BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 95 (2023) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Timon Netzel

QUANTITATIVE PALEOCLIMATE RECONSTRUCTIONS IN THE EUROPEAN REGION BASED ON MULTIPLE PROXIES

BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 95 (2023) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Timon Netzel

QUANTITATIVE PALEOCLIMATE RECONSTRUCTIONS IN THE EUROPEAN REGION BASED ON MULTIPLE PROXIES

Quantitative paleoclimate reconstructions in the European region based on multiple proxies

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Timon Netzel
aus
Hachenburg

Bonn, September 2022

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter:	Prof. Dr. Andreas Hense
Gutachter:	Prof. Dr. Thomas Litt
Tag der Promotion:	14.12.2022
Erscheinungsjahr:	2023

Abstract

The Earth's climate system is highly dimensional and has to be described with the help of probabilities. Such probability descriptions can be estimated from measured data. Relatively large-scale instrumental records, however, date back only about 150 years. Paleoclimatology studies the climate history of the Earth before the period of instrumental measurements. For this purpose, it uses so-called climate proxies, which indirectly provide information about past climate conditions. In order to produce quantitative paleoclimate reconstructions based on these data, several problems must be analyzed and solved in the course of this thesis.

First, we present a new age-depth/distance transformation in a Bayesian formulation by determining the uncertainty information of depths in sediments and distances in speleothems at a given age. This allows us to perform a data-driven transformation to past ages that behaves like a convolution with different kernel smoothers and avoids too much certainty in the statements about these time periods. Another result of this technique is the determination of the age resolution and its projection onto a regular grid. Thus, multiple proxies can be linked in time and spectral analyses can be performed.

Furthermore, we introduce a new way to establish transfer functions that map climate variables to plant distributions. This includes consideration of various machine learning algorithms for solving the classification problem of taxa absence and presence, taking into account uncertainties in the proxy-climate relationship. For the models and plant distributions used in this work, a simple feedforward neural network with one hidden layer wins in 70 % of the cases.

Based on our age-depth/distance transformation and transfer functions, we formulate a new Bayesian Hierarchical Model that produces local paleoclimate reconstructions. This considers various proxy sources such as plant data from lake and mire sediments, isotopic information from speleothems, marine sediments, and ice cores. These are studied not only in temporal space, but also in spectral space using wavelet power spectra. Such a comprehensive use of the spectral behavior of proxy information is possible due to the new age-depth/distance transformation and has therefore not been performed before. In addition, a priori information on the actual climate distribution in specific time periods are incorporated as further constraints. To solve the local reconstruction model, we use two different Markov chain Monte Carlo sampling methods called Metropolis-within-Gibbs and random walk Metropolis-Hastings. During the inference processes, our new method generates taxa weights that provide information about their importance to each site. As a result, over 600 sites in Europe, Northwest Africa, Anatolia, and the Levant are being processed, resulting in final 186 accepted local paleoclimate reconstructions. They show not only small-scale climate changes, which can be identified as Bond, Heinrich, and Dansgaard-Oeschger events, but also large-scale variations such as the last deglacialization and various glacial-interglacial cycles. Human influence on plant information in the lake and mire sediments studied affects our local reconstructions, which can be minimized to some extent with our new method by paying more attention to isotope-based proxies during the inference

process.

Finally, these local paleodata are summarized using spatial reconstruction methods over the European region. In this context, Earth System Models from CMIP5/PMIP3 experiments are linked to our local information, producing assimilated spatial climate fields. The results compared to present-day conditions for the Middle Holocene show warming in northeastern Europe and cooling in southern Europe. The reconstructed annual precipitation indicates an increase in northeastern Europe, a decrease in western Europe, and an increase in the Eastern Mediterranean and Levant. Compared to present-day conditions, the spatial reconstruction of the Last Glacial Maximum reveals a general cooling and an increase in precipitation in the west of the Iberian Peninsula and in northwest Africa, Anatolia, and the Levant. The latter shows a dipole structure with higher precipitation in the southern Levant.

Contents

CHAPTER 1

Introduction

The Earth's climate system is a complex system consisting of the interacting subsystems atmosphere, ocean, cryosphere, lithosphere, and biosphere. External processes such as plate tectonics, solar activity, volcanism, meteorites, orbital parameters, and greenhouse gases play a crucial role (Bradley, 2014). These are on a time scale of $10 - 10^7$ years. In addition, spatio-temporal climate variability, such as El Niño Southern Oscillation (ENSO), Atlantic Meridional Overturning Circulation (AMOC), and North Atlantic Oscillation (NAO), is determined by nonlinear interactions between the subsystems. These are on a time scale of less than 10 years. The description of such a high-dimensional system cannot be deterministic, but must be based on probabilities (Hense, 2005). Such probabilistic descriptions need to be estimated from measured data, which requires certain model assumptions such as ergodicity or transitivity. Relatively large-scale instrumental records, however, date back only about 150 years. In order to make statements about older climate variations, we depend on additional data.

Paleoclimatology studies the climate history of the Earth before the period of instrumental measurements (Bradley, 2014). For this purpose, it uses so-called climate archives or climate proxies. These come from natural sources such as tree rings, stalagmites, ice cores, corals, lake or marine sediments, pollen, or historical records (Bradley, 2014). As climate affects each of these archives, they indirectly provide information about past climatic conditions. Fig. 1.1 shows time series of some common proxies. As part of this, information from archives such as ice cores (Panels B and E), marine sediments (Panel G), and pollen data from lake and mire sediments (Panels C, D, F) are presented. We also see some correlation of variability among proxies in the last ca. 600 calibrated kiloanni before 1950 AD (cal ka BP). AD is used as an abbreviation for anno domini, and calibrated means the calibration of age measurements, e.g. based on radioisotopic methods. The local extreme values of the respective curves in Fig. 1.1 mark the glacial-interglacial cycles, i.e. the transition from warm to cold periods (PAGES, 2016). These are also apparent in the external process of solar radiation in panel A. Other smaller-scale climate variations can be seen in panel G, labeled with a combination of numbers

