) provide the closest surface reconstructed temperature values to the actual (ΔMAAT $< 2°$C, Fig. 10.B on the left panel) compared to the global calibration
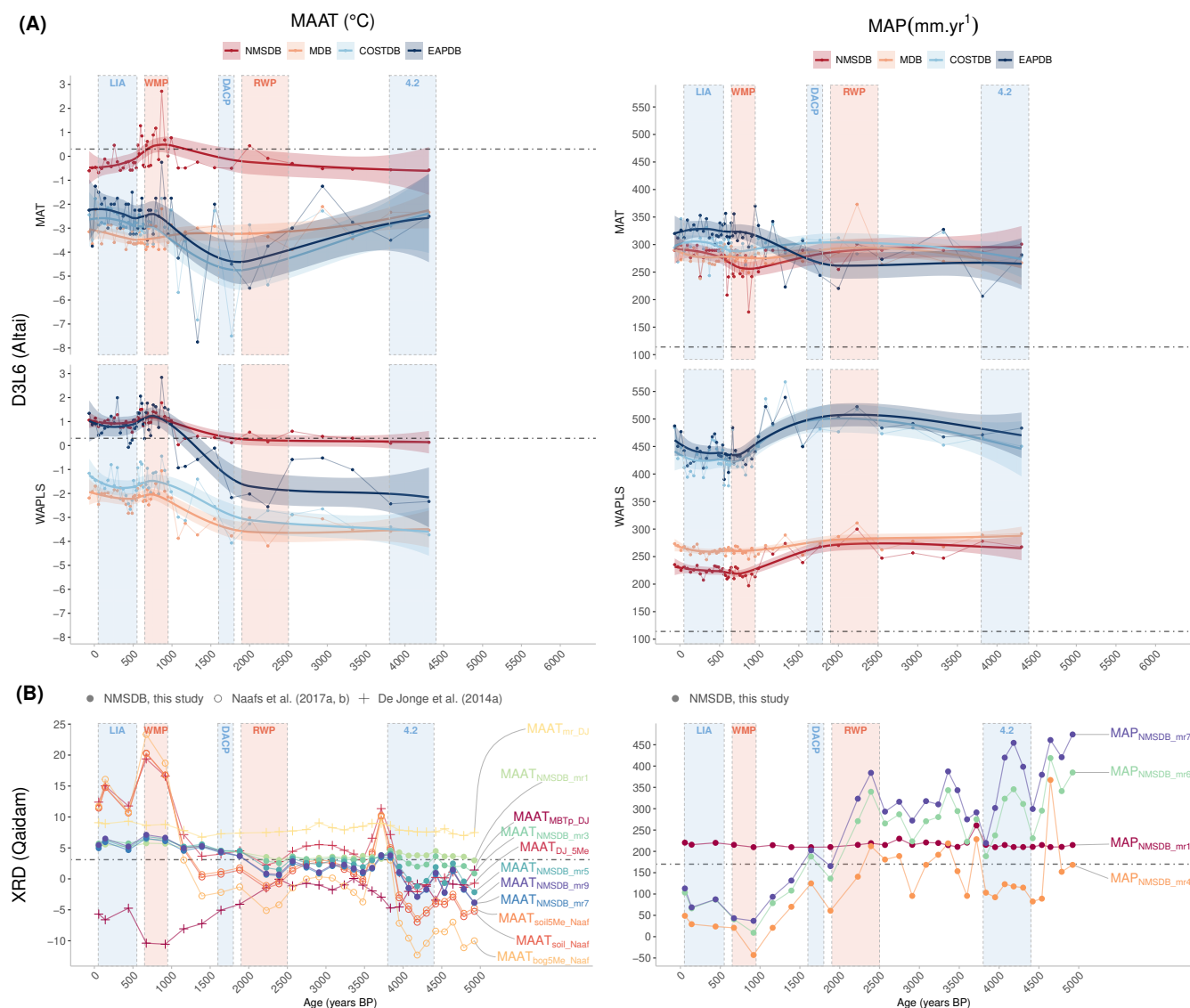
Figure 10. ACA climate reconstruction for the 5000 year BP. (**A**): climate–pollen inferred for the Lake D3L6 (Unkelbach et al., 2019) comparing two transfer function methods (WAPLS and MAT) and the 4 databases (EAPDB, COSTDB, MDB and NMSDB) ; (**B**) climate–brGDGT inferred from the XRD section (Sun et al., 2019) comparing local (NMSDB) and global calibrations (De Jonge et al., 2014a; Naafs et al., 2017a, b). The climate periods correspond to Little Ice Age (LIA), Warm Medieval Period (WMP), Dark Ages Cold Period (DACP), Roman Warm Period (RWP) and 4.2k year event according to Zhang et al. (2008), Aichner et al. (2019) and Sun et al. (2019). Dashed lines represent the actual surface climate parameters foe each cores.

(ΔMAAT $\in [6; 10]°$C, De Jonge et al., 2014a; Naafs et al., 2017a, b). This is explicable by the high similarity between the type of sediment from the XRD section (dry taphonomy) and a large proportion of the NMSDB surface samples (especially

555 from Gobi desert). Moreover, the NMSDB calibrations present a more realistic amplitude: 6 °C over 5000 years (similar to pollen-inferred amplitude) as opposed to 11 to 14 °C amplitude for global calibrations. About the precipitation, the brGDGTs mr–models show a decreasing trend along the Holocene, particularly well-marked between 1000 and 2000 yr BP followed by a bounce on the last 1000 years (tendency consistent with the WAPLS pollen MAP). All brGDGT calibrations exhibit consistent shifts during the LIA (cold-wet), the WMP (warm-dry) and the 4.2ka event (cold-wet). These variations are also exaggerated

560 with global calibrations. To conclude, general trends are consistent for all calibration datasets, except for the drying-warming trend inferred by the calibration from De Jonge et al. (2014a).

Figures 10.A and 10.B show that realistic reconstructed surface values are consistent with literature Holocene trends and validate the application of local calibrations for both pollen and brGDGTs. Furthermore, abrupt oscillations and overall ampli-

565 tudes of temperature and precipitation variations are realistic, in accordance with regional appraisal (Wu et al., 2020). These results permit us to improve our understanding on the Mongolia-China Late Holocene climate variations. Overall, on the 5000 year period the climate appear to follow a dryer-warmer trend. More precisely, both for pollen and brGDGTs the local calibration shows a correlation between temperature and precipitation short-period oscillations: the LIA seems to be colder and wetter than the warm and dry WMP. The same cold-wet behavior is observed for the 4.2ka event. This conclusion is important and

570 situate the D3L6 and the XRD sequences in the same trend that the majority of the ACA paleosequences (Chen et al., 2010, 2015; Wu et al., 2020), connecting the Altai range and the Qaidam basin to the EASM vs. Westerlies Holocene oscillations.

