

UNIVERSITÉ DE LAUSANNE

DOCTORAL THESIS

---

# Software and Numerical Tools for Paleoclimate Analysis

---

*Author:*

Philipp S. SOMMER

*Supervisor:*

Dr. Basil A. S. Davis

*A thesis submitted in fulfillment of the requirements  
for the degree of Doctor of Science*

*in the*

The Davis Group  
Institute of Earth Surface Dynamics (IDYST)

October 9, 2019



# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Learning from the Past – Why we study paleo-climates . . . . .	2
1.2.1 Proxy Data from the Holocene . . . . .	3
1.2.2 Model Simulations of the Holocene . . . . .	3
1.3 Software for Paleoclimatology . . . . .	4
1.3.1 Sofware for Proxy Data Analysis, Visualization and Distribution	5
1.3.2 The Development of Computational Climate Model Analysis .	5
1.3.3 Methods and Workflows in Open-Source Software Development	7
Version Control . . . . .	7
Automated Tests, Test Coverage and Continuous Integration .	8
Automated Documentation . . . . .	9
Distribution through package managers and virtual environments . . . . .	9
1.4 Challenges tackled in this thesis . . . . .	10
References . . . . .	11
<b>Part I New Software Tools for Paleoclimate Analysis</b>	<b>21</b>
<b>2 Psyplot: A flexible framework for interactive data analysis</b>	<b>23</b>
2.1 Summary . . . . .	23
2.2 Introduction . . . . .	23
2.3 The psyplot framework . . . . .	24
2.3.1 Data model . . . . .	24
Psyplot and xarray . . . . .	24
Psyplot core structure . . . . .	26
2.3.2 Psyplot plugins . . . . .	26
psy-simple: The psyplot plugin for simple visualizations . . .	27
psy-maps: The psyplot plugin for visualizations on a map . .	27
psy-reg: The psyplot plugin for visualizing and calculating re-	
gression plots . . . . .	27
psy-strat: A psyplot plugin for stratigraphic plots . . . . .	28
2.3.3 The psyplot Graphical User Interface . . . . .	28
Console . . . . .	28
Help explorer . . . . .	30
Plot creator . . . . .	30
Project content . . . . .	30

Formatoptions . . . . .	30
Figures and plots . . . . .	31
2.4 Conclusions . . . . .	31
2.5 Outlook . . . . .	31
2.A Example call of a plot method . . . . .	33
2.B psy-simple plot methods . . . . .	34
2.C psy-maps plot methods . . . . .	35
2.D psy-reg plot methods . . . . .	35
2.E psy-strat plot methods . . . . .	36
References . . . . .	36
<b>3 Straditize: A digitization software for pollen diagrams</b>	<b>41</b>
<b>4 The EMPD and POLNET web-interfaces</b>	<b>43</b>
4.1 Summary . . . . .	43
4.2 The EMPD web framework . . . . .	44
4.2.1 The EMPD viewer . . . . .	44
The Web Interface . . . . .	45
Implementation details . . . . .	46
4.2.2 The EMPD2 data repository . . . . .	46
4.2.3 The EMPD-admin . . . . .	46
Implementation details . . . . .	49
4.2.4 Distribution of the tools . . . . .	49
4.3 The POLNET viewer . . . . .	49
References . . . . .	51
<b>Part II Numerical Analysis of Paleoclimate Data</b>	<b>53</b>
<b>5 GWGEN v1.0: A globally calibrated scheme for generating daily meteorology from monthly statistics</b>	<b>55</b>
5.1 Introduction . . . . .	55
5.2 Model description . . . . .	57
5.3 Model development . . . . .	58
5.3.1 Development of a global weather station database . . . . .	60
5.3.2 Parameterization . . . . .	61
Precipitation occurrence . . . . .	61
Precipitation amount . . . . .	62
Temperature . . . . .	64
Cloud fraction . . . . .	67
Wind speed . . . . .	69
Cross correlation . . . . .	71
5.3.3 Model Evaluation . . . . .	72
5.3.4 Bias correction . . . . .	73
5.3.5 Sensitivity analysis . . . . .	75
5.4 Limitations . . . . .	76
5.5 Discussion and Outlook . . . . .	77
5.6 Conclusions . . . . .	78
5.7 Code availability . . . . .	78
5.A Supplementary material . . . . .	80
5.A.1 Sensitivity analysis . . . . .	80

References . . . . .	80
<b>6 pyleogrid: An Ensemble method for Gridding Paleo Proxy Climates</b>	<b>87</b>
6.1 Introduction . . . . .	87
6.2 Data . . . . .	88
6.2.1 Pollen database . . . . .	88
6.2.2 Sample site: Tigalmamine . . . . .	88
6.2.3 Site-based holocene temperature estimates . . . . .	90
6.2.4 Age uncertainties . . . . .	91
6.3 Method . . . . .	93
6.3.1 Constrained age sampling . . . . .	93
The intuitive approach . . . . .	93
The random sorting approach . . . . .	96
The Gibbs sampling approach . . . . .	97
6.3.2 Temperature sampling . . . . .	99
6.3.3 Gridding . . . . .	101
6.3.4 Implementation . . . . .	103
6.4 Results . . . . .	104
6.4.1 Site-based realized climate reconstruction: a use-case . . . . .	104
6.4.2 Gridded summer temperature . . . . .	104
Deterministic vs. ensemble approach . . . . .	104
Temperature sampling parameters . . . . .	108
Uncertainties for spatial extrapolation . . . . .	110
6.5 Discussion . . . . .	110
6.6 Conclusions . . . . .	111
6.A Estimated age uncertainties . . . . .	113
6.B Example of generated age distributions . . . . .	113
6.C Maps of uncertainties . . . . .	113
References . . . . .	116
<b>7 Conclusions</b>	<b>121</b>
<b>Appendices</b>	<b>125</b>
<b>Todo list</b>	<b>125</b>
<b>Acronyms</b>	<b>127</b>
<b>A Computing climate-smart urban land use with the Integrated Urban Complexity model (IUCm 1.0)</b>	<b>129</b>
<b>B Publications and Conference contributions</b>	<b>131</b>
B.0.1 Peer-reviewed . . . . .	131
B.0.2 Conference contributions . . . . .	131
<b>C New Software Tools - An Overview</b>	<b>133</b>
C.1 Main packages . . . . .	133
C.2 Other packages . . . . .	133



## Part I

# New Software Tools for Paleoclimate Analysis



## Part II

# Numerical Analysis of Paleoclimate Data





## Chapter 6

# pyleogrid

*An Ensemble method for Gridding Paleo Proxy Climates*

### 6.1 Introduction

Write gridding introduction

- Why gridding
  - Data-Model-Intercomparisons)
  - Easier to handle
  - Energy balance, compatibility with models, grid cell (Area) averages
  - Stability of observation network through time
  - Not just spatial grid, also regular timestep (problems with pseudo gridding, including time ‘windows’ or ‘slices’)
  - Understanding past climates, different forcings – independent of models (difficult from point-cloud)
  - Filling the gaps
  - Spatial scales (Samarthein chironomids etc)
- Importance of uncertainties
  - Necessary for data comparisons, interpretation skill of the data
  - Proxy climate reconstruction uncertainties are higher (inverse modelling, not always properly defined (MAT)) than for instrumental data
  - Age uncertainties can be high (centennial to multi-millennial)
- Why Tps (can also use something else, but has been used before)
  - Extrapolation to the gaps (Previous work by mauri et al and davis)
  - Pseudo-gridding (marcott, Marsicek 2018, Margo, 2009, Bartlein et al ?2013) has holes
  - Data assimilation (pages2k? Need to look the paper up again) – depends on model
  - Bayesian data assimilation (Weitzel 2019) – depends on model
    - Does not require interpolation of time (Marsicek 2018)
- Other
- Introducing constraints; climate, training set size, spatial coverage etc

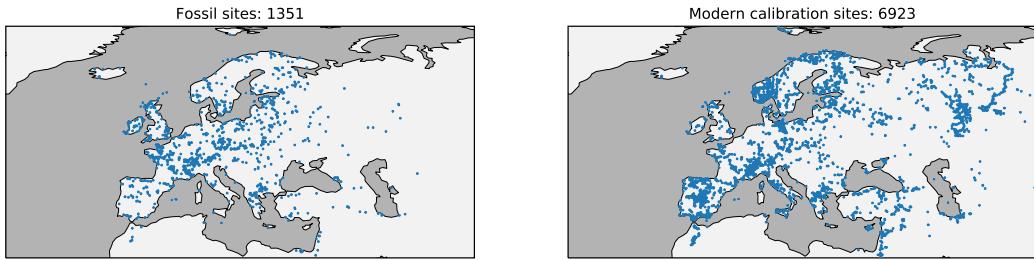


FIGURE 6.1: Site locations of the (left) fossil and (right) modern pollen database.

## 6.2 Data

The ensemble based gridding method is adapted to paleo-climates. In this study, we describe the method using a large set of western Eurasian fossil pollen assemblages that have been transformed to summer (June, July and August) (JJA) temperatures. We focus on pollen data because it is the spatially most widely available proxy during the Holocene, but it is important to mention that the reconstruction method is agnostic to the climate proxy, because it does not explicitly use the pollen assemblages but rather alters the standard climate reconstruction method under the aspect of its methodological uncertainties. As such, the following sections describe the fossil and modern pollen database for this use case (section 6.2.1) and the associated uncertainties of the temperature reconstruction method (section 6.2.3) and the dating of the fossil pollen samples (section 6.2.4).

### 6.2.1 Pollen database

The source data for this study is a subset of the latest development version of the POLNET database, a northern hemispheric, sub-tropical collection of pollen assemblages (Davis and Kaplan, 2017; Sommer et al., 2019). The purpose of this database is to generate the source for large-scale climate reconstruction during the Holocene (past 12'000 years) that can be used for model-data comparisons. The subset that we use in this study to describe and develop the gridding method contains fossil and modern pollen assemblages of western Eurasia, a region that has already been under investigation in the previous study by Mauri et al., 2015.

The fossil database contains raw pollen counts with in total about 1350 datasets that consists of 80500 fossil samples. The majority of the fossil pollen data (left part of figure 6.1) comes from the European Pollen Database (EPD) (Fyfe et al., 2009, ca. 94%) and other publicly available databases. The presented dataset extends the database used by Mauri et al., 2015 especially with a few sites towards the eastern part of the map.

The modern calibration dataset (6900 samples, see right map in figure 6.1) is mainly based on the version 2 of the Eurasian Modern Pollen Database (EMPD) (Davis et al., 2013, ca. 87%, see also chapter 4) and core tops of EPD (10%) that were younger than 250 years cal BP.

### 6.2.2 Sample site: Tigalmamine

We chose the pollen record of Tigalmamine in Morocco (32.9N, 5.34W, 1626m) to evaluate our method. The site was first studied by Lamb and Kaars, 1995, and the pollen data was downloaded from the European Pollen Database. The chronology and choice of control points used here is that from Giesecke et al (Giesecke

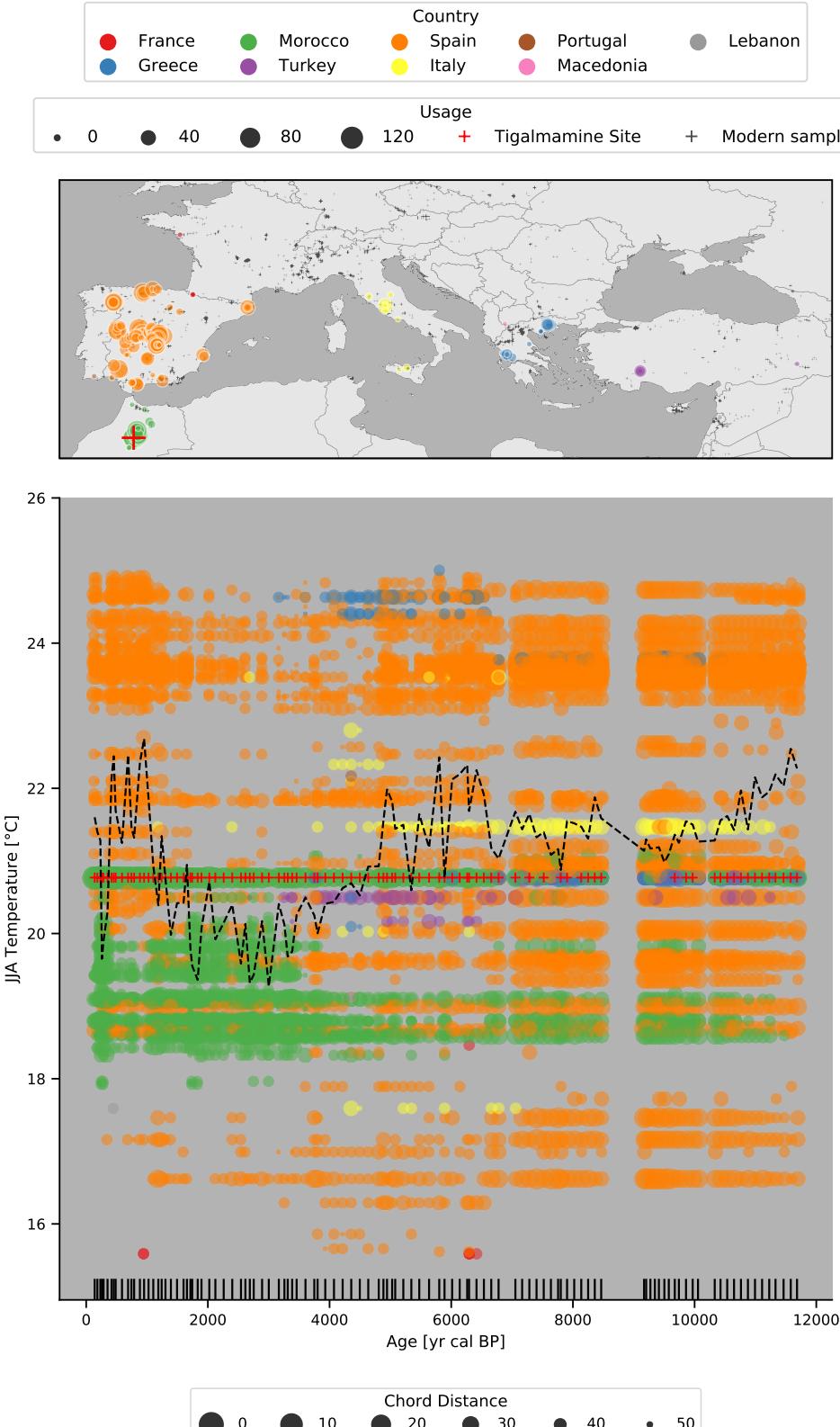


FIGURE 6.2: Climate analogues of the Tigalmamine site (red cross). Every circle corresponds to one modern analogue that was one of the fifties closest analogues in at least one sample within the Tigalmamine dataset. The color-coding of each circle is based its corresponding country (see legend at the top). The marker size in the top plot depends on the usage of the sample as modern analogue. The larger the marker, the more samples in the Tigalmamine dataset use it as modern analogue. Tiny crosses in the map show the locations of the rest of the modern calibration data. The lower plot shows the summer temperature for the analogue (y-axis) at the age of the Tigalmamine sample (x-axis). The marker size in this plot corresponds to the chord distance between modern and fossil pollen assemblage. I.e., the larger the dot, the closer (and more important) the analogue. The dashed line shows the weighted average of all the climate analogues per sample. The age of each Tigalmamine sample is shown with the vertical lines at the bottom of the plot. Red crosses in the lower plot show the Tigalmamine core top sample that has been used as an analogue in 58 out of the 110 samples.

et al., 2013). The site is well dated with 11 radiocarbon dates, and spans the entire Holocene with 110 samples. The data has been used for a previously published pollen reconstruction based on the modern analogue method (Cheddadi et al., 1998), although this study used a calibration dataset that included modern pollen samples from Morocco that have subsequently been found to have geolocation errors (Davis et al., 2013). None of these problematic modern samples have been used in our analysis.