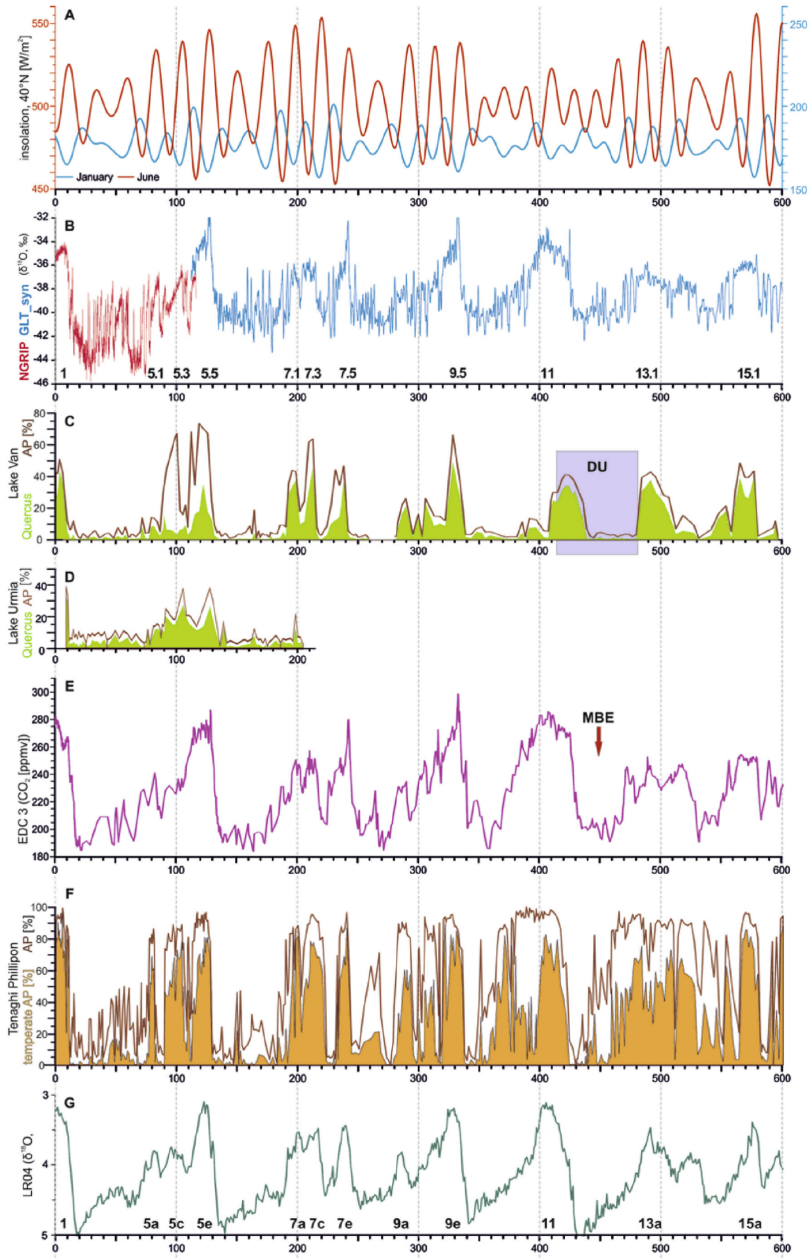


Figure 1.1: This figure is from Litt et al., 2014 and shows time series over the last 600 cal ka BP of different palaeoclimate proxies. Panel A displays solar radiation at 40 °N, B $\delta^{18}O_{VSMOW}$ from Greenland ice cores, E atmospheric CO_2 concentration from Antarctic ice cores, and G $\delta^{18}O_{VPDB}$ from marine sediment cores. In addition, panels C, D and F show the relative amount of arboreal pollen from Lake Van, Lake Urmia and Tenaghi Philippon, respectively.

and letters. These describe some of the so-called Marine Isotope Stages (MISs). In this context, stable oxygen isotopes in marine sediments are used to define further warm and cold periods (Lisiecki and

Raymo, 2005). MIS 1 marks what is known as the Holocene, the current geologic epoch or warm interglacial phase (Bradley, 2014). Between this and the last interglacial (MIS 5e) is the last glacial period. Even shorter climate fluctuations occur there, some of which are called Dansgaard-Oeschger (DO) events (Bond and Lotti, 1995). The corresponding colder phases within a glacial are called stadials and the warmer ones are referred to as interstadials (Pickarski et al., 2015a).

A major challenge in paleoclimatology is estimating the age of the respective proxies. These are determined using a variety of techniques, including radioisotopic methods, paleomagnetism, tephrochronology, and biological dating methods (Bradley, 2014). Age uncertainties play a large role, especially when time scales such as those in Fig. 1.1 are considered. By dealing appropriately with such probability descriptions, an overestimation of statements, e.g. about the timing of past climate changes, and thus misinterpretations can be avoided. Since there is not yet a simple solution to this problem, one task of this work will be to incorporate age uncertainties into paleoclimate reconstructions.

Another difficulty in paleoclimatology is obtaining quantitative climate information from proxies. For this purpose, so-called transfer functions are created, which are based on a calibration of the proxy data to the measurement data to be reconstructed (Kühl et al., 2002). When using multiple proxies, appropriate treatment of their respective uncertainties could lead to more accurate and less uncertain statements about past climate, as we distinguish between more and less informative proxies. There are a number of studies that describe different approaches to determining transfer functions (e.g. Peyron et al., 1998; Lauritzen and Lundberg, 1999; Kühl et al., 2002; Tremaine et al., 2011; Stolzenberger, 2017). To date, however, there is no systematic approach to identifying the most appropriate functions. Therefore, in this work we want to investigate whether something like this is feasible, which is also easy to implement and largely automated.

In Fig. 1.1 we see that the different climate archives show similar patterns to some extent. Therefore, it would be of great advantage to consider information from as many proxies as possible when reconstructing past climate. Furthermore, it may be beneficial to view such data not only in temporal space, but also in spectral space to more easily identify possible periods of past climate signals. Another advantage of spectral analyses is that misinterpretations due to the so-called source effect can be minimized for isotope-based proxies such as speleothems (Cheng et al., 2015). This effect describes the source from which the respective stable isotopes under consideration originate. So far, there have been no attempts to do this using the time series shown as examples. One of the reasons for this could be that such projects often require high computational power. Therefore, in this work, we aim to develop a computationally fast method that can consider multiple proxies simultaneously. It should also be flexible enough to reflect their changes over time in the context of their respective age uncertainties. Thus, it should be possible to consider not only relatively large-scale climate changes such as glacial-interglacial cycles or MISs, but also small-scale ones like the DO, Heinrich (H), and Bond events. The latter is the interglacial counterpart to the DO events and the Heinrich events

denote particularly pronounced stadials of the last glacial (Rasmussen et al., 2014). Another challenge in dealing with proxies is human-influenced signals that are not of climatic origin. If possible, a reconstruction method should be able to filter them out.