6 Conclusions

The palaeoenvironmental and palaeoclimatic signals may present several uncertainties (differential production, preservation...) which can misguide the interpretation of past variations. This study shows how both a multi-proxy approach and an accurate

575 calibration are important in preventing from these biases. We propose a new calibration for Mean Annual Precipitation (MAP) and Mean Annual Air Temperature (MAAT) from brGDGTs as well as a new pollen surface database available for transfer functions. The correlations between pollen rain and climate on one hand and brGDGT soil production and climate on the other are visible but are still mitigated by the complex climate system of Arid Central Asia and the diversity of soils and ecosystems. Precisely, each of our proxies seems to be more narrowly linked to precipitation (MAP) than temperature (MAAT) counter to

580 the majority of calibrations in the literature. This is validated on both modern and past sequences for pollen and brGDGTs. The nature of the samples considered (soil, moss polster and mud from temporary dry pond) also greatly affected these correlations. The calibration attempt for the extreme bio-climates of the Siberian basin and Mongolian plateau is difficult because of the low range of climate values, despite the climate diversity ranging from cold and slightly wet (north) to the arid and warm (south) conditions. Even if global and regional calibrations could be applied in such a setting, local calibrations provide enhanced

585 accuracy and specificity. The MAAT and MAP values do not remarkably spread in the vectorial space, which makes harder to

distinguish the linear correlation against variance noise. Moreover, this range of values is close to the lower saturation limit of the proxies, which makes the accurate local calibration tricky but necessary. The local calibrations also suffer from the reduced size and small geographic extent of the dataset. The vegetation cover, extending from a high cover taiga forest to bare soil desert cover, also buffers the climate signal and the GDGT / pollen response. The correlations between climate parameters and GDGT

590 / pollen proportion are therefore lower than they could be at global scale. Nonetheless, and despite the lower correlation of the local calibration, these local approaches appear to be more accurate to fit the actual climate parameters than the global ones: both for pollen transfer functions and brGDGT multiple regression models. These positive model results have to be considered in light of over-parameterization limits. Too many parameters in mr–brGDGT models or in pollen MAT or WAPLS transfer function can add artificially to the linear relation between climate and proxies and lead to misinterpretation of palaeoclimate

595 records. Akaike's information criterion associated with RMSE and R^2 values is a fair way to select the best climate model. These local calibrations applied to D3L6 and XRD paleosequences highlighted the temperature and precipitation variation throughout the Late Holocene. The next step will be to test our calibrations on pollen and GDGTs records available from the same core. We encourage wider application of this local multi-proxy calibration for a more accurate constraint of these central Asian climatic systems, a crucial improvement to properly model the fluctuations of the Monsoon Line since the Holocene

600 Optimum.

Author contributions. LD conducted the analytical work, LD, SJ, OP, GM designed the study. All the authors contributed to the scientific reflection as well as to the preparation of the manuscript.

Competing interests. The authors declare having no competing interests.

610

References

Aichner, B., Feakins, S. J., Lee, J. E., Herzschuh, U., and Liu, X.: High-resolution leaf wax carbon and hydrogen isotopic record of the late Holocene paleoclimate in arid Central Asia, Climate of the past, 2019.

An, C.-B., Chen, F.-H., and Barton, L.: Holocene environmental changes in Mongolia: a review, Global and Planetary Change, 63, 283–289, 2008.

Arnold, T. W.: Uninformative parameters and model selection using Akaike's Information Criterion, The Journal of Wildlife Management, 74, 1175–1178, 2010.

Atahan, P., Heijnis, H., Dodson, J., Grice, K., Le Metayer, P., Taffs, K., Hembrow, S., Woltering, M., and Zawadzki, A.: Pollen, biomarker and stable isotope evidence of late Quaternary environmental change at Lake McKenzie, southeast Queensland, Journal of Paleolimnology, 53, 139–156, 2015.

Bartlein, P. J., Prentice, I. C., and Webb III, T.: Climatic response surfaces from pollen data for some eastern North American taxa, Journal of Biogeography, p. 35–57, 1986.

Bartlein, P. J., Harrison, S. P., Brewer, S., Connor, S., Davis, B. A. S., Gajewski, K., Guiot, J., Harrison-Prentice, T. I., Henderson, A., and Peyron, O.: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Climate Dynamics, 37, 775–802, 2011.

Bennett, K. D. and Willis, K. J.: Pollen, p. 5–26, Springer, 2002.

Birks, H, H., and Line, J.: The use of rarefaction analysis for estimating palynological richness from Quaternary pollen–analytical data, The Holocene, 2, 1–10, 1992.

Birks, H. J. B., Heiri, O., Seppä, H., and Bjune, A. E.: Strengths and weaknesses of quantitative climate reconstructions based on Late-Quaternary, The Open Ecology Journal, 3, 2010.

Bone, M., Johnson, D., Kelaidis, P., Kintgen, M., Vickerman, L. G., and Gardens, D. B.: Steppes: The Plants and Ecology of the World's Semi-arid Regions, Timber Press, 2015.

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nature Climate Change, 2, 417, 2012.

Brewer, S., Guiot, J., Sánchez-Goñi, M. F., and Klotz, S.: The climate in Europe during the Eemian: a multi-method approach using pollen data, Quaternary Science Reviews, 27, 2303–2315, 2008.

Brun, C.: Anthropogenic indicators in pollen diagrams in eastern France: a critical review, Vegetation history and archaeobotany, 20, 135–142, 2011.

Buckles, L. K., Weijers, J. W. H., Tran, X.-M., Waldron, S., and Sinninghe Damsté, J. S.: Provenance of tetraether membrane lipids in a large temperate lake (Loch Lomond, UK): implications for glycerol dialkyl glycerol tetraether (GDGT)-based palaeothermometry, Biogeosciences, 11, 5539–5563, 2014.

Buckles, L. K., Verschuren, D., Weijers, J. W., Cocquyt, C., Blaauw, M., and Sinninghe Damsté, J. S.: Interannual and (multi-)decadal variability in the sedimentary BIT index of Lake Challa, East Africa, over the past 2200 years, assessment of the precipitation proxy, Climate of the Past, 12, 1243–1262, https://doi.org/http://dx.doi.org/10.5194/cp-12-1243-2016, 2016.

Bunting, M. J. and Hjelle, K. L.: Effect of vegetation data collection strategies on estimates of relevant source area of pollen (RSAP) and relative pollen productivity estimates (relative PPE) for non-arboreal taxa, Vegetation History and Archaeobotany, 19, 365–374, 2010.

Cao, X.-y., Herzschuh, U., Telford, R. J., and Ni, J.: A modern pollen–climate dataset from China and Mongolia: Assessing its potential for climate reconstruction, Review of palaeobotany and palynology, 211, 87–96, 2014.