The site (red cross in figure 6.2) is located on the southern edge of our study region in an area with a montane Mediterranean vegetation and climate. The Mediterranean has traditionally been considered to be a particularly challenging environment for pollen-based reconstructions because of the effects of long term human impact, and the interplay of precipitation and temperature on vegetation distribution. The fossil pollen record of Tigalmamine shows a mainly forested montane Mediterranean assemblage throughout the Holocene, dominated by evergreen oak, but with an important transition between the early Holocene and late Holocene marked by a change from deciduous Oak to Cedrus (see the pollen diagram in Cheddadi et al., 1998). The occurrence of Cedrus represents an interesting challenge for any pollen climate transfer function, since this particular taxa is limited in its distribution (and in our calibration dataset) to Morocco and the Lebanon region, while all of the other taxa in the assemblage are widely distributed across the Mediterranean. The strong presence of evergreen Oak also makes the site interesting, because although this taxa is mainly associated with the Mediterranean region, its distribution extends all the way up the west coast of France to Brittany.

Figure 6.2 illustrates these challenges for the particular site. The climate analogues (see next section 6.2.3) span a large summer temperature regime of about 10 degrees, from 15 to 25 °C. The upper temperature range is dominated by analogue climates from Spain (orange) which in general shows the highest number analogue matches. The early Holocene (12k to 8k BP) is dominated by modern samples from Spain, with a wider and more uniform temperature regime, when compared to the later periods. During the transition in the mid-Holocene (8k to 4k BP), analogues from across the Mediterranean Sea play a more important role, in particular from Greece, Italy and Turkey. The lower temperature regime is then dominated by Moroccan samples (green) that are of particular importance during the late Holocene (4k BP to present) due to the above mentioned presence of Cedrus.

The weighted average of the analogues (black dashed line in figure 6.2) is in general about one to two degrees lower than the one in Cheddadi et al., 1998 (very likely due to the above-mentioned erroneous calibration data they used). The trends are however similar: Higher temperatures in early Holocene (the *spanish analogues* dominate) with a drop around 6k BP (Moroccan climates). Our weighted average however also shows a clear increase during the past 2000 years, again driven by spanish analogues.

The climatic and geographic space that is covered by the analogues is further discussed in section 6.3.2.

### 6.2.3 Site-based holocene temperature estimates

A standard approach for site-based climate reconstruction from fossil pollen assemblages is the modern analogue technique (MAT) (also called *k*-nearest neighbors). This technique estimates the climate of the fossil sample as the (weighted) climate average of the most similar modern samples (i.e. the closest modern analogues). It has the major advantage that it requires little parameterization efforts and can be

applied over a large spatial area that covers many different climate regimes (Mauri et al., 2015).

For this purpose, we follow the standard approach and assign a JJA temperature values to each modern calibration sample (figure ??), taken from the corresponding grid cell in the WorldClim dataset, version 2 at 30 seconds (Fick and Hijmans, 2017).

Every pollen assemblage is then transformed from raw counts to percentages, based on the total sum of terrestrial pollen counts per sample that we deem useful for the reconstruction. This excludes low samples with low counts and taxa with low occurrences. We use squared-chord distance from the R package *rioja* (Juggins, 2017) as a measure of similarity. For a given transformed fossil pollen assemblage  $\{f_t\}$  and a modern pollen assemblage  $\{m_t\}$ , where  $t = \{1, \dots, N\}$  denotes one of the  $N$  individual taxa in the assemblages, this distance measure is defined as

$$d = \sum_{t=1, \dots, N} \left( \sqrt{f_t} - \sqrt{m_t} \right)^2$$

This distance is calculated between every modern and every fossil sample in the entire database (section 6.2.1). The standard, non-probabilistic setup would now compute the climate of the fossil sample as the mean climate of the  $k$  closest analogues (e.g.  $k = 6$ ), eventually weighted by their corresponding distance  $d$ . There are many variations of this technique (see for example Birks et al., 2010, including various measures of similarity, choices about  $k$ , the maximum allowed distance  $d$  between modern and fossil assemblage, subsampling of the calibration dataset to avoid spatial autocorrelation (Guiot and Vernal, 2011; Telford and Birks, 2005, 2009), and by grouping pollen taxa into so-called plant-functional types (PFTs) (Davis et al., 2003; Mauri et al., 2015, e.g.). They all, however, have in common that the categorical, multi-modal distribution of the climate of the modern analogues is simplified into a unimodal distribution, represented by the mean of the analogue climates. Therefore, in our ensemble approach, we do not take the mean but sample the climate of the analogues directly. This is further discussed in the methods section 6.3.2 and 6.4.1.

#### 6.2.4 Age uncertainties

In addition to the methodological uncertainties of the climate reconstruction method (previous section 6.2.3), we provide a framework to handle dating uncertainties. During the gridding step (see next section 6.3.3), every sample is weighted by the age difference to the target reconstruction age. The previous studies by Davis et al., 2003 and Mauri et al., 2015 do not take this uncertainty, that can be as high as multiple centuries, into account although they influence the gridded temperature reconstruction.

The reason is a systematic problem of pollen samples that we overcome here with the recent developments in the pollen community. In palynology, each sample in a sediment core is dated using a so-called age-depth model, a function that maps each depth of the sediment core to an age. This function is based on a few chronological control points where the age has been determined instrumentally (for lake sediments in the Northern Hemisphere, these are commonly radiocarbon ( $^{14}\text{C}$  dates) and interpolates/extrapolates to the depths of the sample locations. Various methodologies exist to define these age-depth models, ranging from simple linear interpolation methods (Bennett, 1994) to the more recently developed bayesian techniques of the Bchron (Haslett and Parnell, 2008) and BACON (Blaauw and Christen, 2011) models.

The early approaches have been proven to provide unreliable uncertainty estimates (Telford et al., 2004) and there has been no standardized way to report the uncertainties, if they are reported at all. For this reason we (and previous studies) cannot rely on the age uncertainties reported in the pollen database. An alternative approach is to recalculate the chronology for every dataset in the database (see Goring, 2019, for instance), but this also requires parameterization for reliable uncertainties and goes beyond the scope of this study.

Instead, we follow an approach that is based on two aspects: age uncertainties are higher for older samples, and samples that are farther away from the radiocarbon dates (i.e. chronological control points). Additionally, samples behave differently if the sample is surrounded by two chronological points (i.e. the sample age is interpolated) or not (sample age is extrapolated). To quantify these relationships, we perform a study based on all datasets (ca. 30'000 samples) from the Neotoma paleoecology database (Williams et al., 2018) that have age-depth models estimated with BACON, a model that has been proven to provide more reliable age uncertainty estimates (Trachsel and Telford, 2016). For the sake of implementation (the age sampling in section 6.3.1 assumes a normal distribution), we apply several assumptions and proximations to the Neotoma samples, in particular:

1. We assume that every dataset with a BACON chronology in Neotoma keeps the defaults and reports the limits of the 95% confidence interval (CI)
2. We keep only the maximal distance of the CI limits from the reported age (i.e. we assume a symmetric distribution)
3. We assume that the distribution is normal (i.e. the 95% CI corresponds to the  $2\sigma$  interval, where  $\sigma^2$  denotes the scale parameter) and a division in half of the maximal distance (see previous assumption) gives the standard deviation  $\sigma$  (which is what we call the age uncertainty)

The resulting data is illustrated in figure 6.3. The grayscale density plots in the background shows the high dispersal of the data and the number of samples decreases strongly with higher distance to the control point or older samples (red lines). Nonetheless, the mean of the data (blue lines) reveals the increasing nature of both relationships, as mentioned before.

Figure 6.3 also shows two models that have been fitted to the data. The first one is a standard simple univariate linear model  $y = a + b \cdot x$  (orange line). This model simulates the increasing trend of both variables although it does not capture the non-linear relationship between age and age-uncertainty. A reason for this non-linearity arises from the time-dependency of the radiocarbon calibration curve and its associated errors. This non-linear behavior gives the motivation to use a constrained linear Generalized Additive Model (GAM), a smooth semi-parametric model of the form

$$\mathbb{E}[y|X] = \beta_0 + f_1(X_1)$$

in the univariate case, or

$$\mathbb{E}[y|X] = \beta_0 + f_1(X_1) + f_2(X_2)$$

in the bivariate case. The feature functions  $f_1$  and  $f_2$  are based on penalized B splines with a constraint for monotonic increasing, the expected value  $\mathbb{E}[y|X]$  is based on a normal distribution. The GAM model has been fitted with the *pyGAM* software package (Servén et al., 2018). This model enables to better simulate the non-linear features as can be seen with the green lines in figure 6.3.

These results approve the initial hypotheses and justify the choice of a bivariate GAM for predicting age uncertainties based on the distance to the chronological control point, and the age of the sample. The two models, together with a bivariate simple linear regression model, and again for interpolated and extrapolated samples, are shown in the central column of figure 6.4. Both model classes (simple linear and GAM) are able to reproduce the general shape of the observed data, although the GAM better resolves the non-linear relationship between the three variables.

The final uncertainties, predicted for the data set presented in the previous section 6.2.1, are shown in the supplementary figure 6.17.

## 6.3 Method

With the intrinsic methodological uncertainties of climate and dating in mind, we present a new ensemble-based approach on gridding the reconstructions from the individual sites. Each ensemble member is generated with randomized sample ages and climate, derived from the corresponding uncertainty measures (see previous sections 6.2.3 and 6.2.4), with additional constraints arising from the integrity of the individual dataset (sediment core). We explain these in more details in sections 6.3.1 and 6.3.2. The final gridding step for each ensemble member is based on a modified setup of Mauri et al., 2015, but can also be extended with other interpolation algorithms, as described in section 6.3.3). We implemented the method as the python package *pyleogrid* that efficiently scales to large datasets and ensemble sizes, and shortly describe it in section 6.3.4.

### 6.3.1 Constrained age sampling

Every dataset has an intrinsic monotonicity constraint that the sample deeper down the core has an older age. An inversion of this constraint is very rare and is usually visible in the stratigraphy of the core, such that affected samples are ruled-out before. As such, a classic unconstrained sampling of ages<sup>1</sup> using a normal distribution centered at reported sample age and a scale corresponding to the estimated age uncertainty (section 6.2.4) violates this constraint. Samples are inverted in such a case when their uncertainty intervals overlap and as such the individual ensemble member would not maintain the integrity of the individual core. We illustrate an example for such a core in section 6.4.1.

*pyleogrid* therefore implements different variants of this constraint with the Gibbs sampling being the one that is finally used.

#### The intuitive approach

The most intuitive approach is to randomly draw a sample age and constraint the age of the neighboring sample with it. This can be done in a *forward* manner, such that every older sample has to be older than the previous younger sample, or in a backward manner, i.e. the younger sample has to be younger than the neighboring older sample. We will show in the paragraphs below that this method is biased, nevertheless we mention it here because of the intuitivity of the approach and because the reason for the failure is non-trivial.

As such, we demonstrate three different algorithms:

---

<sup>1</sup>We call it the unconstrained distribution for convenience, but keeping in mind that every sampled age has to be older than -70 yr cal BP.

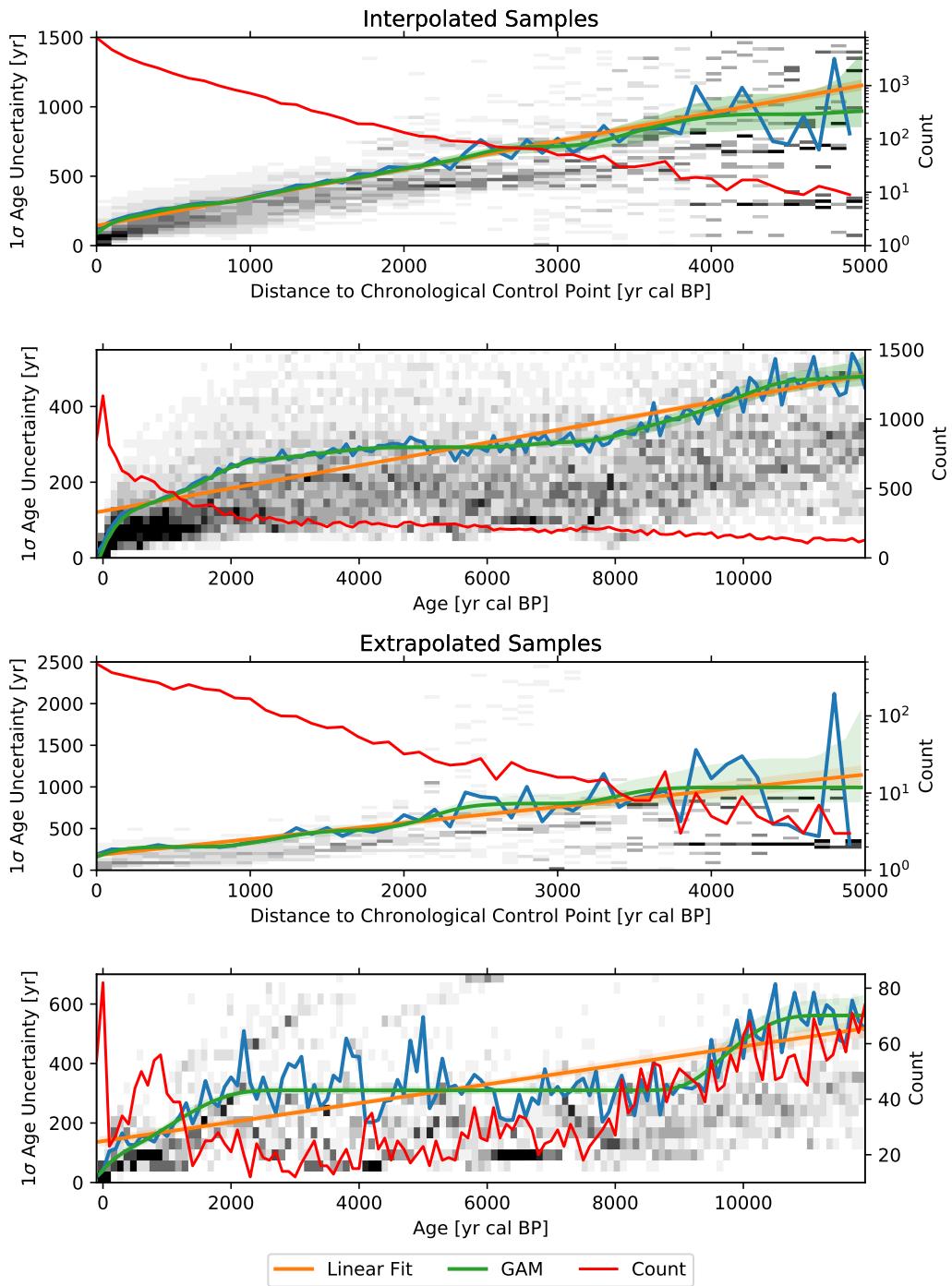


FIGURE 6.3: Univariate regression plots of (first and third) distance to chronological points, and (second and fourth) age to the one sigma dating uncertainty of the sample. The upper two plots contain only interpolated samples (i.e. samples that lie between two chronological control points), the lower extrapolated samples. Blue lines show the mean age uncertainty for the given distance (age). Orange and green lines show the linear and GAM fits of distance (age) to age uncertainty, and red lines show the number of samples for a given distance (age). The grayscale plot in the background shows a two-dimensional histogram (density plot) to illustrate the underlying data of the fits. For the purpose of a better visualization, each vertical bin of this histogram has been normalized to one. Means, counts and histogram are all based on 100 year bins in distance (age). The fits are estimated based on the unbinned data, the source data are all Neotoma datasets with BACON-based age-depth models. Note the logarithmic scale of the right count axis on the first and third plot.