The main goal of this work is to develop an alternative way for calculating quantitative paleoclimate reconstructions that is more automated and relies on statistically based methods and less expert knowledge. This could provide new insights into the importance of the proxies studied and thus extend the knowledge from previous studies. Based on proxy information, there are a variety of qualitative climate reconstructions (e.g. Litt et al., 2014; Panagiotopoulos et al., 2013; Miebach et al., 2016; Schiebel and Litt, 2018; Panagiotopoulos et al., 2020). We want to use these as a comparison for our quantitative statements to show similarities and differences and to check whether our new approach provides realistic results.

In climatology, not only the local-temporal course but also the spatial distribution of climate variables is of interest. One of the main problems is the spatial interpolation of local paleoinformation. There are numerous attempts to do this, with varying advantages and disadvantages (e.g. Gebhardt et al., 2008; Mauri et al., 2015; Weitzel et al., 2019). In this work, we aim to use the new method of Weitzel et al., 2019 to produce physically consistent estimates of past climate state using Earth System Models (ESMs) in conjunction with local climate reconstructions. More specifically, we want

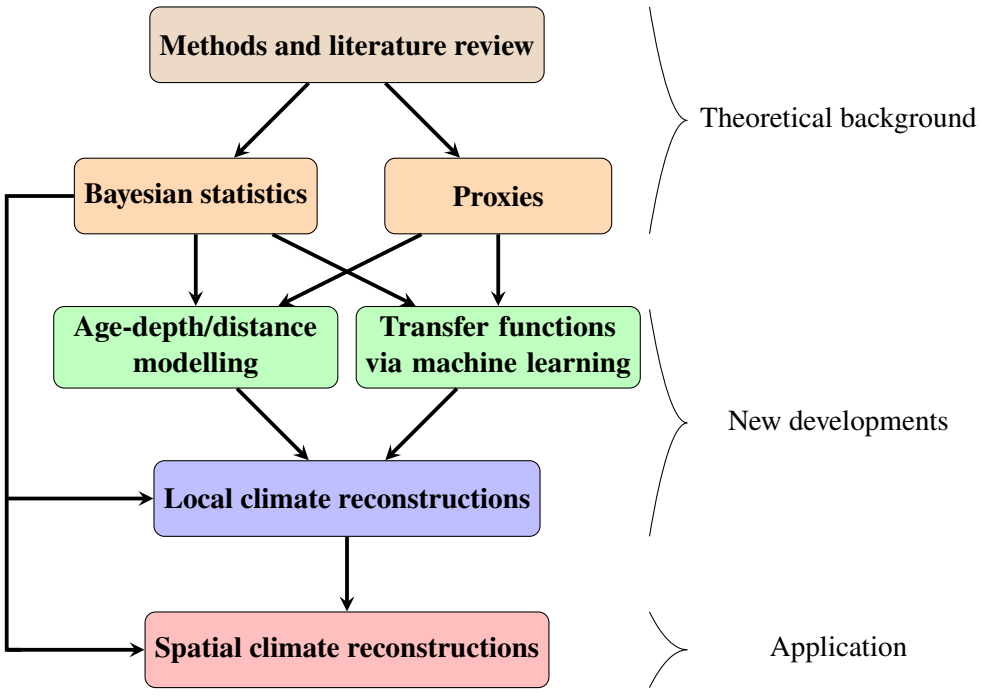


Figure 1.2: Flowchart of the structure of this work. The arrows indicate the reading direction of the diagram and the different work steps are marked by various colors. The comments on the right side summarize certain work steps.

to examine the time slices of the Middle Holocene (MH) around 6000 cal a BP and the Last Glacial Maximum (LGM) around 21000 cal a BP for the European region, North Africa, and the Levant. The periods of the MH and the LGM have already been investigated in some studies (e.g. Simonis et al., 2012; Bartlein et al., 2010; Mauri et al., 2015; Weitzel et al., 2019; Ludwig and Hochman, 2022), so we can compare our new results. In order to test the modified spatial reconstruction method of Weitzel et al., 2019, we also want to generate climate fields for the period $1961 - 1990$ AD. This falls within the historical (HIST) period and can be verified by corresponding climate observations.

The structure of this work is illustrated by the flowchart shown in Fig. 1.2. This includes the research tasks described above, which are marked with the colours light red, light blue and light green (from the bottom to the middle). Furthermore, the theoretical background of this work is presented as a starting point. Specifically, it describes the basics of the various climate proxies and their databases, and covers the principles of Bayesian statistics. With this knowledge, a new approach to age-depth/distance modelling can be developed and transfer functions can be created using various machine learning techniques. Based on these results and with the help of Bayesian statistics, it is possible to develop a new method for local climate reconstructions. There is not only a detailed section on this in the first chapter Methods and literature review, but also in the results of this thesis (see. Sect. 3.3). The situation is similar with spatial climate reconstructions (see. Sect. 3.4). Here, the previous results are combined with the method of Weitzel et al., 2019 to generate new spatial climate data.