Chen, F., Holmes, J., Wünnemann, B., and Yu, Z.: Holocene climate variability in arid Asia: Nature and mechanisms, Quaternary International, 194, 1–5, 00013, 2009.

650 Chen, F.-H., Chen, J.-H., Holmes, J., Boomer, I., Austin, P., Gates, J. B., Wang, N.-L., Brooks, S. J., and Zhang, J.-W.: Moisture changes over the last millennium in arid central Asia: a review, synthesis and comparison with monsoon region, Quaternary Science Reviews, 29, 1055–1068, https://doi.org/10.1016/j.quascirev.2010.01.005, 2010.

Chen, J., Chen, F., Feng, S., Huang, W., Liu, J., and Zhou, A.: Hydroclimatic changes in China and surroundings during the Medieval Climate Anomaly and Little Ice Age: spatial patterns and possible mechanisms, Quaternary Science Reviews, 107, 98–111, 2015.

655 Chen, Y., Zheng, F., Chen, S., Liu, H., Phelps, T. J., and Zhang, C.: Branched GDGT production at elevated temperatures in anaerobic soil microcosm incubations, Organic Geochemistry, 117, 12–21, https://doi.org/10.1016/j.orggeochem.2017.11.015, https://linkinghub.elsevier.com/retrieve/pii/S0146638017304254, 2018.

Chevalier, M.: Enabling possibilities to quantify past climate from fossil assemblages at a global scale, Global and Planetary Change, 175, 27–35, https://doi.org/10.1016/j.gloplacha.2019.01.016, http://www.sciencedirect.com/science/article/pii/S0921818118304910, 2019.

660 Coffinet, S.: Validation and application of lipid biomarkers to reconstruct past environmental changes in East Africa, Ph.D. thesis, Paris 6, 00001, 2015.

Coffinet, S., Huguet, A., Williamson, D., Fosse, C., and Derenne, S.: Potential of GDGTs as a temperature proxy along an altitudinal transect at Mount Rungwe (Tanzania), Organic Geochemistry, 68, 82–89, https://doi.org/10.1016/j.orggeochem.2014.01.004, https://linkinghub.elsevier.com/retrieve/pii/S0146638014000059, 2014.

665 Colcord, D. E., Cadieux, S. B., Brassell, S. C., Castañeda, I. S., Pratt, L. M., and White, J. R.: Assessment of branched GDGTs as temperature proxies in sedimentary records from several small lakes in southwestern Greenland, Organic Geochemistry, 82, 33–41, 2015.

Damsté, J. S. S., Hopmans, E. C., Pancost, R. D., Schouten, S., and Geenevasen, J. A.: Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments, Chemical Communications, p. 1683–1684, 2000.

Dang, X., Yang, H., Naafs, B. D. A., Pancost, R. D., and Xie, S.: Evidence of moisture control on the methylation of branched glycerol 670 dialkyl glycerol tetraethers in semi-arid and arid soils, Geochimica et Cosmochimica Acta, 189, 24–36, 2016.

Dang, X., Ding, W., Yang, H., Pancost, R. D., Naafs, B. D. A., Xue, J., Lin, X., Lu, J., and Xie, S.: Different temperature dependence of the bacterial brGDGT isomers in 35 Chinese lake sediments compared to that in soils, Organic Geochemistry, 119, 72–79, 00005, 2018.

Davtian, N., Ménot, G., Bard, E., Poulenard, J., and Podwojewski, P.: Consideration of soil types for the calibration of molecular proxies for soil pH and temperature using global soil datasets and Vietnamese soil profiles, Organic Geochemistry, 101, 140–153, 2016.

675 Davtian, N., Bard, E., Ménot, G., and Fagault, Y.: The importance of mass accuracy in selected ion monitoring analysis of branched and isoprenoid tetraethers, Organic geochemistry, 118, 58–62, 2018.

De Jonge, C., Hopmans, E. C., Stadnitskaia, A., Rijpstra, W. I. C., Hofland, R., Tegelaar, E., and Sinninghe Damsté, J. S.: Identification of novel penta- and hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC–MS2, GC–MS and GC–SMB-MS, Organic Geochemistry, 54, 78–82, https://doi.org/10.1016/j.orggeochem.2012.10.004, 2013.

680 De Jonge, C., Hopmans, E. C., Zell, C. I., Kim, J.-H., Schouten, S., and Sinninghe Damsté, J. S.: Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction, Geochimica et Cosmochimica Acta, 141, 97–112, 2014a.

De Jonge, C., Stadnitskaia, A., Hopmans, E. C., Cherkashov, G., Fedotov, A., and Sinninghe Damsté, J. S.: In situ produced branched glycerol dialkyl glycerol tetraethers in suspended particulate matter from the Yenisei River, Eastern Siberia, Geochimica et Cosmochimica Acta, 685 125, 476–491, 2014b.

31

Dearing Crampton-Flood, E., Tierney, J. E., Peterse, F., Kirkels, F. M. S. A., and Sinninghe Damsté, J. S.: Global soil and peat branched GDGT compilation dataset, https://doi.org/https://doi.org/10.1594/PANGAEA.907818, https://doi.pangaea.de/10.1594/PANGAEA.907818, 2019.

Dearing Crampton-Flood, E., Tierney, J. E., Peterse, F., Kirkels, F. M. S. A., and Sinninghe Damsté, J. S.: BayMBT: A Bayesian calibration model for branched glycerol dialkyl glycerol tetraethers in soils and peats, Geochimica et Cosmochimica Acta, 268, 142–159, https://doi.org/10.1016/j.gca.2019.09.043, 2020.

Demske, D., Heumann, G., Granoszewski, W., Nita, M., Mamakowa, K., Tarasov, P. E., and Oberhänsli, H.: Late glacial and Holocene vegetation and regional climate variability evidenced in high-resolution pollen records from Lake Baikal, Global and Planetary Change, 46, 255–279, 2005.

Deng, L., Jia, G., Jin, C., and Li, S.: Warm season bias of branched GDGT temperature estimates causes underestimation of altitudinal lapse rate, Organic Geochemistry, 96, 11–17, 2016.

Ding, S., Xu, Y., Wang, Y., He, Y., Hou, J., Chen, L., and He, J.-S.: Distribution of branched glycerol dialkyl glycerol tetraethers in surface soils of the Qinghai–Tibetan Plateau: implications of brGDGTs-based proxies in cold and dry regions, Biogeosciences, 12, 3141–3151, 2015.

Ding, S., Schwab, V. F., Ueberschaar, N., Roth, V.-N., Lange, M., Xu, Y., Gleixner, G., and Pohnert, G.: Identification of novel 7-methyl and cyclopentanyl branched glycerol dialkyl glycerol tetraethers in lake sediments, Organic Geochemistry, 102, 52–58, https://doi.org/10.1016/j.orggeochem.2016.09.009, 2016.