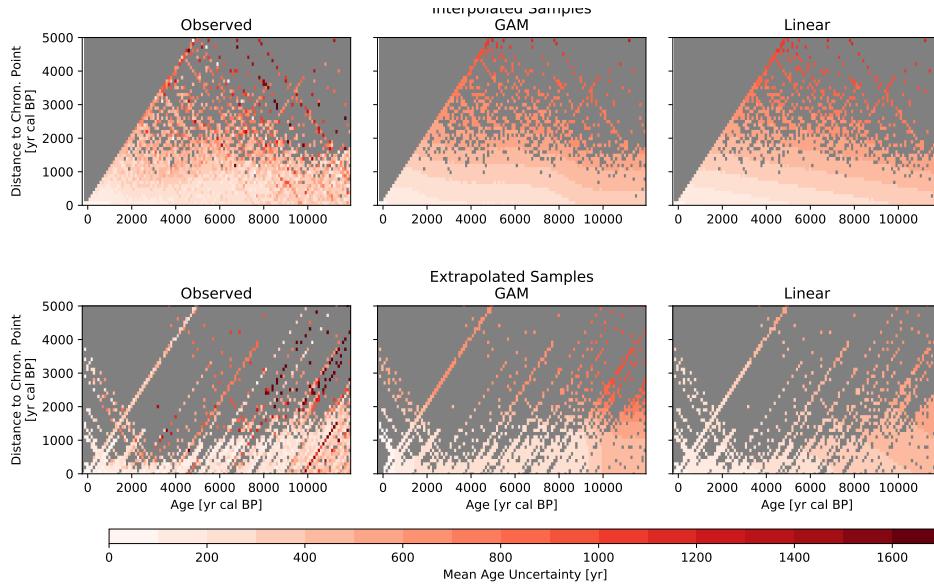


FIGURE 6.4: Bivariate models of age uncertainty. Shown are the mean  $1\sigma$  age uncertainties of ca. 30'000 samples from the Neotoma database sites with BACON-based age-depth models. Each sample has an age uncertainty that depends on the age of the sample (x-axis) and the distance to the closest chronological control point (y-axis). For the purpose of visualization, we grouped the samples into categories of 100 years in x- (age) and 100 years in y- (control point distance) direction, and calculated the average age uncertainty of the groups. These averages are shown with the color coding in the plots, and the gray area represents the space without any observation. The top row shows interpolated samples (i.e. samples that lie between two chronological control points), the bottom row extrapolated samples. Plots in the left column show the observed mean age uncertainties, central and right columns show the means of the predicted age uncertainties from the bivariate linear GAMs and bivariate linear regression models respectively.

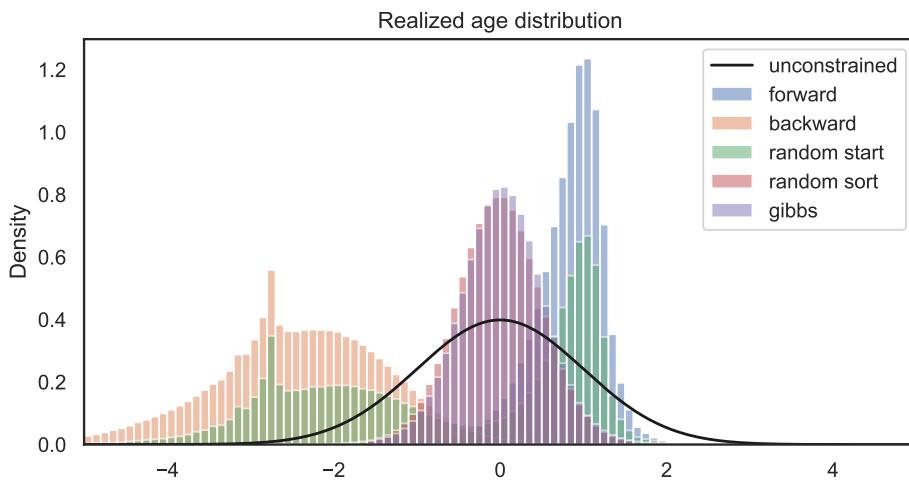


FIGURE 6.5: Histograms of age sampling methods for the site in section 6.2.2 with an ensemble size of 10'000. Every sampled age has been centered at the reported age of the corresponding sample and scaled by its age uncertainty. The black line shows the unconstrained distribution (a standard normal with a standard deviation of 1), the other histograms show the realized distributions for each of the age sampling methods (section 6.3.1). Note that *random sort* and *Gibbs* histograms highly overlap.

**forward** Starting with an unconstrained age distribution for the youngest sample in the core, every consecutive sample has to be older than the previous (i.e. the method works forward in age, but backward in time)

**backward** Starting with an unconstrained age distribution for the oldest sample in the core, every consecutive sample has to be younger than the previous (i.e. the method works backward in age, but forward in time)

**random start** Starting with an unconstrained age distribution of a random sample in the core, we apply the *backward* algorithm for younger and *forward* algorithm for younger samples.

As such, *forward* and *backward* algorithms always start with an unconstrained age distribution of the youngest (oldest) sample for every ensemble member. Within the *random start* algorithm, every sample gets the chance to start with an unconstrained age distribution, because the starting point is random for every ensemble member. The constrained age distributions for the consecutive samples are implemented as truncated normal distributions.

The resulting age distributions from the three algorithms are shown in figure 6.5, together with another method, that is described later in this section. The figure shows the sampled age distributions by the various above-mentioned sampling methods for the site described in section 6.2.2. To make these age distributions comparable, we transformed them to a standard normal distribution (visualized as the unconstrained distribution in figure 6.5) prior to visualization, by subtracting the reported age and dividing by the estimated age uncertainty of the corresponding sample. It is obvious from this figure that all of the above-mentioned algorithms produce an artificial bias to the age distribution. The *forward* approach pushes the samples to the upper tail of the distribution, the *backward* approach pushes everything to the lower tail. The *random start* method produces a bimodal distribution with peaks at the upper and lower tail.

This is also shown with three exemplary samples from the site in the supplementary figure 6.18. The forward method works well for the young sample but pushes all older samples to the upper tail of their distribution. The backward method does the opposite and the random sort method creates a bimodal distribution for the sample in the center of the core, and backward behaves like the forward (backward) algorithm at the older (younger) part of the core.

We explain this initially unexpected results with the overlapping age uncertainties in the core. The site that we describe here has 110 samples. As such, the probability that one sample draws a random age at the lower or upper tail of the distribution is very high. Now, most of the dating uncertainty intervals overlap and this forces all the consecutive samples to the tail of their age distributions. Another problem, that is not shown here, arises from the differing sizes of the age uncertainties which highly depends on the distance to the chronological control point (see section 6.2.4). This can also lead to unsatisfiable requirements, if one sample is close to a control point (and as such has a lower age uncertainty) and the previous sample has been pushed far outside of the 95% confidence interval.

### The random sorting approach

These strong biases of the intuitive approach led to another method, that we also show in red in figure 6.5 and supplementary figure 6.18, the *random sort* method. This method consists of two steps: in the first step we draw random age for each

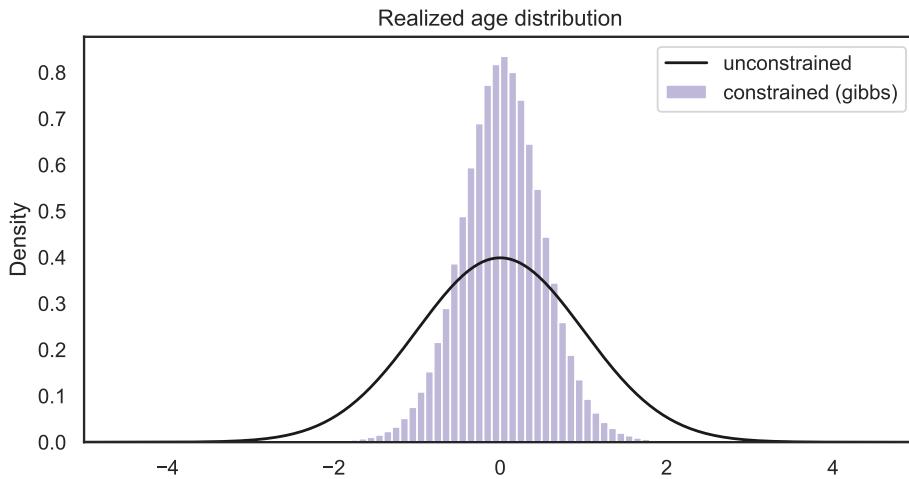


FIGURE 6.6: Realized age distribution for the entire dataset (section 6.2.1) with the *random* method (section 6.3.1). The individual sample distributions have been centered and scaled as in figure 6.5.

sample based on its unconstrained distribution<sup>1</sup>. In the second step, we order these random ages while maintaining the order of samples in each dataset. As such, we assign an age to each sample that is not necessarily drawn from its own distribution, but rather from the one of a neighboring sample. When samples overlap, this then truncates the tails of realized distribution and effectively decreases the reported age uncertainty, as can be seen in the figures 6.5, 6.18. This approach is mathematically difficult to justify because it violates the common methodology that each sample has a unique confidence interval that it needs to explore. Therefore the method might introduce some hidden biases in the sampled distributions that are difficult to quantify. Nevertheless, the algorithm is very fast and much closer to the desired joint distribution, than the previous *intuitive* approach. But in order to avoid any biases and to guarantee a mathematically correct result, we chose to implement a Gibbs sampling algorithm to sample from the desired constrained distribution.

### The Gibbs sampling approach

---

**Algorithm 2** Accept/Reject algorithm.  $\mathcal{N}(\mu, \sigma)$  denotes the normal distribution with location parameter  $\mu$  and shape parameter  $\sigma$ .

---

- 1: Set  $i = 0$
  - 2: Set  $\mu$  as vector of the reported ages in *dataset*
  - 3: Set  $\sigma$  as vector of estimated age uncertainties
  - 4: Set  $\mathbf{a}$  (the target age vector) to be of length  $\mu$
  - 5: **while**  $i < 1$  **or not** *is\_monotonic(a)* **do**
  - 6:    $\mathbf{a} = \mathcal{N}(\mu, \sigma^2)$
  - 7:   Set  $i = i + 1$
  - 8: **end while**
- 

The biases of the above-mentioned algorithms led to the development of a Markov chain Monte Carlo (MCMC) sampling algorithm. An accept/reject algorithm which draws a set of random ages for all unconstrained sample distributions in a core at

once and accepts the draw if the monotonicity condition is satisfied and rejects the sample if the monotonicity condition was initially explored. For one realization of the ages  $\mathbf{a}$  in a given dataset, this is described with the pseudo-code in algorithm 2. This standard approach however did not find a monotonic solution within ten million iterations for a high-resolution site such as it has been used in the previous section.

Therefore we decided to implement a Gibbs sampler, an algorithm that is commonly used in Bayesian inference to obtain a sequence of samples from conditional probability distributions, which generate samples from a multivariate joint distribution when this distribution is unknown and/or cannot be sampled directly. In our case, this distribution is the distribution of all sample ages in one dataset, where each sample age is conditioned by its younger and older neighbor. Let  $\mu = (\mu_1, \mu_2, \dots, \mu_N)$  be the reported ages of the  $N$  pollen samples in one individual dataset with estimated age uncertainties  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$ . The reported ages fulfill the monotonicity constraint, i.e.  $\mu_j \leq \mu_k$  for all  $j, k$  with  $1 \leq j \leq k \leq N$ . The objective of our sampling approach is to generate  $M$  random realizations of  $\mu$ , denoted by  $\mathbf{X}^{(m)} = (X_1^{(m)}, X_2^{(m)}, \dots, X_N^{(m)})$  with  $m = 1, \dots, M$ , that all fulfill the monotonicity constraint. In other words, the realizations  $\mathbf{X}^{(m)}$  are constrained to fulfill

$$X_j^{(m)} \leq X_k^{(m)}, \text{ for all } j, k \text{ with } 1 \leq j \leq k \leq N \text{ and } 1 \leq m \leq M. \quad (6.1)$$

We set the initial value to the reported ages ( $\mathbf{X}^{(1)} = \mu$ ) where we know that the constraint is fulfilled. For the following realizations  $\mathbf{X}^{(m+1)}$  with  $1 < m \leq M$  we sample each component  $X_j^{(m)}$  with  $1 \leq j \leq N$  conditioned by its previous sample  $X_{j-1}^{(m)}$  and, most importantly, conditioned by the next sample, but from the previous realization, i.e.  $X_{j+1}^{(m-1)}$ . As such, we define the sampled age of  $X_j^{(m)}$  with

$$X_j^m = \mathcal{N}(X_{j-1}^{(m)}; X_{j+1}^{(m-1)}; \mu_j, \sigma_j^2) \quad (6.2)$$

where  $\mathcal{N}(a; b; \cdot, \cdot)$  denotes a random variate of the truncated normal distribution with lower limit  $a$  and upper limit  $b$ . Although this algorithm always starts with the youngest sample in the dataset for every realization, such as the *forward* method, it does not push every sample to the lower tail of the distribution because every sampled age is conditioned by the age of the next pollen sample from the previous realization. It is mathematically proven that the combined realizations  $\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(M)}\}$  of this algorithm approximates the joint distribution of the sample ages in the dataset under the given constraint, and that each marginal distribution of the age of a particular pollen sample  $1 \leq j \leq N$  is approximated by  $\{X_j^{(1)}, X_j^{(2)}, \dots, X_j^{(M)}\}$ .

As it is common for a MCMC algorithm, each realization is correlated with nearby realizations. The first samples are particularly correlated with the initial value  $\mu$  and it is therefore common practice to discard the first 1000 realizations, the so-called *burn-in* period.

To avoid an autocorrelation between the successive realizations, we *thin* our set of sampled ages and keep only every tenth realization until we have the desired amount of  $M$  realizations. The value of 10 has been shown to be sufficient using an autocorrelation analysis of the different samples in the Tigalmamine record.

As can be seen in figure 6.5 and 6.18, the outcomes are very close to the above mentioned *random sort* approach. However, a look into the realized distribution of

the last sample in figure 6.18 reveals a negative bias of the distribution sampled with the *random sort* approach.

The realized (and standardized) distribution of the entire database presented in section 6.2.1 is finally shown in figure 6.6. The comparison with the unconstrained distribution in this figure highlights the need for a constrained sampling because the latter significantly reduces the width of the distribution.

### 6.3.2 Temperature sampling

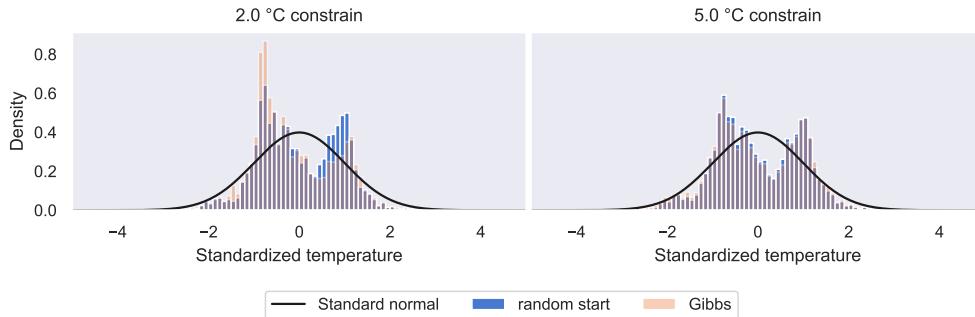


FIGURE 6.7: Histograms of temperature sampling methods for the site in section 6.2.2 with an ensemble size of 10'000 for a climatic constraint of (left) 2 °C, and (right) 5 °C. As in figure 6.5, every sampled temperature has been centered at the weighted average of the corresponding modern analogues and scaled by the corresponding weighted standard deviation. The black line shows the unconstrained distribution (a standard normal with a standard deviation of 1), the other histograms show the realized distributions for *random start* and *Gibbs* temperature sampling method (section 6.3.1).