CHAPTER 2

Methods and literature review

2.1 Proxies

Bernhardi, 1832 was the first to propose a Fennoscandian inland glaciation to explain the transport of the erratic blocks found in northern Germany. Building on this theory, the extent of the glaciers of the Last Glacial period in northern Europe and North America was determined in the following decades (Agassiz, 2012; Prest, 1990). Thus, for example, erratic blocks and end moraines are so-called palaeoclimatic proxies that can provide information about the past climate (Bubenzer and Radtke, 2007). However, such climate archives are quite crude and can only provide information on whether certain areas were glaciated or not. Over time, paleoclimate research has found other proxies that can provide much more detailed information about the climate of the past. Table 2.1 shows the most common climate archives with their characteristic properties. This includes their respective sampling intervals, temporal range and the potential information they can provide. For a more detailed description of the respective proxies, the reader is referred to Bradley, 2014. In this thesis, the focus is on the following climate archives: marine sediments, ice cores, speleothems, lake and mire sediments, and pollen. By looking at these independent sources together, the quantitative climate reconstructions carried out later should reduce the respective weaknesses to a minimum in order to provide a clearer picture of climate history. Each of them offers the possibility of drawing conclusions about the palaeoclimate in terms of temperature and precipitation. Moreover, climate reconstructions on the order of $\sim 10^4 - 10^6$ years are possible, which can cover several glacial cycles. The temporal resolution for these proxies ranges from year/season to 100 years.

Table 2.1: Characteristics of natural archives.
T: Temperature; P: Precipitation, humidity or water balance (P-E); C: chemical composition of air C_a or water C_w; B: information on biomass or vegetation patterns, V: volcanic eruptions; M: geomagnetic field variations; L: sea level; S: solar activity. After Bradley, 2014

Archive	Minimum sampling interval	Temporal range (order year)	Potential information derived
Historical records	Day/h	$\sim 10^3$	T,P,B,V,L,S
Tree rings	Year/season	$\sim 10^4$	T,P,B,V,S
Lake sediments	Year (varves) to 20 years	$\sim 10^4 - 10^6$	T,P,M,B,V,C_w
Corals	Year	$\sim 10^4$	L,T,P,C_w
Ice cores	Year/season	$\sim 10^6$	T,P,M,B,V,S,C_a
Pollen	20 years	$\sim 10^6$	T,P,B
Speleothems	Year	$\sim 5 \times 10^5$	C_w,T,P,V,B
Palaeosols	100 years	$\sim 10^6$	T,P,B
Loess	100 years	$\sim 10^6$	P,B,M
Geomorphic features	100 years	$\sim 10^6$	T,P,V,L
Marine sediments	100 years	$\sim 10^7$	T,C_w,B,M,L,P,S

2.1.1 Stable oxygen isotopes

In palaeoclimatology, oxygen isotopes are most commonly used (Bradley, 2014). They are found in ice cores, marine and lake sediments, and cave finds, among others, and provide information about past climate variability. These fluctuations include not only those with relatively long periods such as glacial-interglacial cycles (Shakun et al., 2015), but also more frequent ones. To obtain an appropriate isotopic signal, a standardized ratio, expressed in parts per thousand, is used that takes advantage of the different molecular weights of the inherent isotopes. Equation 2.1 shows this relationship for the heavier isotope ^{18}O and the lighter, more abundant ^{16}O (Dansgaard, 1964):

$$\delta^{18}O = \left(\frac{\left(\frac{^{18}O}{^{16}O}\right)_{sample}}{\left(\frac{^{18}O}{^{16}O}\right)_{standard}} - 1 \right) \cdot 1000. \tag{2.1}$$

The particular standard used depends on the context in which the isotope ratio is considered. For example, when it comes to snow, ice, rain, groundwater or seawater, Coplen, 1996 suggests the Vienna Standard Mean Ocean Water (VSMOW) as the standard. This refers to the current isotopic composition of the ocean. In case the oxygen present in calcium carbonate $CaCO_3$ is analyzed, the standard is the so-called Vienna Pee Dee Belemnite (VPDB). This name comes from a marine fossil from the Cretaceous period, which consisted of $CaCO_3$ (Ravelo and Hillaire-Marcel, 2007). These standards were established in 1968 by the International Atomic Energy Agency (IAEA) in Vienna to

ensure uniformity of regulation. This can be used, for example, to compare $\delta^{18}O_{VSMOW}$ in different ice deposits or $\delta^{18}O_{VPDB}$ in different stalagmites.

Marine and greenland-ice cores

The first drillings through the Greenland ice sheet took place in the 1960s (Peel, 2005). These were carried out by the US Army, which established bases in the north of Greenland during the Cold War. Even today, these cores are used by scientists to study the climate of the past. In doing so, they rely on the studies of Dansgaard, 1964, who was the first to show how oxygen isotopes can be used as a climate proxy. The cores studied not only show a glacial-interglacial cycle, but also significantly shorter climate fluctuations during the Last Glacial period. These are called Dansgaard-Oeschger events (Bond and Lotti, 1995). The cause of these abrupt climate changes is thought to be the varying extent of sea ice in the North Atlantic (Li et al., 2010). In the process, this freshwater supply influences the AMOC, which in turn determines the meridional heat transport of the ocean (Dokken et al., 2013). Other prominent climate events in the past discovered on the basis of proxy data are the Heinrich events (Heinrich, 1988). He found abrupt and cyclic advances of large amounts of rock that could only have been transported by glaciers of the Laurentide Ice Sheet (LIS) into the northeast Atlantic (ca. 47°N, 19°W). These then deposited in the corresponding ocean sediments, also known as ice rafted debris (IRD). A possible connection between DO and Heinrich events is seen in the fact that the latter are restricted to the stadial phases of DO events (Alvarez-Solas et al., 2011). Further study of marine sediments led Bond et al., 1997 to conclude that there were other climate fluctuations during the Holocene, known as Bond events. These are referred to by the authors as mini-DO events and are considered the interglacial counterpart to these events.

To obtain a climate signal from Eq. 2.1, it is necessary to clarify under what circumstances their relative occurrence changes. Fig. 2.1 shows a corresponding conceptual idea. At the sea surface, water with the lighter isotope ^{16}O evaporates more readily than water with ^{18}O (Dansgaard, 1964). As these air masses are transported toward land, precipitation predominantly releases the heavier isotope ^{18}O to seawater (Fig. 2.1, upper panel). The result of this so-called fractionation is that ^{18}O is now mainly present in the ocean. When climatic conditions are relatively warm and large ice masses are reduced on land, the distribution of these two isotopes balances out again. The precipitation, consisting mainly of ^{16}O, can thus ultimately run off again from the land into the sea. For example, during the Last Glacial, large ice masses were able to spread over large parts of northern Europe and North America (Prest, 1990). These are enriched in the lighter isotope ^{16}O, as shown in Fig. 2.1 (lower panel). This preserves the fractionation of both isotopes relative to the warm period described above. The whole process is also called the ice volume effect.