Dirghangi, S. S., Pagani, M., Hren, M. T., and Tipple, B. J.: Distribution of glycerol dialkyl glycerol tetraethers in soils from two environmental transects in the USA, Organic Geochemistry, 59, 49–60, https://doi.org/10.1016/j.orggeochem.2013.03.009, 2013.

Dray, S. and Dufour, A.-B.: The ade4 package: implementing the duality diagram for ecologists, Journal of statistical software, 22, 1–20, 2007.

Dregne, H. E.: Soils of arid regions, Elsevier Scientific Pub. Co, 1976.

Felauer, T., Schlütz, F., Murad, W., Mischke, S., and Lehmkuhl, F.: Late Quaternary climate and landscape evolution in arid Central Asia: A multiproxy study of lake archive Bayan Tohomin Nuur, Gobi desert, southern Mongolia, Journal of Asian Earth Sciences, 48, 125–135, 2012.

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas: NEW CLIMATE SURFACES FOR GLOBAL LAND AREAS, International Journal of Climatology, 37, 4302–4315, 2017.

Ge, Y., Li, Y., Bunting, M. J., Li, B., Li, Z., and Wang, J.: Relation between modern pollen rain, vegetation and climate in northern China: Implications for quantitative vegetation reconstruction in a steppe environment, Science of The Total Environment, 586, 25–41, 2017.

Ghosh, R., Paruya, D. K., Acharya, K., Ghorai, N., and Bera, S.: How reliable are non-pollen palynomorphs in tracing vegetation changes and grazing activities? Study from the Darjeeling Himalaya, India, Palaeogeography, Palaeoclimatology, Palaeoecology, 475, 23–40, 00003, 2017.

Grimm, E. C.: CONISS: a FORTRAN 77 program for stratigraphically constrained cluster analysis by the method of incremental sum of squares, Computers & geosciences, 13, 13–35, 1987.

Guiot, J.: Methodology of the last climatic cycle reconstruction in France from pollen data, Palaeogeography, Palaeoclimatology, Palaeoecology, 80, 49–69, 1990.

Guiot, J., Torre, F., Jolly, D., Peyron, O., Boreux, J. J., and Cheddadi, R.: Inverse vegetation modeling by Monte Carlo sampling to reconstruct palaeoclimates under changed precipitation seasonality and CO2 conditions: application to glacial climate in Mediterranean region, Ecological modelling, 127, 119–140, 2000.

725 Gul, H. and Ahmad, R.: Effect of salinity on pollen viability of different canola (Brassica napus L.) cultivars as reflected by the formation of fruits and seeds, Pakistan Journal of Botany, 38, 237, 2006.

Haoran, H. and Weihong, Q.: Identifying the northernmost summer monsoon location in East Asia, Progress in Natural Science, 17, 812–820, 2007.

Haynes, R. J. and Swift, R. S.: Effect of rewetting air-dried soils on pH and accumulation of mineral nitrogen, Journal of Soil Science, 40,
730 341–347, https://doi.org/10.1111/j.1365-2389.1989.tb01278.x, 1989.

Hellman, S., Bunting, M., and Gaillard, M.-J.: Relevant Source Area of Pollen in patchy cultural landscapes and signals of anthropogenic landscape disturbance in the pollen record: A simulation approach, Review of Palaeobotany and Palynology, 153, 245–258, 2009a.

Hellman, S., Gaillard, M.-J., Bunting, J. M., and Mazier, F.: Estimating the Relevant Source Area of Pollen in the past cultural landscapes of southern Sweden — A forward modelling approach, Review of Palaeobotany and Palynology, 153, 259–271, 2009b.

735 Herzschuh, U., Kürschner, H., and Ma, Y.: The surface pollen and relative pollen production of the desert vegetation of the Alashan Plateau, western Inner Mongolia, Chinese Science Bulletin, 48, 1488–1493, 2003.

Herzschuh, U., Tarasov, P., Wünnemann, B., and Hartmann, K.: Holocene vegetation and climate of the Alashan Plateau, NW China, reconstructed from pollen data, Palaeogeography, Palaeoclimatology, Palaeoecology, 211, 1–17, 2004.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land
740 areas, International Journal of Climatology, 25, 1965–1978, 2005.

Hilbig, W.: vegetation of Mongolia, SPB Academic Pubishing, 1995.

Hjelle, K. L.: Relationships between pollen and plants in human-influenced vegetation types using presence-absence data in western Norway, Review of Palaeobotany and Palynology, 99, 1–16, 1997.

Hopmans, E. C., Schouten, S., Pancost, R. D., van der Meer, M. T., and Sinninghe Damsté, J. S.: Analysis of intact tetraether lipids in archaeal
745 cell material and sediments by high performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry, Rapid Communications in Mass Spectrometry, 14, 585–589, 2000.

Hopmans, E. C., Schouten, S., and Damsté, J. S. S.: The effect of improved chromatography on GDGT-based palaeoproxies, Organic Geochemistry, 93, 1–6, 2016.

Huguet, A., Fosse, C., Laggoun-Défarge, F., Delarue, F., and Derenne, S.: Effects of a short-term experimental microclimate warming on the
750 abundance and distribution of branched GDGTs in a French peatland, Geochimica et Cosmochimica Acta, 105, 294–315, 2013.

Huguet, C., Hopmans, E. C., Febo-Ayala, W., Thompson, D. H., Sinninghe Damste, J. S., and Schouten, S.: An improved method to determine the absolute abundance of glycerol dibiphytanyl glycerol tetraether lipids, Organic Geochemistry, 37, 1036–1041, 00301, 2006.

Jackson, S. T. and Williams, J. W.: Modern analogs in Quaternary paleoecology: here today, gone yesterday, gone tomorrow?, Annual Review of Earth and Planetary Sciences, 32, 2004.