As already mentioned in 6.2.3, our sampling approach does not use the temperature and uncertainty reported for every single variable. Instead, it samples the underlying distribution. As such, our method can be adapted to multiple site-specific reconstruction methods, such as weighted averaging (WA), weighted-averaging partial-least squares (WAPLS) (Birks et al., 1990; Braak and Juggins, 1993) or other approaches (e.g. Birks et al., 2010; Brewer et al., 2007; Juggins, 2013). In this study, we use a modern analogue technique (MAT) approach (see section 6.2.3) and sample the discrete set of climate analogues for each sample. The probability to select an analogue (i.e. its weight) is thereby determined by the chord distance between the fossil and modern pollen assemblages. The closer the assemblages (relative to the other potential analogues), the higher the weight.

This methodology is substantially different from the standard approach, such that it takes the multimodality of the analogues into account, whereas the standard approach (weighted average of the  $k$  closest analogues) estimates a unimodal distribution. It additionally better represents the discrete nature of the analogue approach whereas the standard method intrinsically assumes a continuity in the distribution. In fact, only a small part of the available climate space is actually represented by the modern analogues. The 5'500 modern analogues for Tigalmamine, for instance (110 samples with 50 analogues each), are represented by only 240 distinct modern pollen samples with only 131 distinct JJA temperatures and they eventually span a large climate space (see figure 6.2).

The latter gives the motivation for a climatic constraint that ensures the integrity of the individual dataset. It is, for instance, impossible that two samples from the

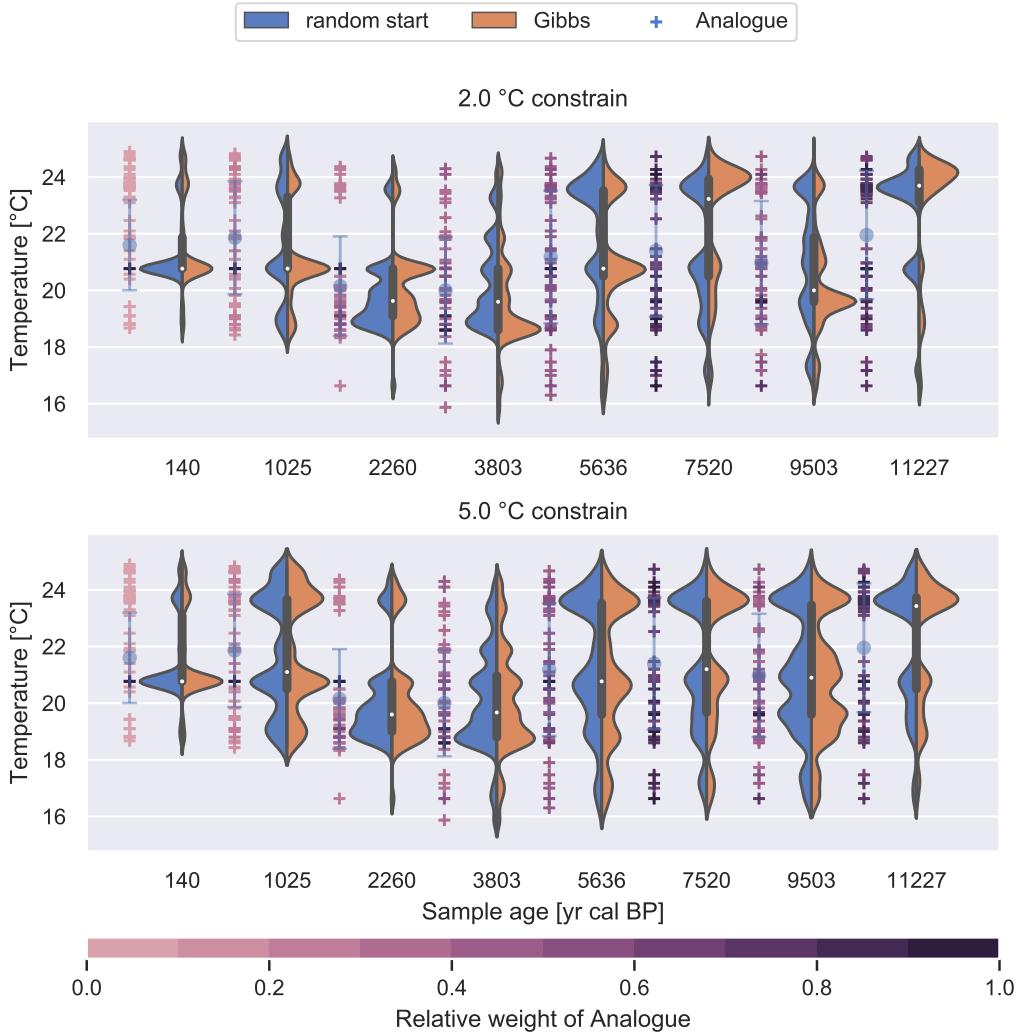


FIGURE 6.8: Violin plots for sampled temperature distributions of some samples in the Tigalmamine record with a climatic threshold of (top) 2 and (bottom) 5 °C. Blue (left) distributions are realized with the *random start* method, ochre (right) areas with Gibbs sampling. The crosses to the left of each violin shows the locations of the climate analogues. Each cross is color coded by its chord distance relative to the chord distance of the closest analogue (i.e. the closest analogue has always a weight of 1).

same dataset but 200 years apart experience a temperature difference of five degrees or more between each other. This is, however, a possible combination, considering the underlying set of analogues (see figure 6.2 for instance). We therefore perform a constrained sampling, as in section 6.3.1, and implement a fixed temperature threshold  $T$ . Every sampled analogue in each dataset (i.e. every choice of the discrete distribution for each pollen sample) is constraint to not differ by more than  $T$  degrees celsius from its temporally neighboring samples. The exact choice of  $T$  is a critical assumption and has a major impact on the realized temperature distribution for each sample. We decided for a very conservative estimate of five degrees, which is only applied if the samples do not differ by more than 1000 years. These choices are further discussed in section 6.4.

In the remainder of this section, we focus on the implementation of this conditional sampling, because of its substantial impact on the realized distribution, as

already shown for the sampled ages in section 6.3.1. We briefly discuss the same approaches as in section 6.3.1 (without the *random sort* algorithm because we do not enforce monotonicity here). The core of the method is the same for all approaches: If a climate analogue differs by more than  $T$  degrees from the temperature of the conditioning sample, its probability is set to zero. The choice about the *conditioning sample* is dependent on sampling algorithm. Here, we discuss following methods:

**forward** The temperature of every older sample must not differ by more than  $T$  from its younger sample

**backward** The temperature of every younger sample must not differ by more than  $T$  from its older sample

**random start** Starting with a random sample in a data, we apply the forward method to older and the backward method to younger samples

**Gibbs** The choice for each sample is constrained by the younger sample and the older sample from the previous realization of the dataset.

We described these algorithms already in detail in section 6.3.1 and therefore focus only on the comparison of results. The only difference is that now, without the monotonicity constraint, *forward* and *backward* methods give the same result as the *random start* method. Therefore we will only focus on the last two methods.

Figure 6.7 shows the realized distributions for the two methods. As in the corresponding figure for the age sampling (figure 6.5), we subtracted the weighted average of the climate analogues of the corresponding pollen sample by each of the randomly sampled temperature values, and afterwards divided by the weighted standard deviation, in order to make the drawn temperature values at the different ages comparable. Both methods realize a bimodal distribution (a feature that is also visible in the spread of the climate analogues in figure 6.2) and result in the similar distributions when considering a climate constraint of 5 °C. This does not hold for the stronger 2 °C constraint, where the Gibbs method gives more weight to samples below the weighted average. This is caused by the additional constraint of the Gibbs sampling approach where each sample is constrained by the sample of the previous realization. The algorithm always starts with the youngest sample which has a higher probability in the Moroccan regime (green area, figure 6.2). Due to the constraint of the sample in the previous realization, it then is more likely that we stay in this regime. As such, the distribution tends to get more unimodal compared to the other method, where each realization is entirely independent of the other.

This difference between the two methods and the two climatic constraints is further illustrated in figure 6.8, which shows the violin plots for a selection of samples in the Tigalmamine record, separated by method and climatic constraint. As already shown with the standardized temperatures in figure 6.7, the two methods are approximately equal under the 5 °C constraint and significantly reduce the realized temperature regime with a more multimodal distribution under the 2 °C constraint. The Gibbs method tends to sharpen the distribution at the locations of the analogue climates stronger than the *random start* method, which implies that the latter is again prone to (potentially) large and unknown biases.

### 6.3.3 Gridding

The underlying gridding algorithm is the same as in Mauri et al., 2015, the thin plate spline regression method (*Tps*) of the *fields* R-package (Nychka et al., 2017; R Core

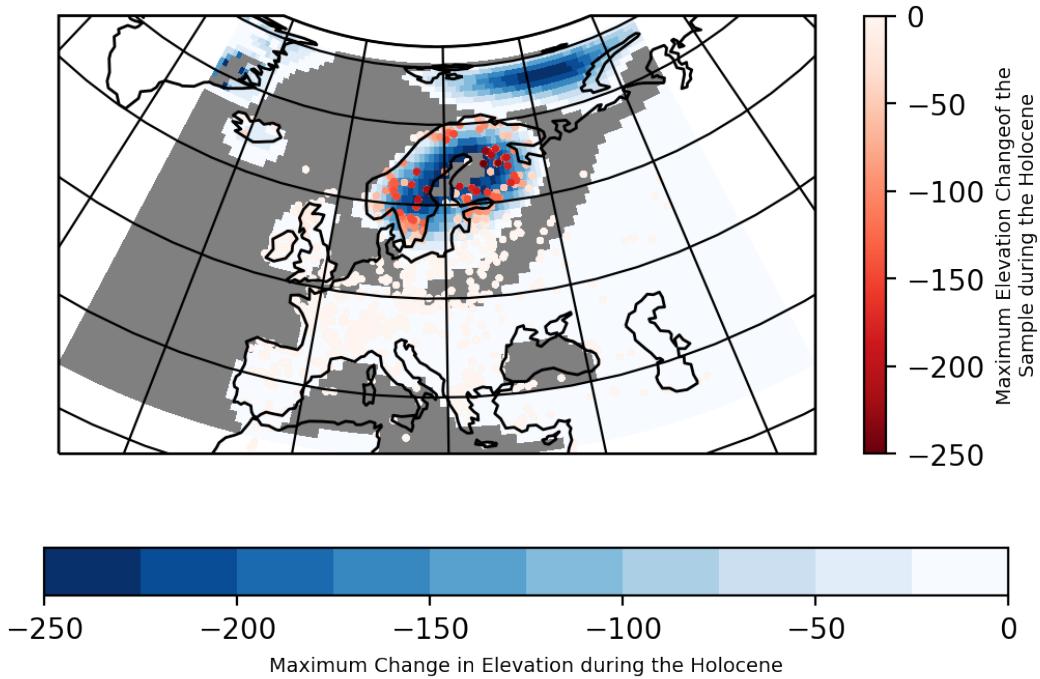


FIGURE 6.9: Maximum change of elevation during the Holocene per grid cell (blue background), based on the data from the ICE-6G model. The dots show the applied (maximal) corrections to the samples in these locations.

Elevation figure has to be made nicer.

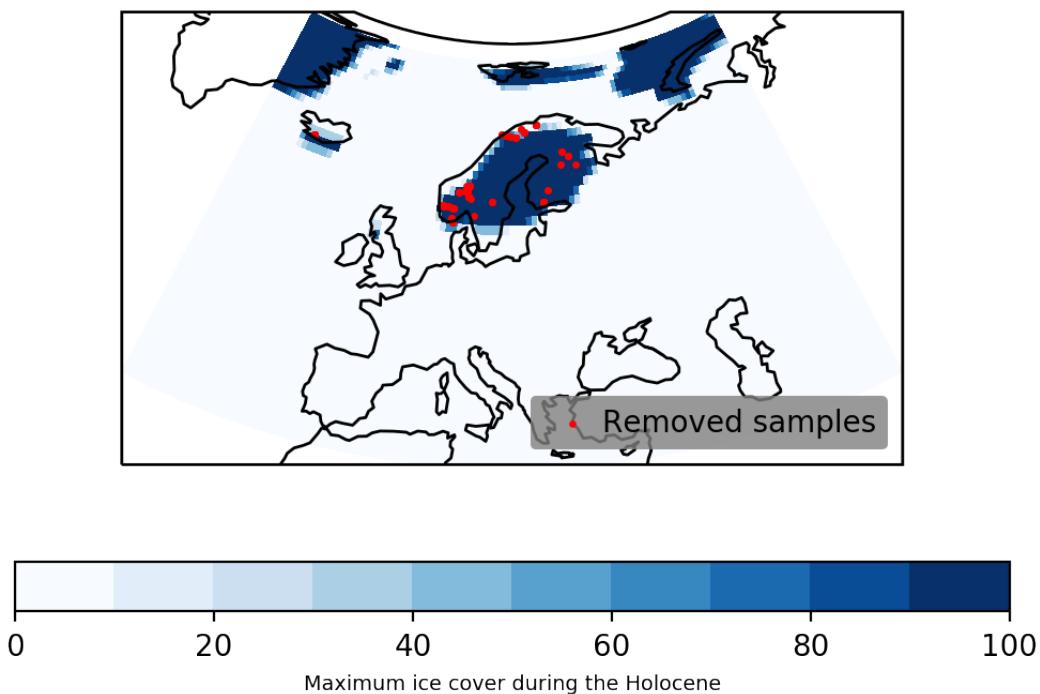


FIGURE 6.10: Locations of samples (red dots) that have been removed from the input data because they are covered by ice (according to the ICE-6G model, blue background).

ice sheet mask has to be made nicer.

Team, 2019). This method interpolates the scalar variable (temperature) from the irregularly spaced four dimensional sample data onto a two dimensional surface (defined by latitude, longitude and elevation) in 3D space. The time component is considered by giving a higher weight to samples that are temporally closer to the target time of the interpolation.

The major difference compared to Mauri et al., 2015 is the ensemble approach, which can also be interpreted as a bootstrapping approach. We apply the gridding many times to different realizations of the input data. The major advantage of this ensemble approach is the sophisticated distribution of temperatures per grid cell. This allows a better estimate of the uncertainty of the gridded climate reconstruction, compared to Mauri et al., 2015. But it should be noted that this ensemble approach is structurally independent of the underlying regression algorithm. Hence, it can potentially also be extended to other gridding methods.

Another difference between our study and Mauri et al., 2015 is the type of climatic variable. Mauri et al., 2015 calculated interpolated anomalies, whereas we interpolate the absolute climate variable as it is derived from the pollen-climate reconstruction method. The advantage is that we can directly interpolate to the elevation at a given timestep, whereas Mauri et al., 2015 applied an a posteriori isostatic correction.