We can see this in Fig. 2.2 for 3 very well-known $\delta^{18}O$-curves. Panel (a) shows the data from the so-called North Greenland Ice Core Project (NGRIP) (Andersen et al., 2004). This dates back to the last warm period, which was characterized by an increase in $\delta^{18}O$ about 120,000 years before today. In

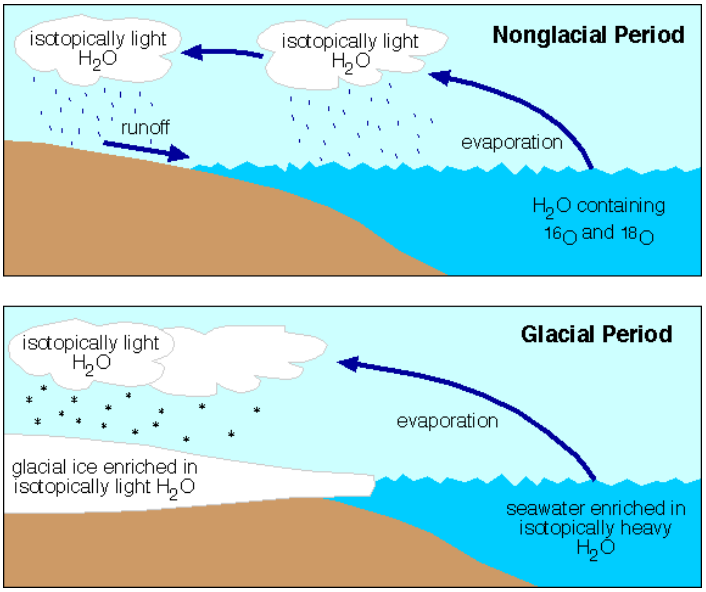


Figure 2.1: This figure is from DiVenere, 2017 and represents a conceptual model of the fractionation of lighter and heavier oxygen isotopes under glacial and non-glacial conditions.

addition, strong fluctuations can be observed during the Last Glacial period, which represent the DO cycles mentioned above. The described fluctuations are detectable due to the ice volume effect. Using $\delta^{18}O$ from Greenland ice, the mean annual temperature (T_{ANN}) can be derived (Bradley, 2014). Panel (b) shows the isotope profile of the calcium carbonate-containing shells of the planktonic foraminifera (Ravelo and Hillaire-Marcel, 2007). The term Medstack means that data from several sediment cores studied in the Mediterranean were combined (Wang et al., 2010). An analogous approach was taken by Lisiecki and Raymo, 2005, who determined the isotope profile shown in (c) using 52 sediment cores distributed worldwide. Most of the cores come from the Atlantic, but also from the Pacific and Indian Ocean. In contrast to Medstack, they studied benthic foraminifera, i.e. those that occur on the deep sea floor. The $\delta^{18}O$ of these organisms is a function of the temperature and the $\delta^{18}O$ of the water, which in turn is influenced by the ice volume and salinity (Ravelo and Hillaire-Marcel, 2007). Thus, higher uncertainties must be taken into account here if the atmospheric T_{ANN} is to be reconstructed from the curves (b) and (c). It is also noticeable that the ocean data are anti-correlated with the ice core data. This is a consequence of the fractionation described above. If a certain isotope is less abundant in the ocean, it is correspondingly more abundant in the Greenland ice. Nevertheless, the broad characteristics are similar to those of NGRIP. Here, too, it is easy to distinguish between warm and cold periods. However, due to the coarser resolution and bioturbation of the curves in (b) and (c), it is more difficult to detect shorter climate fluctuations such as the DO or Heinrich events (Bradley, 2014).

As we have seen, paleotemperatures have already been reconstructed from the above isotopic data.

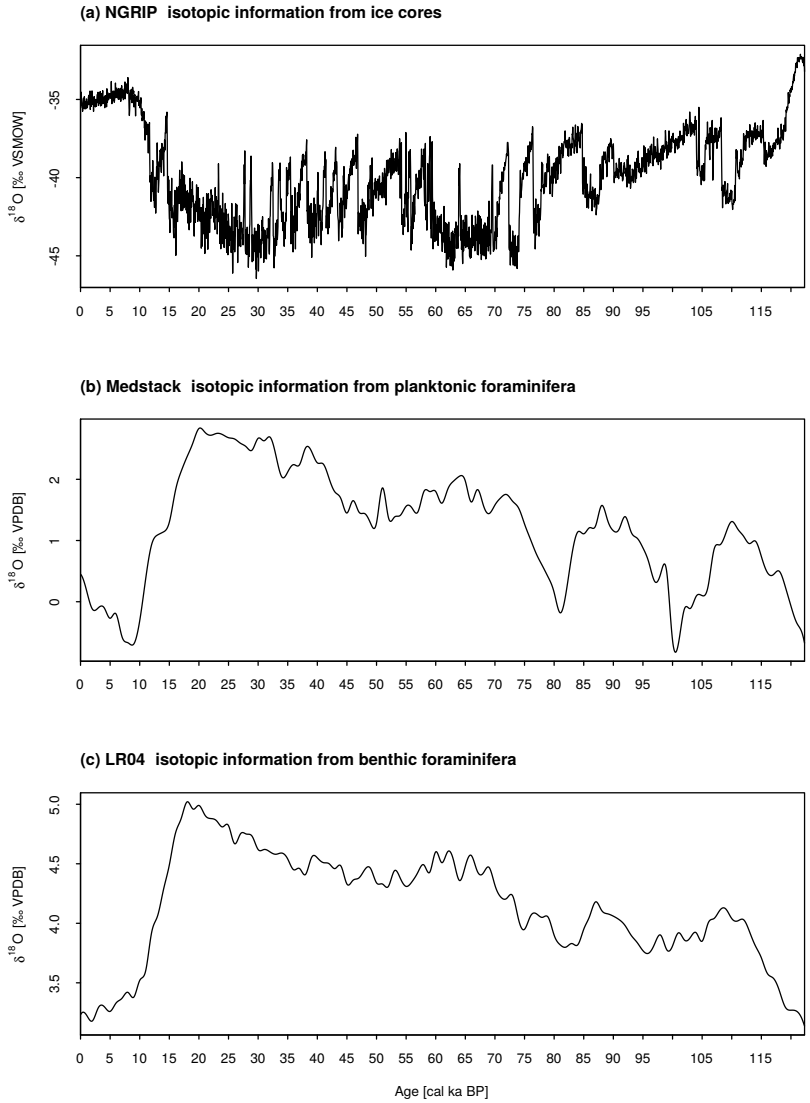


Figure 2.2: Shown are 3 time series based on isotope data from different sources. Panel (a) depicts δ^{18}O of ice cores from the North Greenland Ice Core Project. In contrast, panel (b) is based on isotopes of planktonic foraminifera from cores in the Mediterranean Sea, the so-called Medstack. δ^{18}O of LR04 in (c) are from benthic foraminifera from marine cores distributed worldwide.