755 Juggins, S. and Juggins, M. S.: Package 'rioja', RCRAN, 2019.

Kaufman, D., McKay, N., Routson, C., Erb, M., Davis, B., Heiri, O., Jaccard, S., Tierney, J., Dätwyler, C., Axford, Y., Brussel, T., Cartapanis, O., Chase, B., Dawson, A., de Vernal, A., Engels, S., Jonkers, L., Marsicek, J., Moffa-Sánchez, P., Morrill, C., Orsi, A., Rehfeld, K., Saunders, K., Sommer, P. S., Thomas, E., Tonello, M., Tóth, M., Vachula, R., Andreev, A., Bertrand, S., Biskaborn, B., Bringué, M., Brooks, S., Caniupán, M., Chevalier, M., Cwynar, L., Emile-Geay, J., Fegyveresi, J., Feurdean, A., Finsinger, W., Fortin, M.-C., Foster, L.,

760   Fox, M., Gajewski, K., Grosjean, M., Hausmann, S., Heinrichs, M., Holmes, N., Ilyashuk, B., Ilyashuk, E., Juggins, S., Khider, D., Koinig,
K., Langdon, P., Larocque-Tobler, I., Li, J., Lotter, A., Luoto, T., Mackay, A., Magyari, E., Malevich, S., Mark, B., Massaferro, J., Montade,
V., Nazarova, L., Novenko, E., Pařil, P., Pearson, E., Peros, M., Pienitz, R., Płóciennik, M., Porinchu, D., Potito, A., Rees, A., Reinemann,
S., Roberts, S., Rolland, N., Salonen, S., Self, A., Seppä, H., Shala, S., St-Jacques, J.-M., Stenni, B., Syrykh, L., Tarrats, P., Taylor, K.,
van den Bos, V., Velle, G., Wahl, E., Walker, I., Wilmshurst, J., Zhang, E., and Zhilich, S.: A global database of Holocene paleotemperature
765   records, Scientific Data, 7, 115, https://doi.org/10.1038/s41597-020-0445-3, http://www.nature.com/articles/s41597-020-0445-3, 2020.

Klinge, M. and Sauer, D.: Spatial pattern of Late Glacial and Holocene climatic and environmental development in Western Mongolia - A
critical review and synthesis, Quaternary Science Reviews, 210, 26–50, 00000, 2019.

Klinge, M., Dulamsuren, C., Erasmi, S., Karger, D. N., and Hauck, M.: Climate effects on vegetation vitality at the treeline of boreal forests
of Mongolia, Biogeosciences, 15, 1319–1333, 00003, 2018.

770   Klotz, S., Guiot, J., and Mosbrugger, V.: Continental European Eemian and early Würmian climate evolution: comparing signals using
different quantitative reconstruction approaches based on pollen, Global and Planetary Change, 36, 277–294, 2003.

Klotz, S., Müller, U., Mosbrugger, V., de Beaulieu, J.-L., and Reille, M.: Eemian to early Würmian climate dynamics: history and pattern of
changes in Central Europe, Palaeogeography, Palaeoclimatology, Palaeoecology, 211, 107–126, 2004.

Knappy, C. S., Nunn, C. E. M., Morgan, H. W., and Keely, B. J.: The major lipid cores of the archaeon Ignisphaera aggregans: im-
775   plications for the phylogeny and biosynthesis of glycerol monoalkyl glycerol tetraether isoprenoid lipids, Extremophiles, 15, 517,
https://doi.org/10.1007/s00792-011-0382-3, 2011.

Kröpelin, S., Verschuren, D., Lézine, A.-M., Eggermont, H., Cocquyt, C., Francus, P., Cazet, J.-P., Fagot, M., Rumes, B., and Russell, J. M.:
Climate-driven ecosystem succession in the Sahara: the past 6000 years, science, 320, 765–768, https://doi.org/10/chj4rg, 00467, 2008.

Kusch, S., Winterfeld, M., Mollenhauer, G., Höfle, S. T., Schirrmeister, L., Schwamborn, G., and Rethemeyer, J.: Glycerol dialkyl glycerol
780   tetraethers (GDGTs) in high latitude Siberian permafrost: Diversity, environmental controls, and implications for proxy applications,
Organic Geochemistry, 136, 103 888, 2019.

Kühl, N., Gebhardt, C., Litt, T., and Hense, A.: Probability density functions as botanical-climatological transfer functions for climate
reconstruction, Quaternary Research, 58, 381–392, 2002.

Laldinthar, R. and Dkhar, M. S.: Relationship between soil bacterial population and various physico-chemical properties at two broadleaved
785   forest stands of meghalaya differing in altitudes, Transcriptomics, 3, 1–7, 2015.

Lebreton, V., Messager, E., Marquer, L., and Renault-Miskovsky, J.: A neotaphonomic experiment in pollen oxidation and its implications
for archaeopalynology, Review of Palaeobotany and Palynology, 162, 29–38, 2010.

Legendre, P.: Spatial Autocorrelation: Trouble or New Paradigm?, Ecology, 74, 1659–1673, https://doi.org/10.2307/1939924, https://
esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1939924, 1993.

790   Lehmkuhl, F., Hilgers, A., Fries, S., Hülle, D., Schlütz, F., Shumilovskikh, L., Felauer, T., and Protze, J.: Holocene geomorphological
processes and soil development as indicator for environmental change around Karakorum, Upper Orkhon Valley (Central Mongolia),
Catena, 87, 31–44, 2011.

Lei, Y., Yang, H., Dang, X., Zhao, S., and Xie, S.: Absence of a significant bias towards summer temperature in branched tetraether-based
paleothermometer at two soil sites with contrasting temperature seasonality, Organic Geochemistry, 94, 83–94, 2016.

795   Li, Q., Wu, H., Yu, Y., Sun, A., and Luo, Y.: Quantifying regional vegetation changes in China during three contrasting temperature intervals
since the last glacial maximum, Journal of Asian Earth Sciences, https://doi.org/10.1016/j.jseaes.2018.10.013, 00000, 2018.

Li, Y., Bunting, M. J., Xu, Q., Jiang, S., Ding, W., and Hun, L.: Pollen–vegetation–climate relationships in some desert and desert-steppe communities in northern China, The Holocene, 21, 997–1010, https://doi.org/10/brh6qc, 00039, 2011.

Li, Y.-C., Author, Q.-H. X. C., Yang, X.-L., Chen, H., and Lu, X.-M.: Pollen-vegetation relationship and pollen preservation on the North-
800   eastern Qinghai-Tibetan Plateau, Grana, 44, 160–171, 2005.

Liang, J., Russell, J. M., Xie, H., Lupien, R. L., Si, G., Wang, J., Hou, J., and Zhang, G.: Vegetation effects on tem-
perature calibrations of branched glycerol dialkyl glycerol tetraether (brGDGTs) in soils, Organic Geochemistry, 127, 1–11,
https://doi.org/10.1016/j.orggeochem.2018.10.010, 2019.

Lin, Y.-T., Whitman, W. B., Coleman, D. C., Shi, S.-Y., Tang, S.-L., and Chiu, C.-Y.: Changes of soil bacterial communities in bamboo
805   plantations at different elevations, FEMS microbiology ecology, 91, 2015.

Liu, H., Wang, Y., Tian, Y., Zhu, J., and Wang, H.: Climatic and anthropogenic control of surface pollen assemblages in East Asian steppes,
Review of Palaeobotany and Palynology, 138, 281–289, 2006.

Liu, X.-L., Lipp, J. S., Schröder, J. M., Summons, R. E., and Hinrichs, K.-U.: Isoprenoid glycerol dialkanol diethers: A series of novel
archaeal lipids in marine sediments, Organic Geochemistry, 43, 50–55, https://doi.org/10/bzfj97, 00043, 2012a.