For the reconstruction presented in section 6.4, we use the data of the ICE-6G-C model (Argus et al., 2014; Peltier et al., 2015), that is also used in the PMIP4 experiments (Ivanovic et al., 2016; Kageyama et al., 2018; Otto-Bliesner et al., 2017). To account for the change in elevation, *pyleogrid* also implements a method to correct the elevation of the samples based on the elevation difference in the given input raster for the different time steps. The results of this correction can be seen in figure 6.9. In addition to this elevation correction, we removed samples from the input data where the ICE-6G model reports an ice coverage of more than 50% in the grid cell (figure 6.10). Affected samples are all during the early Holocene in northern Europe.

#### 6.3.4 Implementation

The ensemble method presented in this thesis is available as the python package *pyleogrid* from [github.com/Chilipp/pyleogrid](https://github.com/Chilipp/pyleogrid). [The documentation of the package](#) is available at [pyleogrid.readthedocs.io](https://pyleogrid.readthedocs.io). This module also contains the models for predicting age uncertainties (section 6.2.4), that are based on the *pyGAM* software (Servén et al., 2018). The github repository and the documentation contains the notebooks that are used to run the analysis of this study.

add DOI

The sampling methods of *pyleogrid* that generate entirely independent realizations of the input data (forward, backward, random start and random sort methods) are built using the functionalities of the numerical numpy and scipy packages (Jones et al., 2001; Oliphant, 2006) to efficiently generate thousands of constrained realizations of the input data. The computationally more expensive Gibbs sampling algorithms (section 6.3.1 and 6.3.2) were every realization is constrained by the previous realizations, is implemented in Cython (Behnel et al., 2011), an optimising static compiler for Python, and uses the corresponding Cython Application programming interfaces (APIs) of numpy and scipy.

As already mentioned in section 6.3.3, *pyleogrid* uses the *Tps* method of the *fields* package (Nychka et al., 2017) and as such interfaces into an R environment for the gridding.

*pyleogrid* is designed to scale to large amounts of input data (e.g. for a global reconstruction or hemispheric reconstruction) on a local computer or a large parallelized cluster, using the xarray and dask packages for parallel computing and out-of-core computation (Dask Development Team, 2016; Hoyer and Hamman, 2017).

## 6.4 Results

In this section we briefly present the results of the final temperature reconstruction, both for the site-based reconstruction at Tigalmamine (section 6.4.1), and the western Eurasian gridded temperature record for selected time slices (section 6.4.2). The purpose of this section is present the results of the ensemble approach with a special focus on the choices that have been made in the methods section, particularly the number of analogues and the climatic constraint. We focus on the ensemble mean, but it is important to keep in mind that our method approximates the joint spatio-temporal distribution of the climate.

### 6.4.1 Site-based realized climate reconstruction: a use-case

The final reconstruction of the Tigalmamine site for an ensemble size of 10'000 realizations is displayed in figure 6.11. It shows a density plot of the realized temperature and age regime, as well as the weighted average of the *standard approach*. For comparison we also show the mean of the marginal temperature distribution record from our method (we call this the *realized mean* from now). The plot shows how our method realizes the site within the ensemble. It shows the *Moroccan* regime during the late Holocene (lower left dark green area between 0 and 4'000 years cal BP, see also figure 6.2) and the *Spanish* regime during the early to mid Holocene (upper right dark green area in figure 6.11). The figure also displays the differences between the two temperature constraints. As already mentioned in section 6.3.2, the 2 °C constraint results in a more multimodal distribution, especially towards the end of the record. The the realized mean is therefore significantly different from the weighted average of the standard approach.

Figure 6.12 shows the same figure but for a varying number of analogues. The trends of the three images behave the same and show all a decrease in temperature during the early holocene and a small plateau during the last millenium. Also the temperature regimes conform with a the above-mentioned dominance of the *Moroccan regime* in the late Holocene the *Spanish regime* during the early Holocene. With a lower number of analogues, however, (top plot) the temperature changes in the means are much smaller over a short period of time. The distribution for this scenario (green 2D histogram in the plot) shows a strong multimodality with a few temperature values that have a very high probability. The scenarios with higher analogues (20 or 50) result in a much smoother mean because the underlying distribution covers more of the available age-temperature space.

### 6.4.2 Gridded summer temperature

#### Deterministic vs. ensemble approach

The gridded reconstruction for the entire dataset is shown in figure 6.13 for 3k, 6k, 9k and 12k BP. The figure shows the temperature anomaly to the modern time (0k BP), both for the ensemble method (10'000 realizations, 5 °C climate constraint) and the deterministic case, i.e. without climate and temperature sampling. For the deterministic case, we followed the approach by Mauri et al., 2015 and mask cells that are

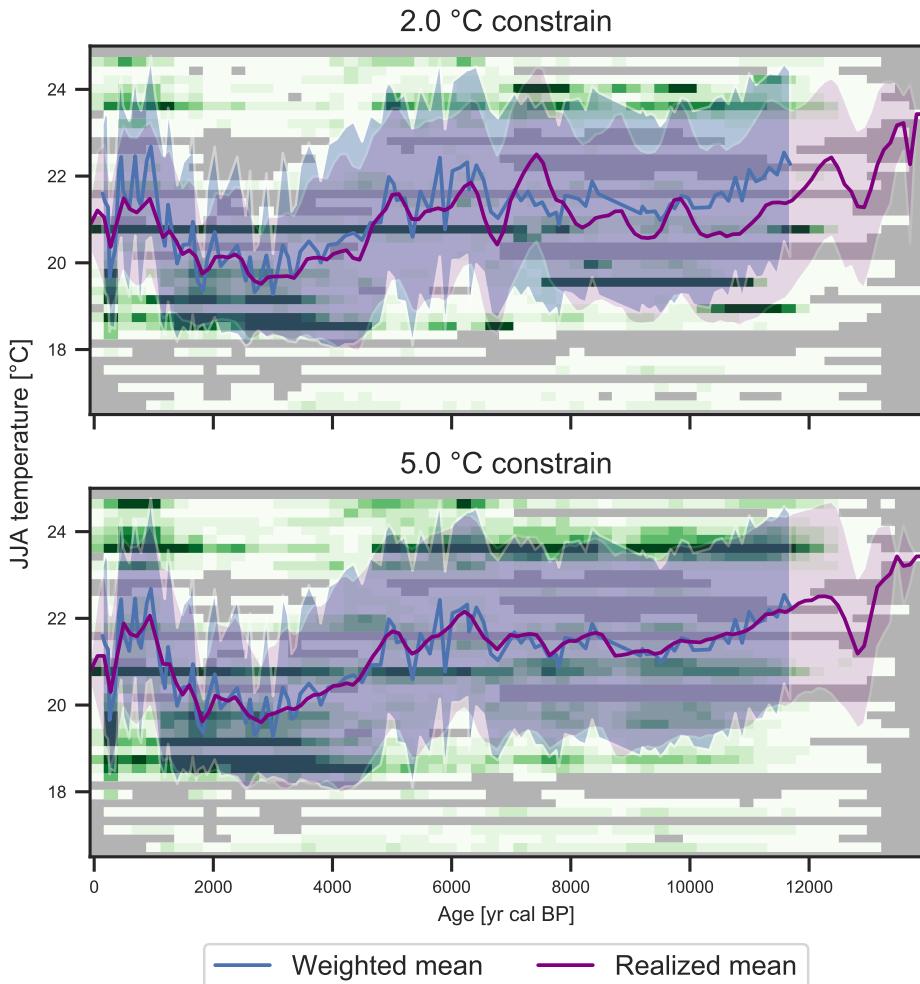


FIGURE 6.11: Realized summer temperature reconstruction of Tigalmamine from the Gibbs sampling with a temperature constraint of (top) 2 °C and (bottom) 5 °C. The background shows a density plot of the sampled age-temperature pairs within the ensemble of 10'000 realizations. The purple line is the mean of all sampled temperatures within 100-year bins, i.e. it represents the mean of the marginal temperature distribution at a given age. The blue curve is the weighted average of the standard MAT approach (black dashed line in figure 6.2). The shaded areas correspond to the standard deviations of the corresponding line.

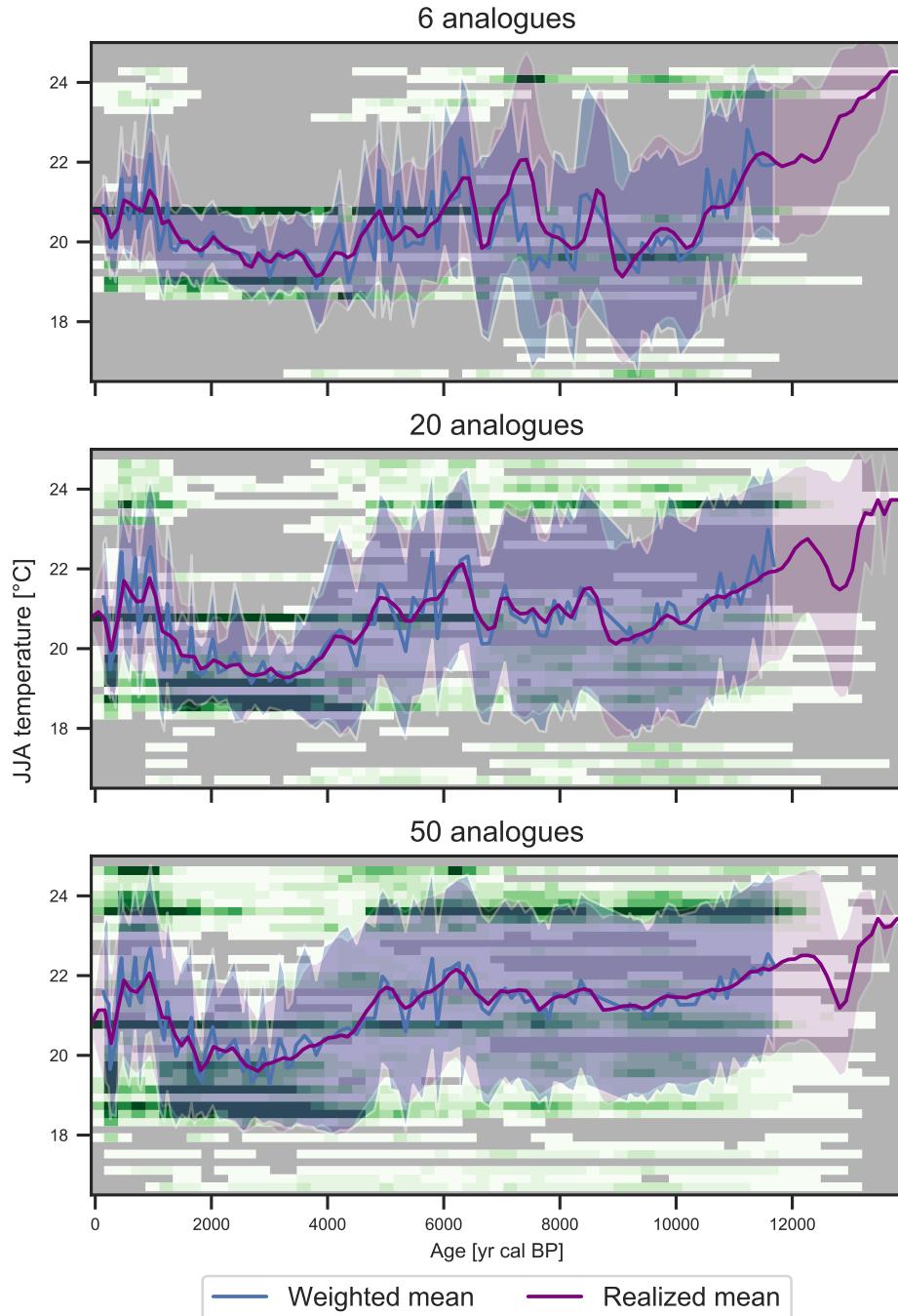


FIGURE 6.12: Realized summer temperature reconstruction of Tigalmamine from the Gibbs sampling with 5 °C constraint for (top) 6 analogues, (middle) 20 analogues and (bottom) the default 50 analogues. The weighted means (blue curves) have been calculated using only the 6, 20 or 50 analogues. See figure 6.11 for a description of the elements in the plot.

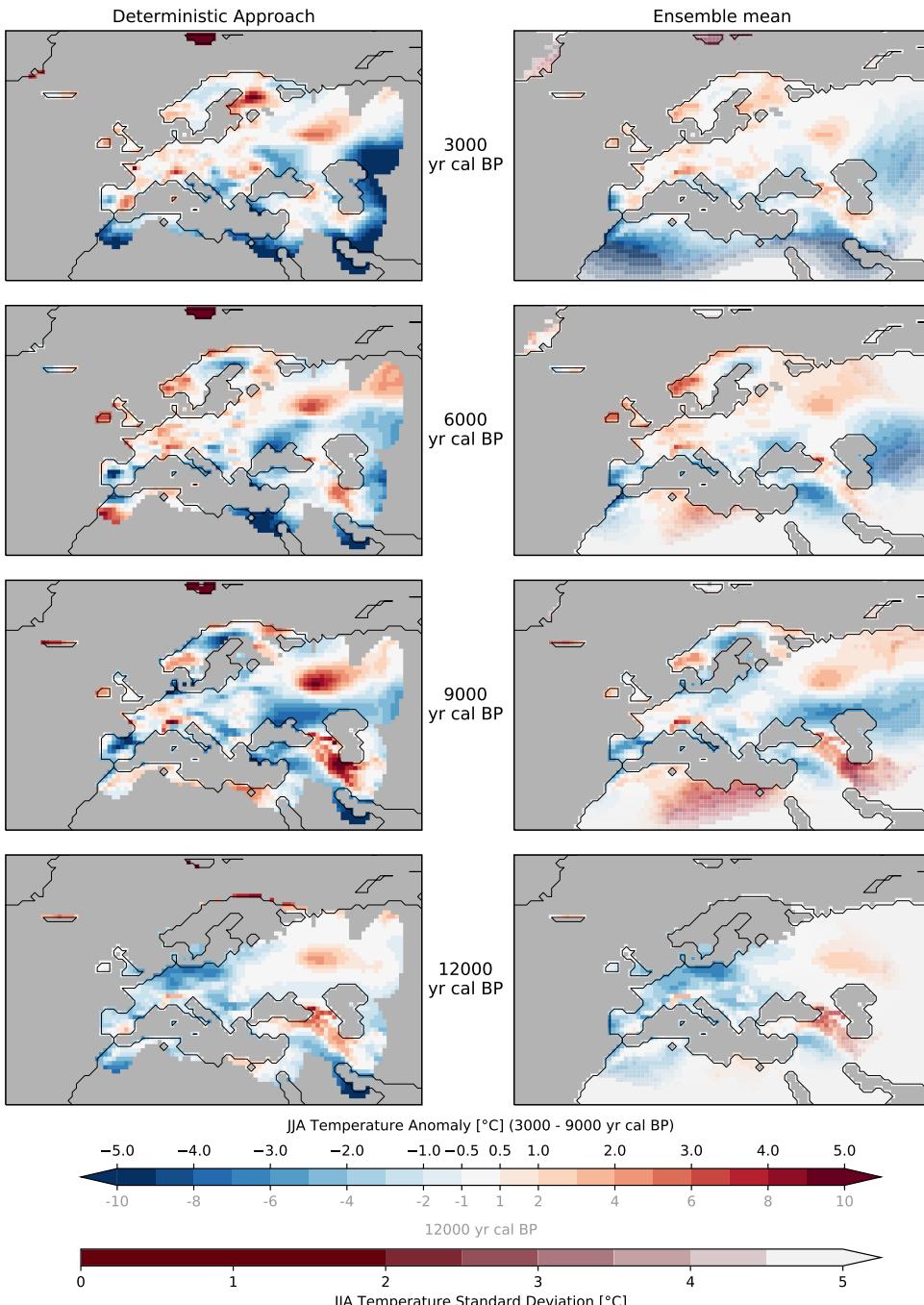


FIGURE 6.13: Deterministic approach vs. ensemble mean (10'000 realizations) at 3k, 6k, 9k and 12k BP (note the different color coding for 12k BP). Each map shows the anomaly with respect to the gridded reconstruction at 0k BP. The deterministic approach (left column) is the input data with temperatures as weighted averages of the 50 closest analogues and without age sampling. The gridded reconstruction has been masked when more than 500km away from the closest sample. The ensemble mean (right column) is overlayed by the corresponding ensemble standard deviation of the anomaly (suppl. figure 6.19). This overlay is transparent for standard deviations smaller than 2.5 °C and afterwards gets more opaque until it reaches 5 °C. Grid cells in the maps that were covered by more than 50 percent with ice, according to the ICE-6G data, have been masked.