Although Table 2.1 shows that these proxies also contain the potential for precipitation information, we will only reconstruct it indirectly using these isotopes. That is, we will later (see Sect. 2.5) compare the curves from Fig. 2.2 with the temperatures reconstructed in this work and link them to precipitation information in a further reconstruction step.

A common approach to visualize the climate variations described above is the analysis of so-called wavelet power spectrum (WPS). Here, time series in different frequencies/periods are analyzed with

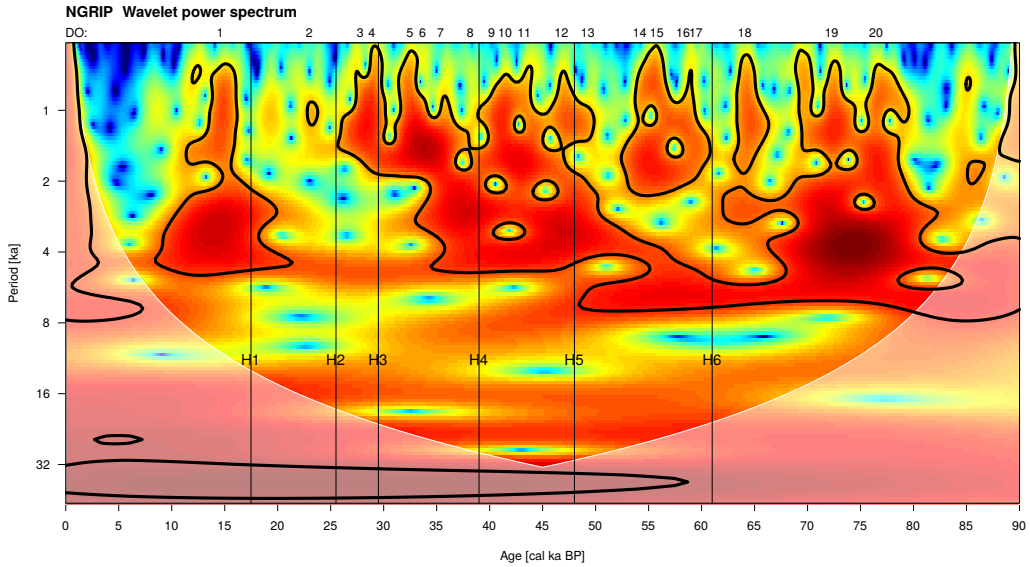


Figure 2.3: Wavelet power spectrum based on the time series in Fig. 2.2 (a). The vertical lines indicate the timing of Heinrich events $1-6$, and DO events $1-20$ are marked above the graph. Areas within the thicker black lines represent significant power and the not grayed out part, the regions within the so-called cone of influence.

respect to their time dimension. An example of this is shown in Fig. 2.3. This WPS is based on the first 90 cal ka BP of the NGRIP time series from Fig. 2.2 (a). On the y-axis, the periods of these data are plotted against time on the x-axis. The redder the colors displayed, the greater the variability in the time series, i.e. the power of the wavelet spectrum. The opposite is true for the blue colors. For simplicity, the corresponding normalized color bar is not shown. We see in this WPS that the DO events $1-20$ occur with relatively high and significant power in periods between 1 and 2 ka. In addition, Heinrich events $1-6$ are intermediate between some DO events because they occur at their longer stadials (Rasmussen et al., 2014). The technical background of the calculation of such a WPS is described in the following paragraph.

Time series analysis Since a large number of time series are analyzed in this thesis, the mathematical convolution and some of its applications are briefly described. The convolution is defined as the integral of the product of the two functions f and g after one of them has been inverted and shifted:

$$(f * g)(x) := \int_{\mathbb{R}} f(\tau)\overline{g}(x-\tau)\mathrm{d}\tau. \tag{2.2}$$

The bar denotes the complex conjugate. Here, for the mean value $f * g$, the function value $f(\tau)$ is weighted with $g(x-\tau)$. This form of superposition can be used, among other things, to form a moving

average. We see that $(f * g)(x)$ depends only on x. The initial functions contain the information of τ on the one hand and the information of the relation between τ and x on the other hand. Convolution is used in the fields of probability, statistics and signal analysis, among others (Honerkamp, 2002).

In this work, a particular interest is the temporal changes of the frequency spectra of a time series. Therefore, the so-called wavelet transformation is used. In this method, the output signal is convolved with scaled and translated forms of a temporally localized wave function, the so-called mother wavelets (Debret et al., 2007). We use the well-known Morlet wavelet, which represents a Gaussian-modulated sine wave

$$\psi_\eta(t) = \pi^{-\frac{1}{4}}(e^{i\eta t} - e^{-\eta^2/2})e^{-t^2/2}, \tag{2.3}$$

with the reference frequency η. From this mother wavelet we can derive a family of so-called daughter wavelets with the help of

$$\psi_{ab}(t) = \frac{1}{a}\psi\left(\frac{t-b}{a}\right). \tag{2.4}$$

The translation parameter b scans the signal $s(t)$ in the time dimension and the scaling parameter a over different frequency ranges. The continuous wavelet transform can be described using the convolution from equation 2.2 as follows:

$$\mathcal{W}_\psi(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}\overline{\psi}\left(\frac{t-b}{a}\right)s(t)\,dt. \tag{2.5}$$

The WPS shown in Fig. 2.7 is finally calculated with

$$\mathrm{WPS} = \left|\mathcal{W}_\psi(a,b)\right|^2. \tag{2.6}$$

More specifically, in this work, the discretized forms of the above equations are used to perform wavelet-based time series analysis. The shaded area in the WPS of Fig. 2.3 describes the region where the boundary effect can no longer be neglected. This happens when stretched wavelets go beyond the boundaries of the observation interval (Torrence and Compo, 1998). The unshaded area, called the cone of influence (COI), shows an accurate time-frequency representation of the data. Outside this region, the patterns should be considered suspect because of possible boundary effects.