810   Liu, X.-L., Lipp, J. S., Simpson, J. H., Lin, Y.-S., Summons, R. E., and Hinrichs, K.-U.: Mono- and dihydroxyl glycerol dibiphytanyl glycerol
tetraethers in marine sediments: Identification of both core and intact polar lipid forms, Geochimica et Cosmochimica Acta, 89, 102–115,
https://doi.org/10/f323xm, 00050, 2012b.

Ma, Y., Liu, K.-b., Feng, Z., Sang, Y., Wang, W., and Sun, A.: A survey of modern pollen and vegetation along a south–north transect in
Mongolia, Journal of Biogeography, 35, 1512–1532, 2008.

815   Marsicek, J., Shuman, B. N., Bartlein, P. J., Shafer, S. L., and Brewer, S.: Reconciling divergent trends and millennial variations in Holocene
temperatures, Nature, 554, 92–96, https://doi.org/10.1038/nature25464, https://www.nature.com/articles/nature25464, number: 7690 Pub-
lisher: Nature Publishing Group, 2018.

Martin, C., Menot, G., Thouveny, N., Peyron, O., Andrieu-Ponel, V., Montade, V., Davtian, N., Reille, M., and Bard, E.: Early Holocene
Thermal Maximum recorded by branched tetraethers and pollen in Western Europe (Massif Central, France), Quaternary Science Reviews,
820   228, 106 109, 2020.

Martínez-Sosa, P., Tierney, J. E., and Meredith, L. K.: Controlled lacustrine microcosms show a brGDGT response to environmental perturba-
tions, Organic Geochemistry, 145, 104 041, https://doi.org/10.1016/j.orggeochem.2020.104041, https://linkinghub.elsevier.com/retrieve/
pii/S0146638020300760, 2020.

Masson-Delmotte, V.: Global Warming of 1.5 degree C: An IPCC Special Report on the Impacts of Global Warming of 1.5 degree C Above
825   Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the
Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty, World Meteorological Organization, 2018.

Menges, J., Huguet, C., Alcaniz, J. M., Fietz, S., Sachse, D., and Rosell-Melé, A.: Influence of water availability in the dis-
tributions of branched glycerol dialkyl glycerol tetraether in soils of the Iberian Peninsula, Biogeosciences, 11, 2571–2581,
https://doi.org/https://doi.org/10.5194/bg-11-2571-2014, 2014.

830   Naafs, B., Gallego-Sala, A., Inglis, G., and Pancost, R.: Refining the global branched glycerol dialkyl glycerol tetraether (brGDGT) soil
temperature calibration, Organic Geochemistry, 106, 48–56, 2017a.

Naafs, B. D. A.: Global biomarker (GDGT) database for peatlands, https://doi.org/https://doi.org/10.1594/PANGAEA.883765, 2017.

Naafs, B. D. A., Inglis, G. N., Zheng, Y., Amesbury, M. J., Biester, H., Bindler, R., Blewett, J., Burrows, M. A., Del Castillo Torres, D., and Chambers, F. M.: Introducing Global Peat-Specific Temperature and pH Calibrations Based on brGDGT Bacterial Lipids, Geochimica et Cosmochimica Acta, 208, 285–301, https://doi.org/10/f99n9b, 2017b.

Naafs, B. D. A., McCormick, D., Inglis, G. N., and Pancost, R. D.: Archaeal and bacterial H-GDGTs are abundant in peat and their relative abundance is positively correlated with temperature, Geochimica et Cosmochimica Acta, 227, 156–170, 00004, 2018.

Ohlwein, C. and Wahl, E. R.: Review of probabilistic pollen-climate transfer methods, Quaternary Science Reviews, 31, 17–29, 2012.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., et al.: Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity, BioScience, 51, 933–938, 2001.

Overpeck, J. T., Webb, T., and Prentice, I. C.: Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs, Quaternary Research, 23, 87–108, 1985.

Pearson, E. J., Juggins, S., Talbot, H. M., Weckström, J., Rosén, P., Ryves, D. B., Roberts, S. J., and Schmidt, R.: A lacustrine GDGT-temperature calibration from the Scandinavian Arctic to Antarctic: Renewed potential for the application of GDGT-paleothermometry in lakes, Geochimica et Cosmochimica Acta, 75, 6225–6238, 2011.

Peterse, F., van der Meer, J., Schouten, S., Weijers, J. W., Fierer, N., Jackson, R. B., Kim, J.-H., and Damsté, J. S. S.: Revised calibration of the MBT–CBT paleotemperature proxy based on branched tetraether membrane lipids in surface soils, Geochimica et Cosmochimica Acta, 96, 215–229, 2012.

Peyron, O., Guiot, J., Cheddadi, R., Tarasov, P., Reille, M., de Beaulieu, J.-L., Bottema, S., and Andrieu, V.: Climatic reconstruction in Europe for 18,000 yr BP from pollen data, Quaternary research, 49, 183–196, 1998.

Peyron, O., Magny, M., Goring, S., Joannin, S., De Beaulieu, J.-L., Bruggiapaglia, E., Sadori, L., Garfi, G., Kouli, K., and Ioakim, C.: Contrasting patterns of climatic changes during the Holocene across the Italian Peninsula reconstructed from pollen data, Climate of the Past, 9, 1233–1252, 2013.

Peyron, O., Combourieu-Nebout, N., Brayshaw, D., Goring, S., Andrieu-Ponel, V., Desprat, S., Fletcher, W., Gambin, B., Ioakim, C., Joannin, S., et al.: Precipitation changes in the Mediterranean basin during the Holocene from terrestrial and marine pollen records: a model–data comparison, Climate of the Past, 13, 249–265, https://doi.org/https://doi.org/10.5194/cp-13-249-2017, 2017.

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, Climate Dynamics, 12, 185–194, 1996.

Prentice, I. C.: Pollen representation, source area, and basin size: toward a unified theory of pollen analysis, Quaternary Research, 23, 76–86, 1985.

Pérez-Angel, L., Sepúlveda, J., Rajagopalan, B., Montes, C., Molnar, P., Snell, K., Dildar, N., and González-Arango, C.: An improved regional branched GDGT- based soil temperature calibration for the Tropical Andes of Colombia: Towards a global calibration for the tropics., Goldschmidt Abstract, 2019.

Reddy, P. R. and Goss, J. A.: Effect of salinity on pollen I. Pollen viability as altered by increasing osmotic pressure with NaCl, MgCl2, and CaCl2, American Journal of Botany, 58, 721–725, 1971.

Rudaya, N., Tarasov, P., Dorofeyuk, N., Solovieva, N., Kalugin, I., Andreev, A., Daryin, A., Diekmann, B., Riedel, F., Tserendash, N., et al.: Holocene Environments and Climate in the Mongolian Altai Reconstructed from the Hoton-Nur Pollen and Diatom Records: A Step towards Better Understanding Climate Dynamics in Central Asia, Quaternary Science Reviews, 28, 540–554, https://doi.org/10/fjsk99, 2009.