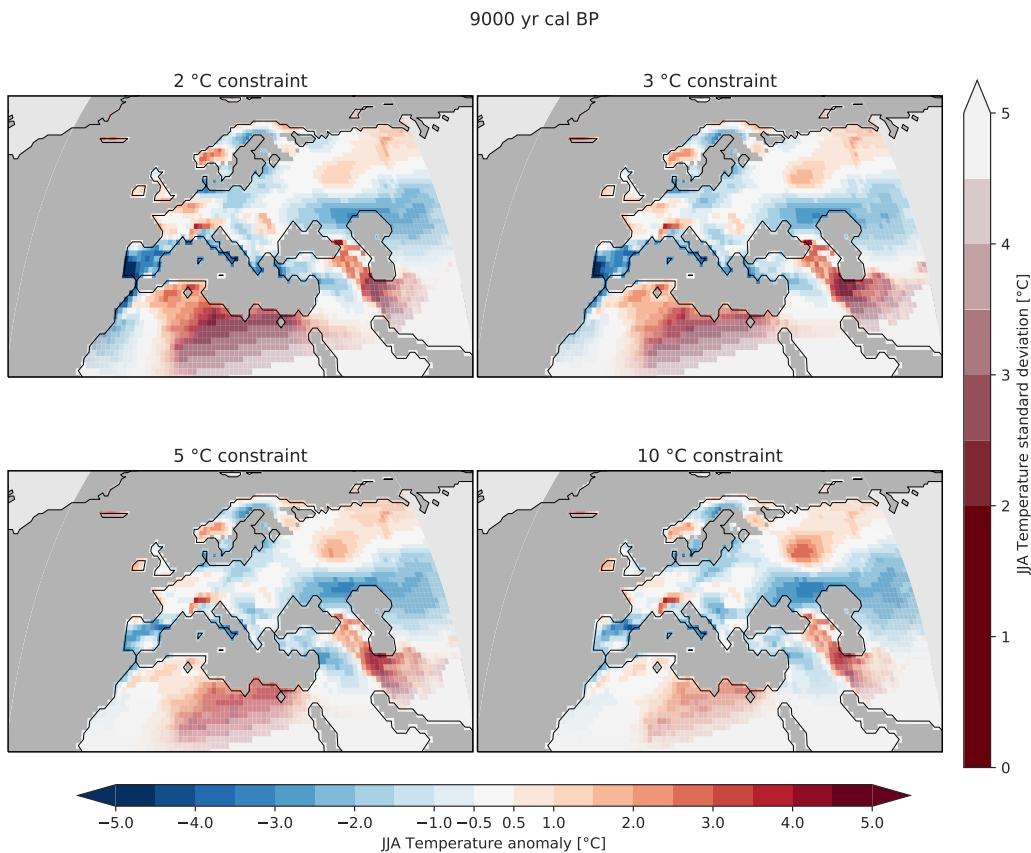


FIGURE 6.14: Ensemble mean summer temperature anomaly at 9k BP (1'000 realizations) for different temperature constraints. The standard deviation is visualized on top of the mean with an increasing opaqueness (same as in figure 6.13).

further away than 500km from any of the pollen sites. For the ensemble approach on the other hand we used the ensemble standard deviation as a gray overlay that is getting more opaque for higher standard deviations.

Both methods show similar patterns and span the same temperature range at the onset of the Holocene. The temperature ranges in the other maps, however, are different with the deterministic approach being more extreme in certain locations, particularly in Spain, eastern Russia and the western Mediterranean region. The method predicts absolute temperature anomalies of more than four degrees, whereas the anomalies of the ensemble approach commonly range between -2.5 and 2.5 °C are more coherent. The ensemble standard deviation of the anomaly (suppl. figure 6.19) in Europe ranges between 0.5 and 1 °C for the mid- to late Holocene timesteps at 6k and 3k BP, and between 1 and 2 °C at 12k and 9k BP. The standard deviation is particularly high at the map boundaries of the 12k reconstruction, very likely due to the smaller availability of fossil samples in this period.

### Temperature sampling parameters

Figure 6.14 shows the influence of the choice for the climatic constraint (see section 6.3.2) on the gridded reconstruction for a selected timestep at 9k BP. The comparison of the 5 °C and the very relaxed 10 °C constraint shows only minor differences. Stronger constraints (2 °C and 3 °C) result in a strengthening of the anomaly, particularly in Spain and towards the southern and south-eastern borders of the map.

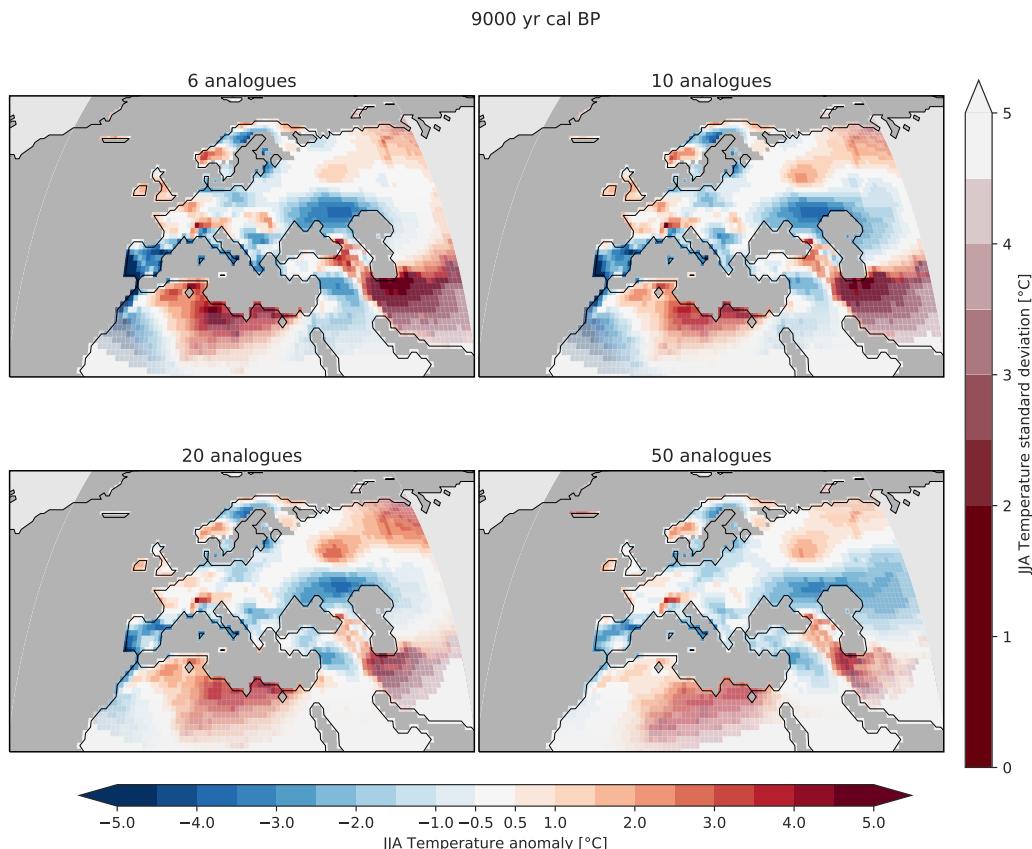


FIGURE 6.15: Ensemble mean summer temperature anomaly at 9k BP for numbers of analogues. The standard deviation is visualized on top of the mean with an increasing opaqueness (same as in figure 6.13).

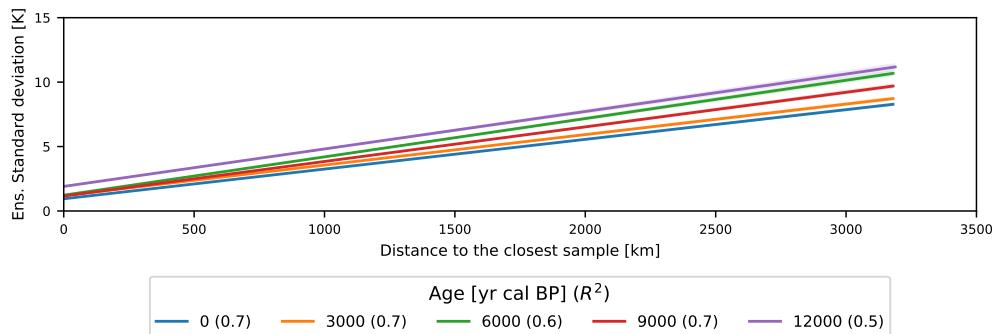


FIGURE 6.16: Relationship between distance to proxy sites and the ensemble standard deviation. Each line corresponds to a linear fit where the distance to the closest site in the database for a given grid cell is regressed against the ensemble standard deviation in this cell (figure 6.19). The different lines represent the different time steps from figure 6.13, the value in brackets after each legend label shows the coefficient of determination ( $R^2$ ) for the given regression fit.

These impacts are coherent with a reduced number of analogues, as shown in figure 6.15. Lower numbers of analogues, e.g. 6 or 10, result in a similar strengthening of the anomaly in these geographic regions.

### Uncertainties for spatial extrapolation

One important issue of gridding is the question about how far we can extrapolate spatially outside of the geographic domain of our proxy database. Mauri et al., 2015 for instance chose a fixed threshold of 500km to the closest pollen site (also shown in the deterministic plots of figure 6.13). Grid cells that are further away than this, are masked.

The ensemble method we present here does not require such a formal definition and should rather be interpreted with respect to the calculated uncertainties. Figure 6.16 shows a clear linear relationship between this distance and the ensemble uncertainty. It also reveals the caveats of using fixed threshold because the relationship between standard deviation and distance to the closest site is time dependent, likely because of the smaller amount of available samples during the early Holocene.

## 6.5 Discussion

In this section, we will discuss the methodological uncertainties of the method that are not necessarily easy to quantify. They arise from first, the underlying site-based proxy-climate reconstruction method (in this case MAT), second the constrained climate sampling to approximate the joint distribution, and third, the gridding algorithm (*Tps*).

The results from the site-based reconstruction in section 6.4.1, and the gridded reconstruction in section 6.4.2, show that the first two aspects (site-based reconstruction method uncertainty and climate sampling) are closely linked, and potentially have a strong influence on the outcome. The effects of a reduced number of analogues (figures 6.12 and 6.15), as well as a stronger climatic constraint (figures 6.11 and 6.14) resulted in more extreme temperature anomalies over the same geographic regions. This result is coherent with the sampled temperatures presented in the

methods section (figure 6.8) where we show that a higher climatic constraint particularly strengthens the closer climate analogues and as such effectively reduces the number of analogues that is used in the reconstruction. There is no clear and definitive answer to the question how many analogues one should use and how strong the climatic constraint should be. Based on the results however, we favor a less strong climatic constraint of 5 °C and a high number of analogues (50). The philosophy behind is to let the method choose which analogue it uses. If there are inconsistencies within the set of modern analogues for a particular dataset, then this should be reflected by the ensemble standard deviation, and can eventually be compensated by the spatially neighboring samples in the gridding process.

These are all problems that arise from the underlying discrete distribution of modern analogues. An alternative would be to use a PDF based method (Chevalier et al., 2014; Chevalier, 2019, for instance) that can potentially overcome these weaknesses by providing a more continuous distribution for sampling the climate.

The third aspect of the uncertainty is related to the gridding algorithm itself which can be added on top of the uncertainty of our ensemble method. For the *Tps* method, this uncertainty can be estimated conveniently using the *predictSE* function of the *fields* R-package that approximates the covariances of the prediction based on a linear combination of the observed data under the assumption of fixed covariance parameters (see Nychka et al., 2017 for details). A calculation of this standard error for 20 ensemble members revealed that it is rather independent of the individual realization. As such, we present the averaged standard error of these 20 members in the supplementary figure 6.20. We can see that this uncertainty estimate is high towards boundaries of the interpolation domain, but smaller than the ensemble standard deviations in between.

## 6.6 Conclusions

With *pyleogrid* we are presenting a new methodological framework that transforms multiple site-based proxy-climate reconstructions into a joint spatio-temporal probabilistic climate reconstruction. Our method exploits the climatic and temporal space that is spanned by the intrinsic uncertainties related to the proxy-climate reconstruction method and the dating of the samples, in order to approximate the distribution of potential climate states in the geographic area of interest. Our approach requires little parametrization, is computationally efficient and can be scaled to large hemispheric or even global areas. The generic ensemble approach we present is in principle agnostic to the underlying proxy-climate reconstruction method and to the gridding method and can therefore be extended to a wide range of potential applications.

Compared to previous approaches of a large-scale gridded reconstruction, our method therefore provides a more reliable uncertainty estimate based on our constrained sampling approaches, which is essential for the comparison with climate model output.

The methodology comes with two side-products that are essential for the spatio-temporal ensemble approach. The first one is a methodology to estimate dating uncertainties based on a bivariate model of the age of the sample and its temporal distance to the closest chronological control point in the dataset. The other side-product, is a probabilistic variant of the site-based modern analogue technique (MAT) that uses a Gibbs sampling algorithm for the ages and analogue climates in order to approximate the joint distribution within a single dataset. This sampling algorithm

is constrained for the individual dataset through first, the monotonicity of the sample ages, and second, a climatic threshold that must not be overcome between too temporally neighboring samples. We compare this sampling algorithm to computationally faster algorithms that involve a forward sampling (climate/age of a sample is constrained by its younger predecessor), its inversion, the backward sampling, as well as a combined approach that uses forward and backward sampling. For age sampling we also test an algorithm that starts with an unconstrained sampling of the ages and then applies an *a posteriori* sorting in order to maintain the correct distribution of samples in the core (sections 6.3.1 and 6.3.2). All of these approaches, however reveal biases when compared to the computationally more demanding, but stationary correct distribution of the Gibbs sampler. The software package *pyleogrid* therefore implements a computationally efficient version of this Gibbs sampling algorithm that efficiently scales from one single dataset to tenth of thousands of realizations of more than a thousand individual datasets, as it has been used in this study.

This sampling successfully reconstructs a probabilistic version of the individual dataset and provides reliable uncertainty estimates. An evaluation of this realization with respect to the number of modern analogues, and the climatic constraint of the sampling algorithm reveals a close linkage of the two parameters. Further developments might therefore explore the usage of other proxy-climate reconstruction methods that provide a more continuous distribution to sample from.