Speleothems

Moore, 1952 coined the term speleothem, by which he meant deposits in caves such as stalagmites, stalactites, and flowstones. They are mostly composed of the chemical compound calcium carbonate $CaCO_3$ (Hendy, 1971), which is formed by a series of chemical reactions. As can be seen in Fig. 2.4, the water H_2O in the soil zone above the cave combines with carbon dioxide CO_2 to form acidic water H_2CO_3. This continues to react with calcium carbonate in the karst zone to form $Ca^{2+} + 2\,HCO_3^-$. The dissolved carbon dioxide in this solution may degas at a lower CO_2 concentration within the

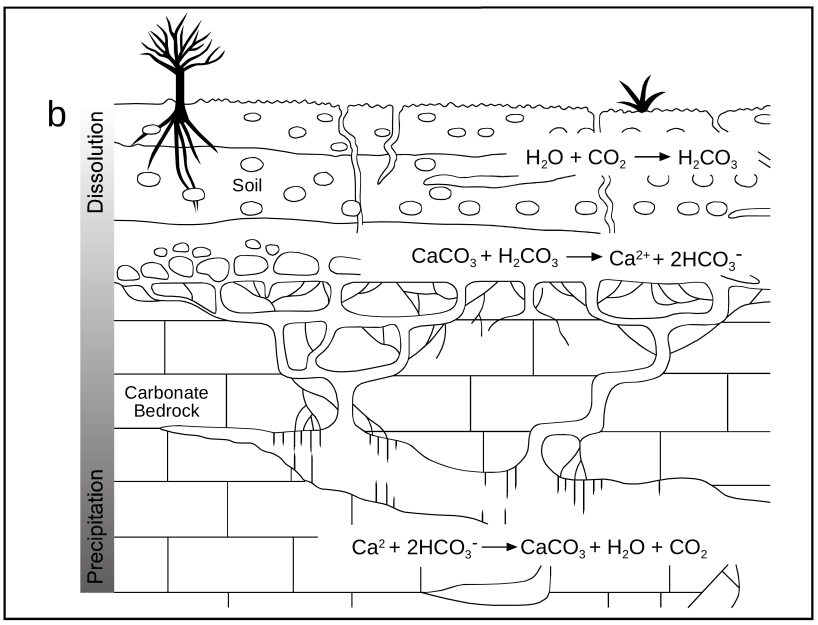


Figure 2.4: This figure is from Fairchild et al., 2006 and shows the dissolution and precipitation regimes of the karst system.

cave. The resulting supersaturation leads to the precipitation of calcium carbonate, which provides the components for speleothems.

The whole process reveals that the molecules of calcium carbonate have different sources. The calcium comes from the karst zone. Carbon originates from the overlying soil zone and reflects surface vegetation activity (Bar-Matthews et al., 1997). However, only part of the oxygen stems from the soil zone. Another source is water infiltrating into the ground, mainly from precipitation events. Thus, the calcium carbonate of speleothems only partially reflects local conditions. It is not only the source of precipitation that is important. Another factor are processes in the atmosphere that influence the oxygen in the water as the air masses are transported to the cave. In Fig. 2.5 the way of δ^{18}O from the ocean to the cave entrance is depicted. The $\delta^{18}O_{vapor}$ reflects that of water and Sea Surface Temperature (SST), among others. During transport to the cave location, the following factors affect $\delta^{18}O_{cloud}$: altitude, longitude, latitude, distance from the source, amount of precipitation, and air temperature (McDermott et al., 1999).

Hendy, 1971 describes how speleothems can be used as paleoclimate archives. If the δ^{18}O of the calcite is in isotopic equilibrium with the δ^{18}O of the drip water, a relationship can be found between these two ratios, which in turn is temperature-dependent (Kim and O'Neil, 1997). Therefore, isotopic information from speleothems may allow inferences about cave temperature, which is related to mean annual surface temperature due to ventilation effects (Lauritzen and Lundberg, 1999; Sánchez-Fernández et al., 2018; Bradley, 2014). This idea of a transfer function must then be calibrated with data from the appropriate caves, which Tremaine et al., 2011 did for some caves in Florida and

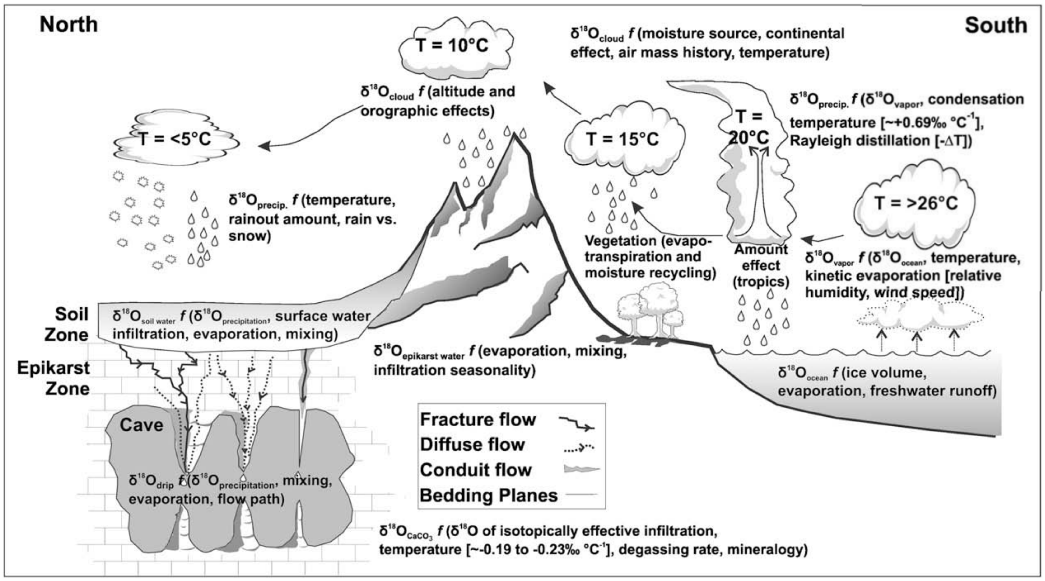


Figure 2.5: This figure is from Lachniet, 2009 and shows the complexity of the processes that determine the composition of δ^{18}O with respect to speleothems. In particular, fluctuations in temperature and relative humidity affect the δ^{18}O values due to different processes and phase changes in the atmosphere, ocean, hydrosphere, soil, and epikarst zones, and finally in the caves.