Räsänen, S., Hicks, S., and Odgaard, B. V.: Pollen deposition in mosses and in a modified 'Tauber trap' from Hailuoto, Finland: what exactly do the mosses record?, Review of Palaeobotany and Palynology, 129, 103–116, 2004.

Salonen, J. S., Luoto, M., Alenius, T., Heikkilä, M., Seppä, H., Telford, R. J., and Birks, H. J. B.: Reconstructing palaeoclimatic variables from fossil pollen using boosted regression trees: comparison and synthesis with other quantitative reconstruction methods, Quaternary Science Reviews, 88, 69–81, https://doi.org/10.1016/j.quascirev.2014.01.011, http://www.sciencedirect.com/science/article/pii/S0277379114000183, 2014.

Salonen, J. S., Korpela, M., Williams, J. W., and Luoto, M.: Machine-learning based reconstructions of primary and secondary climate variables from North American and European fossil pollen data, Scientific reports, 9, 1–13, 2019.

Schouten, S., Rijpstra, W. I. C., Durisch-Kaiser, E., Schubert, C. J., and Sinninghe Damsté, J. S.: Distribution of glycerol dialkyl glycerol tetraether lipids in the water column of Lake Tanganyika, Advances in Organic Geochemistry 2011: Proceedings of the 25th International Meeting on Organic Geochemistry, 53, 34–37, 2012.

Seppä, H., Birks, H., Odland, A., Poska, A., and Veski, S.: A modern pollen–climate calibration set from northern Europe: developing and testing a tool for palaeoclimatological reconstructions, Journal of Biogeography, 31, 251–267, 2004.

Sharkhuu, N.: Recent changes in the permafrost of Mongolia, in: Proceedings of the 8th International Conference on Permafrost, 21-25 July 2003, Zurich, Switzerland, p. 1029–1034, AA Balkema, Lisse, the Netherlands, 2003.

Shen, C., Xiong, J., Zhang, H., Feng, Y., Lin, X., Li, X., Liang, W., and Chu, H.: Soil pH drives the spatial distribution of bacterial communities along elevation on Changbai Mountain, Soil Biology and Biochemistry, 57, 204–211, 2013.

Sinninghe Damsté, J. S.: Spatial heterogeneity of sources of branched tetraethers in shelf systems: The geochemistry of tetraethers in the Berau River delta (Kalimantan, Indonesia), Geochimica et Cosmochimica Acta, 186, 13–31, 2016.

Stevens, G. C. and Fox, J. F.: The causes of treeline, Annual review of ecology and systematics, 22, 177–191, 1991.

Sun, Q., Chu, G., Liu, M., Xie, M., Li, S., Ling, Y., Wang, X., Shi, L., Jia, G., and Lü, H.: Distributions and temperature dependence of branched glycerol dialkyl glycerol tetraethers in recent lacustrine sediments from China and Nepal, Journal of Geophysical Research: Biogeosciences, 116, 2011.

Sun, W., Zhao, S., Pei, H., and Yang, H.: The Coupled Evolution of Mid- to Late Holocene Temperature and Moisture in the Southeast Qaidam Basin, Chemical Geology, 528, 119 282, https://doi.org/10.1016/j.chemgeo.2019.119282, 2019.

Symonds, M. R. and Moussalli, A.: A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion, Behavioral Ecology and Sociobiology, 65, 13–21, 2011.

Telford, R. and Birks, H.: Evaluation of transfer functions in spatially structured environments, Quaternary Science Reviews, 28, 1309–1316, 2009.

Telford, R. J. and Birks, H. J. B.: The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance, Quaternary Science Reviews, 24, 2173–2179, https://doi.org/10.1016/j.quascirev.2005.05.001, http://www.sciencedirect.com/science/article/pii/S027737910500168X, 2005.

Telford, R. J. and Birks, H. J. B.: A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages, Quaternary Science Reviews, 30, 1272–1278, publisher: Elsevier, 2011.

Ter Braak, C. J. and Juggins, S.: Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages, in: Twelfth International Diatom Symposium, p. 485–502, Springer, 1993.

Ter Braak, C. J. F., Juggins, S., Birks, H. J. B., and Van der Voet, H.: Weighted averaging partial least squares regression (WA-PLS): definition and comparison with other methods for species-environment calibration, chap. 25, p. 525–560, Elsevier, 1993.

Climate
of the Past
Discussions

Thomas, E. K., Clemens, S. C., Sun, Y., Huang, Y., Prell, W., Chen, G., Liu, Z., and Loomis, S.: Midlatitude land surface temperature impacts the timing and structure of glacial maxima, Geophysical Research Letters, 44, 984–992, 2017.

Tian, F., Herzschuh, U., Mischke, S., and Schlütz, F.: What drives the recent intensified vegetation degradation in Mongolia–Climate change or human activity?, The Holocene, 24, 1206–1215, 2014.

Tierney, J. E. and Russell, J. M.: Distributions of branched GDGTs in a tropical lake system: Implications for lacustrine application of the MBT/CBT paleoproxy, Organic Geochemistry, 40, 1032–1036, https://doi.org/10.1016/j.orggeochem.2009.04.014, 2009.

Unkelbach, J., Kashima, K., Enters, D., Dulamsuren, C., Punsalpaamuu, G., and Behling, H.: Late Holocene (Meghalayan) Palaeoenvironmental Evolution Inferred from Multi-Proxy-Studies of Lacustrine Sediments from the Dayan Nuur Region of Mongolia, Palaeogeography, Palaeoclimatology, Palaeoecology, 530, 1–14, https://doi.org/10.1016/j.palaeo.2019.05.021, 2019.

Wagner, B., Vogel, H., Francke, A., Friedrich, T., Donders, T., Lacey, J., Leng, M., Regattieri, E., Sadori, L., Wilke, T., Zanchetta, G., Albrecht, C., Bertini, A., Combourieu-Nebout, N., Cvetkoska, A., Giaccio, B., Grazhdani, A., Hauffe, T., Holtvoeth, J., and Zhang, X.: Mediterranean winter rainfall in phase with African monsoons during the past 1.36 million years, Nature, 573, 256–260, https://doi.org/10.1038/s41586-019-1529-0, 2019.

Wang, H., Liu, W., and Lu, H.: Appraisal of branched glycerol dialkyl glycerol tetraether-based indices for North China, Organic Geochemistry, 98, 118–130, 2016.

Wang, J.-T., Cao, P., Hu, H.-W., Li, J., Han, L.-L., Zhang, L.-M., Zheng, Y.-M., and He, J.-Z.: Altitudinal distribution patterns of soil bacterial and archaeal communities along Mt. Shegyla on the Tibetan Plateau, Microbial ecology, 69, 135–145, 2015.