## Supplementary material

### 6.A Estimated age uncertainties

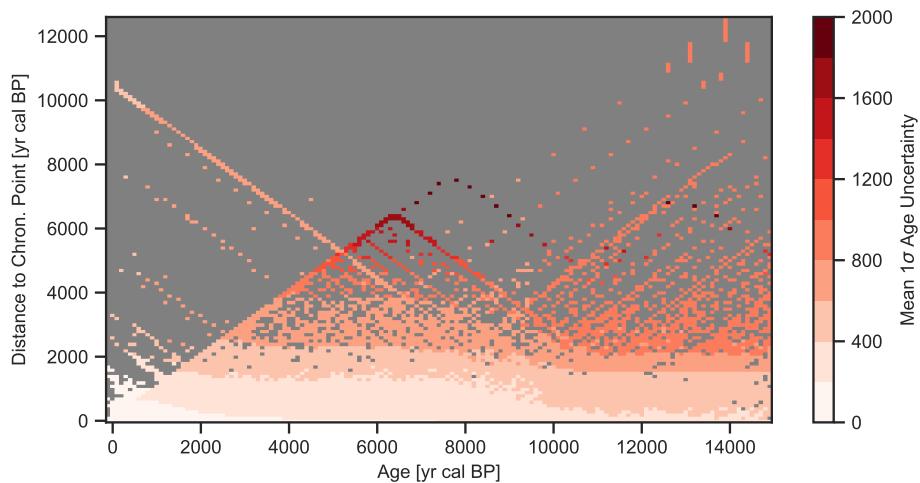


FIGURE 6.17: Estimated age uncertainties for the Eurasian dataset from section 6.2.1 with the same formatting as in figure 6.4.

### 6.B Example of generated age distributions

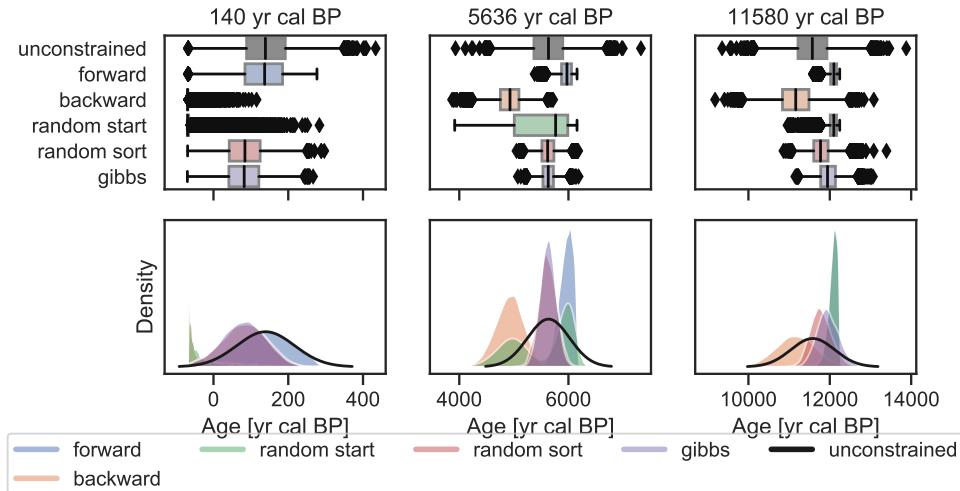


FIGURE 6.18: Example of three samples from the site in section 6.4.1 and their realized distributions. Sampling algorithms are explained in section 6.3.1. Top plots show the box plots of the realized distribution that are visualized with a kernel density estimate in the lower row.

### 6.C Maps of uncertainties

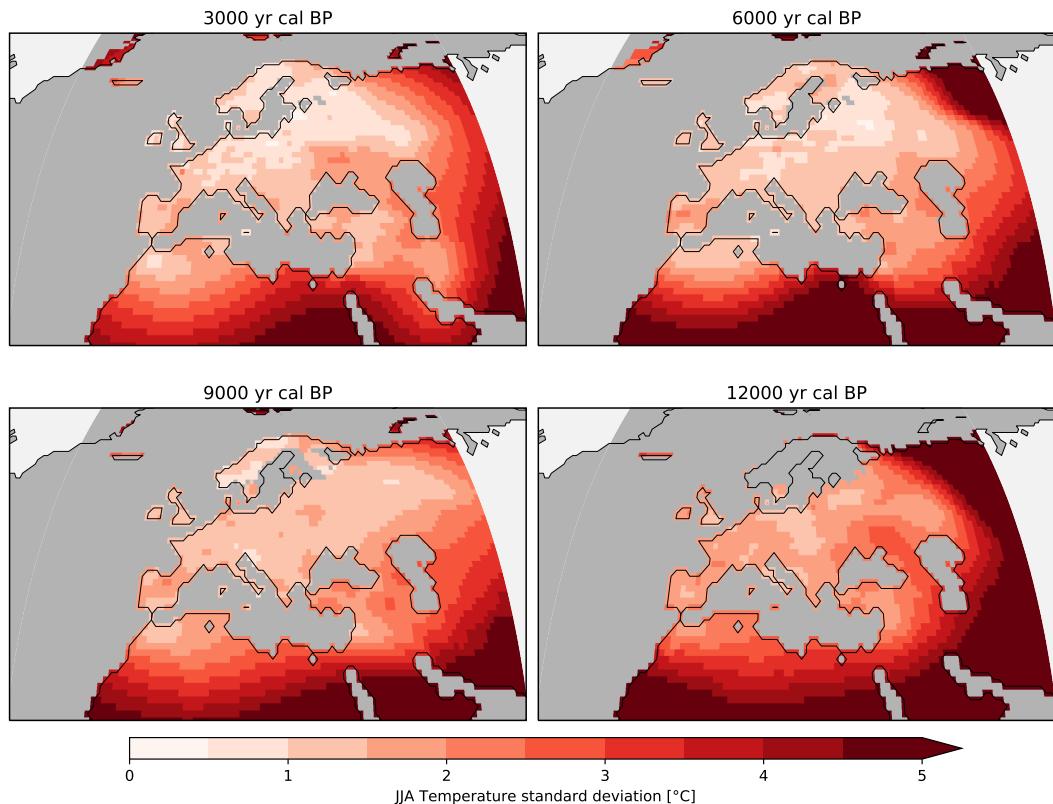


FIGURE 6.19: Standard deviation of the ensemble mean summer temperature anomaly (10'000 realizations) at 3k, 6k, 9k and 12k BP.

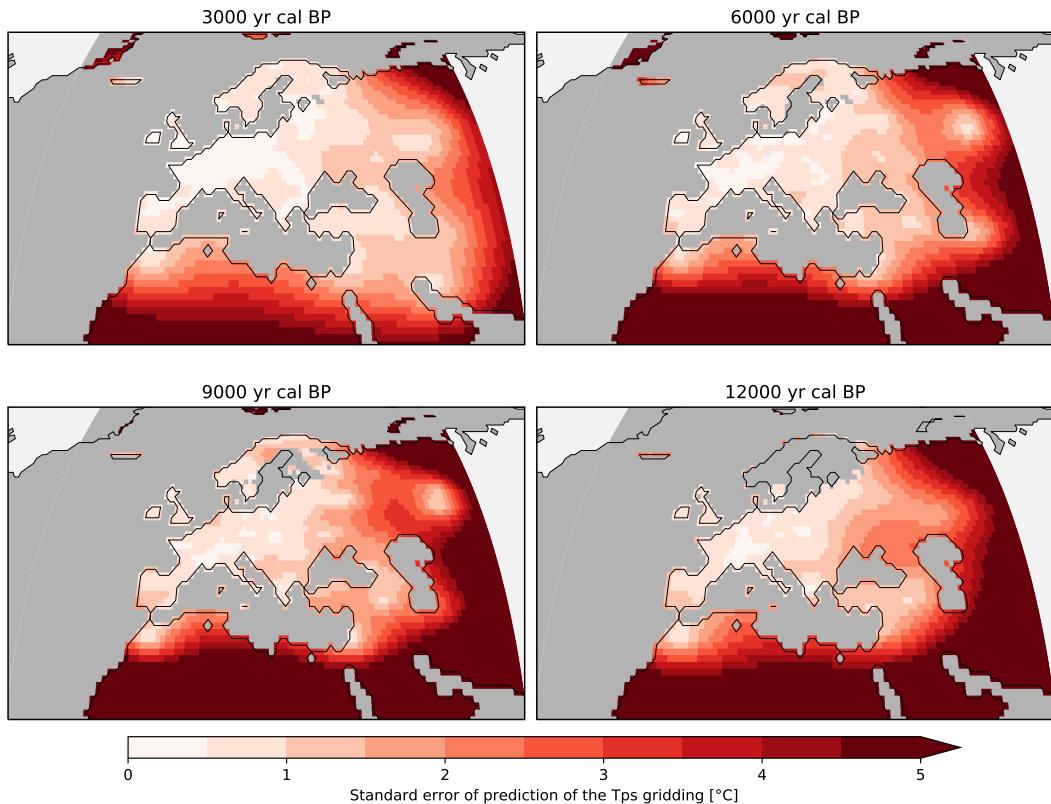


FIGURE 6.20: Standard error of the gridding algorithm, predicted with the *predictSE* method of the *fields* package (Nychka et al., 2017) for the different timeslices. The standard errors are an average of 20 ensemble members, although there is only very little difference between the individual maps.

## References

- Argus, Donald F., W. R. Peltier, R. Drummond, and Angelyn W. Moore (2014). "The Antarctica component of postglacial rebound model ICE-6G\_C (VM5a) based on GPS positioning, exposure age dating of ice thicknesses, and relative sea level histories". In: *Geophysical Journal International* 198.1, pp. 537–563. ISSN: 0956-540X. DOI: [10.1093/gji/ggu140](https://doi.org/10.1093/gji/ggu140). eprint: <http://oup.prod.sis.lan/gji/article-pdf/198/1/537/17366620/ggu140.pdf>. URL: <https://doi.org/10.1093/gji/ggu140>.
- Behnel, Stefan, Robert Bradshaw, Craig Citro, Lisandro Dalcin, Dag Sverre Seljebotn, and Kurt Smith (2011). "Cython: The Best of Both Worlds". In: *Computing in Science & Engineering* 13.2, pp. 31–39. ISSN: 1521-9615. DOI: [10.1109/mcse.2010.118](https://doi.org/10.1109/mcse.2010.118).
- Bennett, K.D. (1994). "Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences". In: *The Holocene* 4.4, pp. 337–348. DOI: [10.1177/095968369400400401](https://doi.org/10.1177/095968369400400401). eprint: <https://doi.org/10.1177/095968369400400401>. URL: <https://doi.org/10.1177/095968369400400401>.
- Birks, H. J. B., J. M. Line, S. Juggins, A. C. Stevenson, and C. J. F. T. Braak (1990). "Diatoms and pH Reconstruction". In: *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 327.1240, pp. 263–278. DOI: [10.1098/rstb.1990.0062](https://doi.org/10.1098/rstb.1990.0062). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.1990.0062>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.1990.0062>.
- Birks, H. John B., Oliver Heiri, Heikki Seppä, and Anne E. Bjune (2010). "Strengths and Weaknesses of Quantitative Climate Reconstructions Based on Late-Quaternary Biological Proxies". In: *The Open Ecology Journal* 3.1, pp. 68–110. DOI: [10.2174/1874213001003020068](https://doi.org/10.2174/1874213001003020068).
- Blaauw, Maarten and J. Andrés Christen (2011). "Flexible paleoclimate age-depth models using an autoregressive gamma process". In: *Bayesian Analysis* 6.3, pp. 457–474. ISSN: 1931-6690. DOI: [10.1214/11-ba618](https://doi.org/10.1214/11-ba618).
- Braak, Cajo J. F. ter and Steve Juggins (1993). "Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages". In: *Hydrobiologia* 269.1, pp. 485–502. ISSN: 1573-5117. DOI: [10.1007/bf00028046](https://doi.org/10.1007/bf00028046). URL: <https://doi.org/10.1007/BF00028046>.
- Brewer, S., Joel Guiot, and Doris Barboni (2007). "Pollen data as climate proxies". In: *Encyclopedia of Quaternary Science*. Elsevier, pp. 2497–2508. URL: <https://hal.archives-ouvertes.fr/hal-00995404>.
- Cheddadi, R., H. F. Lamb, J. Guiot, and S. van der Kaars (1998). "Holocene climatic change in Morocco: a quantitative reconstruction from pollen data". In: *Climate Dynamics* 14.12, pp. 883–890. DOI: [10.1007/s003820050262](https://doi.org/10.1007/s003820050262).
- Chevalier, M., R. Cheddadi, and B. M. Chase (2014). "CREST (Climate REconstruction SofTware): a probability density function (PDF)-based quantitative climate reconstruction method". In: *Climate of the Past* 10.6, pp. 2081–2098. DOI: [10.5194/cp-10-2081-2014](https://doi.org/10.5194/cp-10-2081-2014).
- Chevalier, Manuel (2019). "Enabling possibilities to quantify past climate from fossil assemblages at a global scale". In: *Global and Planetary Change* 175, pp. 27–35. DOI: [10.1016/j.gloplacha.2019.01.016](https://doi.org/10.1016/j.gloplacha.2019.01.016).
- Dask Development Team (2016). *Dask: Library for dynamic task scheduling*. URL: <https://dask.org>.

- Davis, B. A. S., S. Brewer, A. C. Stevenson, and J. Guiot (2003). "The temperature of Europe during the Holocene reconstructed from pollen data". In: *Quat. Sci. Rev.* 22.15-17, pp. 1701–1716. ISSN: 02773791. DOI: [10.1016/s0277-3791\(03\)00173-2](https://doi.org/10.1016/s0277-3791(03)00173-2).
- Davis, B. A. S., M. Zanon, P. Collins, A. Mauri, J. Bakker, D. Barboni, A. Barthelmes, C. Beaudouin, A. E. Bjune, E. Bozilova, R. H. W. Bradshaw, B. A. Brayshay, S. Brewer, E. Brugia paglia, J. Bunting, S. E. Connor, J. L. de Beaulieu, K. Edwards, A. Ejarque, P. Fall, A. Florenzano, R. Fyfe, D. Galop, M. Giardini, T. Giesecke, M. J. Grant, J. Guiot, S. Jahns, V. Jankovska, S. Juggins, M. Kahrmann, M. Karpinska-Kolaczek, P. Kolaczek, N. Kuhl, P. Kunes, E. G. Lapteva, S. A. G. Leroy, M. Leydet, J. Guiot, S. Jahns, V. Jankovska, S. Juggins, M. Kahrmann, M. Karpinska-Kolaczek, P. Kolaczek, N. Kuehl, P. Kunes, E. G. Lapteva, S. A. G. Leroy, M. Leydet, J. A. L. Saez, A. Masi, I. Matthias, F. Mazier, V. Meltssov, A. M. Mercuri, Y. Miras, F. J. G. Mitchell, J. L. Morris, F. Naughton, A. B. Nielsen, E. Novenko, B. Odgaard, E. Ortú, M. V. Overballe-Petersen, H. S. Pardoe, S. M. Peglar, I. A. Pidék, L. Sadori, H. Seppä, E. Severova, H. Shaw, J. Swieta-Musznicka, M. Theuerkauf, S. Tonkov, S. Veski, W. O. van der Knaap, J. F. N. van Leeuwen, J. Woodbridge, M. Zimny, and J. O. Kaplan (2013). "The European Modern Pollen Database (EMPD) project". In: *Vegetation History and Archaeobotany* 22.6, pp. 521–530. ISSN: 0939-6314. DOI: [10.1007/s00334-012-0388-5](https://doi.org/10.1007/s00334-012-0388-5). URL: <http://link.springer.com/article/10.1007/s00334-012-0388-5>.
- Davis, Basil A. S. and Jed O. Kaplan (2017). *HORNET Holocene Climate Reconstruction for the Northern Hemisphere Extra-tropics*. SNF-Research-Plan. last accessed Jan, 30th, 2018. URL: <http://p3.snf.ch/project-169598#>.
- Fick, Stephen E. and Robert J. Hijmans (2017). "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas". In: *International Journal of Climatology* 37.12, pp. 4302–4315. DOI: [10.1002/joc.5086](https://doi.org/10.1002/joc.5086). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5086>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.5086>.
- Fyfe, Ralph M., Jacques-Louis de Beaulieu, Heather Binney, Richard H. W. Bradshaw, Simon Brewer, Anne Le Flao, Walter Finsinger, Marie-José Gaillard, Thomas Giesecke, Graciela Gil-Romera, Eric C. Grimm, Brian Huntley, Petr Kunes, Norbert Kühl, Michelle Leydet, André F. Lotter, Pavel E. Tarasov, and Spassimir Tonkov (2009). "The European Pollen Database: past efforts and current activities". In: *Vegetation History and Archaeobotany* 18.5, pp. 417–424. DOI: [10.1007/s00334-009-0215-9](https://doi.org/10.1007/s00334-009-0215-9).
- Giesecke, Thomas, Basil Davis, Simon Brewer, Walter Finsinger, Steffen Wolters, Maarten Blaauw, Jacques-Louis de Beaulieu, Heather Binney, Ralph M. Fyfe, Marie-José Gaillard, Graciela Gil-Romera, W. O. van der Knaap, Petr Kuneš, Norbert Kühl, Jacqueline F. N. van Leeuwen, Michelle Leydet, André F. Lotter, Elena Ortú, Malte Semmler, and Richard H. W. Bradshaw (2013). "Towards mapping the late Quaternary vegetation change of Europe". In: *Vegetation History and Archaeobotany* 23.1, pp. 75–86. DOI: [10.1007/s00334-012-0390-y](https://doi.org/10.1007/s00334-012-0390-y).
- Goring, S.J. (2019). *Bulk Baconizing*. <https://github.com/NeotomaDB/bulk-baconizing>. DOI: [10.5281/zenodo.2545891](https://doi.org/10.5281/zenodo.2545891).
- Guiot, Joel and Anne de Vernal (2011). "QSR Correspondence "Is spatial autocorrelation introducing biases in the apparent accuracy of palaeoclimatic reconstructions?" Reply to Telford and Birks". In: *Quaternary Science Reviews* 30.21, pp. 3214–3216. ISSN: 0277-3791. DOI: [10.1016/j.quascirev.2011.07.023](https://doi.org/10.1016/j.quascirev.2011.07.023). URL: <http://www.sciencedirect.com/science/article/pii/S0277379111002344>.
- Haslett, John and Andrew Parnell (2008). "A simple monotone process with application to radiocarbon-dated depth chronologies". In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 57.4, pp. 399–418. DOI: [10.1111/j.1467-7187.2008.00650.x](https://doi.org/10.1111/j.1467-7187.2008.00650.x).