Wang, M., Zheng, Z., Zong, Y., Man, M., and Tian, L.: Distributions of soil branched glycerol dialkyl glycerol tetraethers from different climate regions of China, Scientific reports, 9, 1–8, 2019.

Wang, M., Yang, H., Zheng, Z., and Tian, L.: Altitudinal climatic index changes in subtropical China indicated from branched glycerol dialkyl glycerol tetraethers proxies, Chemical Geology, p. 119579, 2020.

Wang, W. and Feng, Z.: Holocene moisture evolution across the Mongolian Plateau and its surrounding areas: A synthesis of climatic records, Earth-Science Reviews, 122, 38–57, 2013.

Wang, Y., Liu, X., and Herzschuh, U.: Asynchronous evolution of the Indian and East Asian Summer Monsoon indicated by Holocene moisture patterns in monsoonal central Asia, Earth-Science Reviews, 103, 135–153, 2010.

Watson, B. I., Williams, J. W., Russell, J. M., Jackson, S. T., Shane, L., and Lowell, T. V.: Temperature variations in the southern Great Lakes during the last deglaciation: Comparison between pollen and GDGT proxies, Quaternary Science Reviews, 182, 78–92, 2018.

Weijers, J. W., Schouten, S., van der Linden, M., van Geel, B., and Sinninghe Damsté, J. S.: Water table related variations in the abundance of intact archaeal membrane lipids in a Swedish peat bog, FEMS Microbiology Letters, 239, 51–56, 2004.

Weijers, J. W., Schouten, S., Sluijs, A., Brinkhuis, H., and Damsté, J. S. S.: Warm arctic continents during the Palaeocene–Eocene thermal maximum, Earth and Planetary Science Letters, 261, 230–238, 2007a.

Weijers, J. W., Schouten, S., van den Donker, J. C., Hopmans, E. C., and Damsté, J. S. S.: Environmental controls on bacterial tetraether membrane lipid distribution in soils, Geochimica et Cosmochimica Acta, 71, 703–713, 00503, 2007b.

Wen, R., Xiao, J., Ma, Y., Feng, Z., Li, Y., and Xu, Q.: Pollen–climate transfer functions intended for temperate eastern Asia, Quaternary International, 311, 3–11, 2013.

Wesche, K., Ambarlı, D., Kamp, J., Török, P., Treiber, J., and Dengler, J.: The Palaearctic steppe biome: a new synthesis, Biodiversity and conservation, 25, 2197–2231, 2016.

Wickham, H.: ggplot2: elegant graphics for data analysis, Springer, 2016.

Wu, D., Cao, J., Jia, G., Guo, H., Shi, F., Zhang, X., and Rao, Z.: Peat brGDGTs-Based Holocene Temperature History of the Altai Mountains in Arid Central Asia, Palaeogeography, Palaeoclimatology, Palaeoecology, 538, 109 464, https://doi.org/10.1016/j.palaeo.2019.109464, 2020.

Xie, S., Pancost, R. D., Chen, L., Evershed, R. P., Yang, H., Zhang, K., Huang, J., and Xu, Y.: Microbial lipid records of highly alkaline deposits and enhanced aridity associated with significant uplift of the Tibetan Plateau in the Late Miocene, Geology, 40, 291–294, 00080, 2012.

Xie, W., Zhang, C., and Ma, C.: Temporal variation in community structure and lipid composition of Thaumarchaeota from subtropical soil: Insight into proposing a new soil pH proxy, Organic geochemistry, 83, 54–64, 2015.

Xu, Q.-h., Li, Y.-c., Tian, F., Cao, X.-y., and Yang, X.-l.: Pollen assemblages of tauber traps and surface soil samples in steppe areas of China and their relationships with vegetation and climate, Review of Palaeobotany and Palynology, 153, 86–101, 2009.

Yang, H., Pancost, R. D., Dang, X., Zhou, X., Evershed, R. P., Xiao, G., Tang, C., Gao, L., Guo, Z., and Xie, S.: Correlations between microbial tetraether lipids and environmental variables in Chinese soils: Optimizing the paleo-reconstructions in semi-arid and arid regions, Geochimica et Cosmochimica Acta, 126, 49–69, 2014.

Yang, H., Lü, X., Ding, W., Lei, Y., Dang, X., and Xie, S.: The 6-methyl branched tetraethers significantly affect the performance of the methylation index (MBT') in soils from an altitudinal transect at Mount Shennongjia, Organic Geochemistry, 82, 42–53, https://doi.org/10.1016/j.orggeochem.2015.02.003, 2015.

Yang, H., Xiao, W., Słowakiewicz, M., Ding, W., Ayari, A., Dang, X., and Pei, H.: Depth-dependent variation of archaeal ether lipids along soil and peat profiles from southern China: Implications for the use of isoprenoidal GDGTs as environmental tracers, Organic Geochemistry, 128, 42–56, 00000, 2019.

Zang, J., Lei, Y., and Yang, H.: Distribution of glycerol ethers in Turpan soils: implications for use of GDGT-based proxies in hot and dry regions, Frontiers of Earth Science, 12, 862–876, 00000, 2018.

Zhang, P., Cheng, H., Edwards, R. L., Chen, F., Wang, Y., Yang, X., Liu, J., Tan, M., Wang, X., and Liu, J.: A test of climate, sun, and culture relationships from an 1810-year Chinese cave record, science, 322, 940–942, 2008.

Zheng, Y., Yu, G., Wang, S. M., Xue, B., Zhuo, D. Q., Zeng, X. M., and Liu, H. Q.: Simulation of paleoclimate over East Asia at six ka BP and twenty one ka BP by a regional climate model, Climate Dynamics, 23, 513–529, 00002, 2004.

Zheng, Y., Li, Q., Wang, Z., Naafs, B. D. A., Yu, X., and Pancost, R. D.: Peatland GDGT records of Holocene climatic and biogeochemical responses to the Asian Monsoon, Organic Geochemistry, 87, 86–95, https://doi.org/10.1016/j.orggeochem.2015.07.012, 2015.

Zheng, Z., Huang, K., Xu, Q., Lu, H., Cheddadi, R., Luo, Y., Beaudouin, C., Luo, C., Zheng, Y., and Li, C.: Comparison of climatic threshold of geographical distribution between dominant plants and surface pollen in China, Science in China Series D: Earth Sciences, 51, 1107–1120, 2008.

Zheng, Z., Wei, J., Huang, K., Xu, Q., Lu, H., Tarasov, P., Luo, C., Beaudouin, C., Deng, Y., Pan, A., et al.: East Asian pollen database: modern pollen distribution and its quantitative relationship with vegetation and climate, Journal of biogeography, 41, 1819–1832, 2014.