- 9876.2008.00623.x. eprint: <https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9876.2008.00623.x>. URL: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9876.2008.00623.x>.
- Hoyer, S. and J. Hamman (2017). “xarray: N-D labeled arrays and datasets in Python”. In: *Journal of Open Research Software* 5.1. DOI: [10.5334/jors.148](https://doi.org/10.5334/jors.148). URL: <http://doi.org/10.5334/jors.148>.
- Ivanovic, R. F., L. J. Gregoire, M. Kageyama, D. M. Roche, P. J. Valdes, A. Burke, R. Drummond, W. R. Peltier, and L. Tarasov (2016). “Transient climate simulations of the deglaciation 21–9 thousand years before present (version 1) – PMIP4 Core experiment design and boundary conditions”. In: *Geosci. Model Dev.* 9.7, pp. 2563–2587. DOI: [10.5194/gmd-9-2563-2016](https://doi.org/10.5194/gmd-9-2563-2016). URL: <https://www.geosci-model-dev.net/9/2563/2016/>.
- Jones, Eric, Travis Oliphant, Pearu Peterson, et al. (2001). *SciPy: Open source scientific tools for Python*. [Online; accessed 2017-02-18]. URL: <http://www.scipy.org/>.
- Juggins, Steve (2013). “Quantitative reconstructions in palaeolimnology: new paradigm or sick science?” In: *Quaternary Science Reviews* 64, pp. 20–32. ISSN: 0277-3791. DOI: [10.1016/j.quascirev.2012.12.014](https://doi.org/10.1016/j.quascirev.2012.12.014). URL: <http://www.sciencedirect.com/science/article/pii/S0277379112005422>.
- (2017). *rioja: Analysis of Quaternary Science Data*. R package version 0.9-21. URL: <http://www.staff.ncl.ac.uk/stephen.juggins/>.
- Kageyama, Masa, Pascale Braconnot, Sandy P. Harrison, Alan M. Haywood, Jo-hann H. Jungclaus, Bette L. Otto-Bliesner, Jean-Yves Peterschmitt, Ayako Abe-Ouchi, Samuel Albani, Patrick J. Bartlein, Chris Brierley, Michel Crucifix, Aisling Dolan, Laura Fernandez-Donado, Hubertus Fischer, Peter O. Hopcroft, Ruza F. Ivanovic, Fabrice Lambert, Daniel J. Lunt, Natalie M. Mahowald, W. Richard Peltier, Steven J. Phipps, Didier M. Roche, Gavin A. Schmidt, Lev Tarasov, Paul J. Valdes, Qiong Zhang, and Tianjun Zhou (2018). “The PMIP4 contribution to CMIP6 – Part 1: Overview and over-arching analysis plan”. In: *Geoscientific Model Development* 11.3, pp. 1033–1057. DOI: [10.5194/gmd-11-1033-2018](https://doi.org/10.5194/gmd-11-1033-2018). URL: <https://www.geosci-model-dev.net/11/1033/2018/>.
- Lamb, Henry F. and Sander van der Kaars (1995). “Vegetational response to Holocene climatic change: pollen and palaeolimnological data from the Middle Atlas, Morocco”. In: *The Holocene* 5.4, pp. 400–408. DOI: [10.1177/095968369500500402](https://doi.org/10.1177/095968369500500402). eprint: <https://doi.org/10.1177/095968369500500402>. URL: <https://doi.org/10.1177/095968369500500402>.
- Mauri, A., B. A. S. Davis, P. M. Collins, and J. O. Kaplan (2015). “The climate of Europe during the Holocene: a gridded pollen-based reconstruction and its multi-proxy evaluation”. In: *Quat. Sci. Rev.* 112, pp. 109–127. ISSN: 0277-3791. DOI: [10.1016/j.quascirev.2015.01.013](https://doi.org/10.1016/j.quascirev.2015.01.013). URL: <http://www.sciencedirect.com/science/article/pii/S0277379115000372>.
- Nychka, Douglas, Reinhard Furrer, John Paige, and Stephan Sain (2017). *fields: Tools for Spatial Data*. R package version 9.8-3. Boulder, CO, USA: University Corporation for Atmospheric Research. DOI: [10.5065/d6w957ct](https://doi.org/10.5065/d6w957ct). URL: <https://github.com/NCAR/Fields>.
- Oliphant, Travis E (2006). *A guide to NumPy*. Vol. 1. Trelgol Publishing USA. URL: <http://www.numpy.org/>.
- Otto-Bliesner, B. L., P. Braconnot, S. P. Harrison, D. J. Lunt, A. Abe-Ouchi, S. Albani, P. J. Bartlein, E. Capron, A. E. Carlson, A. Dutton, H. Fischer, H. Goelzer, A. Govin, A. Haywood, F. Joos, A. N. LeGrande, W. H. Lipscomb, G. Lohmann, N. Mahowald, C. Nehrbass-Ahles, F. S. R. Pausata, J.-Y. Peterschmitt, S. J. Phipps, H. Renssen, and Q. Zhang (2017). “The PMIP4 contribution to CMIP6 – Part 2:

- Two interglacials, scientific objective and experimental design for Holocene and Last Interglacial simulations". In: *Geosci. Model Dev.* 10.11, pp. 3979–4003. DOI: [10.5194/gmd-10-3979-2017](https://doi.org/10.5194/gmd-10-3979-2017). URL: <https://www.geosci-model-dev.net/10/3979/2017/>.
- Peltier, W. R., D. F. Argus, and R. Drummond (2015). "Space geodesy constrains ice age terminal deglaciation: The global ICE-6G\_C (VM5a) model". In: *Journal of Geophysical Research: Solid Earth* 120.1, pp. 450–487. DOI: [10.1002/2014jb011176](https://doi.org/10.1002/2014jb011176). eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014JB011176>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JB011176>.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Servén, Daniel, Charlie Brummitt, and Hassan Abedi (2018). *Dswah/Pygam*: V0.8.0. DOI: [10.5281/zenodo.1476122](https://doi.org/10.5281/zenodo.1476122).
- Sommer, Philipp S., Basil A. S. Davis, Manuel Chevalier, Jian Ni, and John Tipton (2019). "The HORNET project: applying 'big data' to reconstruct the climate of the Northern Hemisphere during the Holocene". In: *20th Congress of the International Union for Quaternary Research (INQUA)*. International Union for Quaternary Research. URL: <https://app.oxfordabstracts.com/events/574/program-app/submission/94623>.
- Telford, R. J., E. Heegaard, and H. J. B. Birks (2004). "All age-depth models are wrong: but how badly?" In: *Quaternary Science Reviews* 23.1, pp. 1–5. ISSN: 0277-3791. DOI: [10.1016/j.quascirev.2003.11.003](https://doi.org/10.1016/j.quascirev.2003.11.003). URL: <http://www.sciencedirect.com/science/article/pii/S0277379103003160>.
- Telford, R.J. and H.J.B. Birks (2005). "The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance". In: *Quaternary Science Reviews* 24.20, pp. 2173–2179. ISSN: 0277-3791. DOI: [10.1016/j.quascirev.2005.05.001](https://doi.org/10.1016/j.quascirev.2005.05.001). URL: <http://www.sciencedirect.com/science/article/pii/S027737910500168X>.
- Telford, R.J. and H.J.B. Birks (2009). "Evaluation of transfer functions in spatially structured environments". In: *Quaternary Science Reviews* 28.13, pp. 1309–1316. ISSN: 0277-3791. DOI: [10.1016/j.quascirev.2008.12.020](https://doi.org/10.1016/j.quascirev.2008.12.020). URL: <http://www.sciencedirect.com/science/article/pii/S0277379108003806>.
- Trachsel, Mathias and Richard J Telford (2016). "All age-depth models are wrong, but are getting better". In: *The Holocene* 27.6, pp. 860–869. DOI: [10.1177/0959683616675939](https://doi.org/10.1177/0959683616675939). eprint: <https://doi.org/10.1177/0959683616675939>. URL: <https://doi.org/10.1177/0959683616675939>.
- Williams, John W., Eric C. Grimm, Jessica L. Blois, Donald F. Charles, Edward B. Davis, Simon J. Goring, Russell W. Graham, Alison J. Smith, Michael Anderson, Joaquin Arroyo-Cabralles, Allan C. Ashworth, Julio L. Betancourt, Brian W. Bills, Robert K. Booth, Philip I. Buckland, B. Brandon Curry, Thomas Giesecke, Stephen T. Jackson, Claudio Latorre, Jonathan Nichols, Timshel Purdum, Robert E. Roth, Michael Stryker, and Hikaru Takahara (2018). "The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource". In: *Quaternary Research* 89.1, pp. 156–177. DOI: [10.1017/qua.2017.105](https://doi.org/10.1017/qua.2017.105).



# Appendices



# Todo list

█	rewrite the motivation – does not read well . . . . .	1
█	1: Add reference. . . . .	1
█	2: Add reference. . . . .	1
█	3: Add reference. <a href="https://pangaea.de/">https://pangaea.de/</a> . . . . .	1
█	4: Add reference. EMPD paper . . . . .	1
█	5: Add reference. ICON . . . . .	1
█	6: Add reference. POLNET-gridding paper . . . . .	1
█	7: Add reference. . . . .	2
█	8: Add reference. . . . .	2
█	9: Add reference. cite World bank report? . . . . .	2
█	10: Add reference. . . . .	2
█	11: Add reference. check these references! taken from Achilles PhD thesis, there might be better ones . . . . .	2
█	12: Add reference. Check these . . . . .	2
█	13: Add reference. check Walker et al., 2009 . . . . .	2
█	Add some background on the Holocene. How did it change (global mean temperature estimate?), how was the insolation? CO <sub>2</sub> effects, impact of the ice sheets during the early holocene, changes in altitude, large-scale atmospheric circulation, human influences. . . . .	2
█	14: Add reference. PMIP paper . . . . .	2
█	15: Add reference. check Wanner et al., 2008 . . . . .	2
█	16: Add reference. . . . .	3
█	17: Add reference. . . . .	3
█	18: Add reference. . . . .	3
█	19: Add reference. . . . .	3
█	20: Add reference. . . . .	3
█	21: Add reference. . . . .	3
█	22: Add reference. . . . .	3
█	23: Add reference. . . . .	3
█	24: Add reference. Don't know about Birks and Birks, 1980, took it from Manus review paper... . . . . .	3
█	25: Add reference. Manus review paper . . . . .	3
█	26: Add reference. Don't know about Wodehouse, 1935, took it from Manus review paper... . . . . .	3
█	27: Add reference. Other review papers? . . . . .	3
█	28: Add reference. add more..., Climate12K . . . . .	3
█	29: Add reference. cite some MAT papers . . . . .	3
█	30: Add reference. that North-US/South-US discrepancy... . . . . .	3
█	31: Add reference. . . . .	4
█	32: Add reference. cite some open-data publications . . . . .	4
█	33: Add reference. add more? . . . . .	5
	Figure: Visualize multiple grids on the same map, e.g. by using the grid spec- ifications from Treut et al., 2007. . . . .	6

34: Add reference. . . . .	6
Look into Dasgupta et al., 2016 . . . . .	6
35: Add reference. . . . .	7
36: Add reference. . . . .	7
37: Add reference. jupyter qtconsole . . . . .	7
6 that address two use-cases tackling the combination of observations and models . . . . .	10
Finally, in chapter 6 I investigate the question to what extent large-scale atmospheric circulation features can be estimated from proxy data. In this analysis I analyze the long-term stability of spatial correlation patterns between surface temperature and northern hemispheric teleconnections based on three Earth System Models (ESMs). . . . .	10
not sure if I include this here . . . . .	11
Need to write chapter 3 . . . . .	41
needs implementation . . . . .	45
Check this number . . . . .	49
Figure: POLNET pollen diagram . . . . .	50
Figure: POLNET Climate reconstructino . . . . .	50
Write gridding introduction . . . . .	87
Elevation figure has to be made nicer. . . . .	102
ice sheet mask has to be made nicer. . . . .	102
add DOI . . . . .	103
Need to write chapter A . . . . .	129
Need to write chapter C . . . . .	133

# Acronyms

- APD** African Pollen Database. 1, 4
- API** Application programming interface. 1, 2, 6
- CDO** Climate Data Operator. 1, 2
- CI** Continuous Integration. 1, 4
- EMIC** Earth System Model of Intermediate Complexity. 1
- EMPD** Eurasian Modern Pollen Database. 1, 4, 6
- EPD** European Pollen Database. 1, 6
- ESM** Earth System Model. 1, 2,
- FOSS** Free and Open-Source Software. 1
- GAM** Generalized Additive Model. 6
- GCM** General Circulation Model. 1
- GUI** graphical user interface. 1, 2
- JJA** summer (June, July and August). 6
- LAPD** Latin American Pollen Database. 1, 4
- LGM** Last Glacial Maximum. 1
- LIG** Last Interglacial. 1
- MAT** modern analogue technique. 6
- MCMC** Markov chain Monte Carlo. 6
- NAPD** North American Pollen Database. 1
- NCDC** National Climatic Data Center. 1
- NCO** netCDF Operator. 1
- NOAA** National Oceanic and Atmospheric Administration. 1
- PMIP** Paleoclimate Modelling Intercomparison Project. 1