compared them with older calibrations like in, e.g. Kim and O'Neil, 1997. If a quantitative climate reconstruction is to be performed, it is necessary to take into account the ice volume effect mentioned above. An example of this can be found in Lauritzen and Lundberg, 1999, where temperature during the Holocene is reconstructed using a speleothem in Norway. Overall, only a few quantitative climate reconstructions based on speleothems have been carried out so far, which Lachniet, 2009 and McDermott et al., 2011 attribute to the complexities of the hydrologic water cycle shown in Fig. 2.5.

Since the last decades, the number of speleothems studied worldwide has increased significantly (Atsawawaranunt et al., 2018) and a large part of them has been made easily accessible through the speleothem database **S**peleothem **I**sotopes **S**ynthesis and **A**na**l**ysis (SISAL). As a result, a huge amount of proxy information can be easily viewed for specific areas, such as great parts of Europe (Lechleitner et al., 2018) or China (Zhang et al., 2019). Other advantages include the precise datability of speleothems (Richards and Dorale, 2003) and their high density of proxy material (Treble et al., 2007). For these reasons, it is advisable to consider this wealth of information in climate reconstructions.

Particularly striking is the simultaneous occurrence of prominent climatic events. For example, Genty, 2003 describe speleothems from Villars Cave in France and Soreq Cave in Israel that show similar cycles to the oceanic sediments and ice cores of Greenland during the Last Glacial (see Fig. 2.2). The δ^{18}O curve of the Soreq cave is shown in the Fig. 2.6. Bar-Matthews et al., 1999 infer from these data some Heinrich events, the last deglaciation and the Younger Dryas (YD), characterized by cold and dry climatic conditions in the Eastern Mediterranean. At Jeita Cave in Lebanon, Cheng

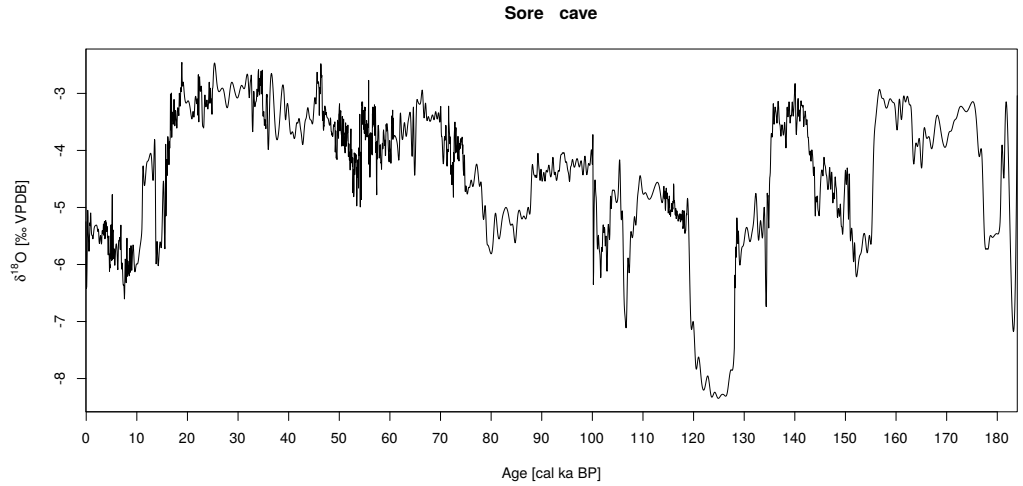


Figure 2.6: Time series of δ^{18}O of a speleothem from Soreq cave, Israel.

et al., 2015 compare their data with these from Israel and find similarities with respect to the Bond events, the Dryas events, the Bølling-Allerød (BA) interstadial, and the Heinrich stadials. To detect events during the Holocene, they used spectral analysis. In addition, Bar-Matthews et al., 2003 describe that MIS 7 – 7.5 are detectable in Soreq and Pequin caves. Bar-Matthews et al., 1997 and Bar-Matthews et al., 2003 find DO cycles not only in speleothems but also in marine cores of the Eastern Mediterranean. Moreover, these cycles appear to be globally synchronized overall during the Last Glacial (Bar-Matthews, 2013). Some caves in China, for example, show spectral patterns during this period and the Holocene that are consistent with those from the North Atlantic and thus have an influence on the Asian summer monsoon (ASM) (Wan et al., 2011; Cheng et al., 2015). Fleitmann et al., 2003 also find a link between Greenland ice cores during the Holocene with speleothems from Oman. The 8.2 event is found, for example, in Romania (Constantin et al., 2007), Norway (Lauritzen and Lundberg, 1999), or in various caves in Germany (Waltgenbach et al., 2020). Poleva Cave in Romania (Constantin et al., 2007) additionally shows a correlation with Villars and Soreq Caves and other Bond events such as the 4.2 and 3.2 events. The latter is also detectable in the Ascunsă Cave, also located in Romania (Drăguşin et al., 2014). Furthermore, Budsky et al., 2019 find evidence for DO cycles during MIS 3 – 5e in the western Mediterranean. They associate these DO events with wetter climatic conditions in much of Western Europe. Many of these aforementioned events are associated with AMOC disturbances and IRD. For example, Deininger et al., 2016 demonstrate a connection between 11 speleothems over Europe and important North Atlantic Ocean circulations for the late Holocene. Demény et al., 2021 also reveal a correlation with the SST of the North Atlantic and thus with the AMOC during the mid to late Holocene. Their argument is based on a principal component analysis (PCA) based on speleothem data. In Niggemann et al., 2003, spectral analysis is used to

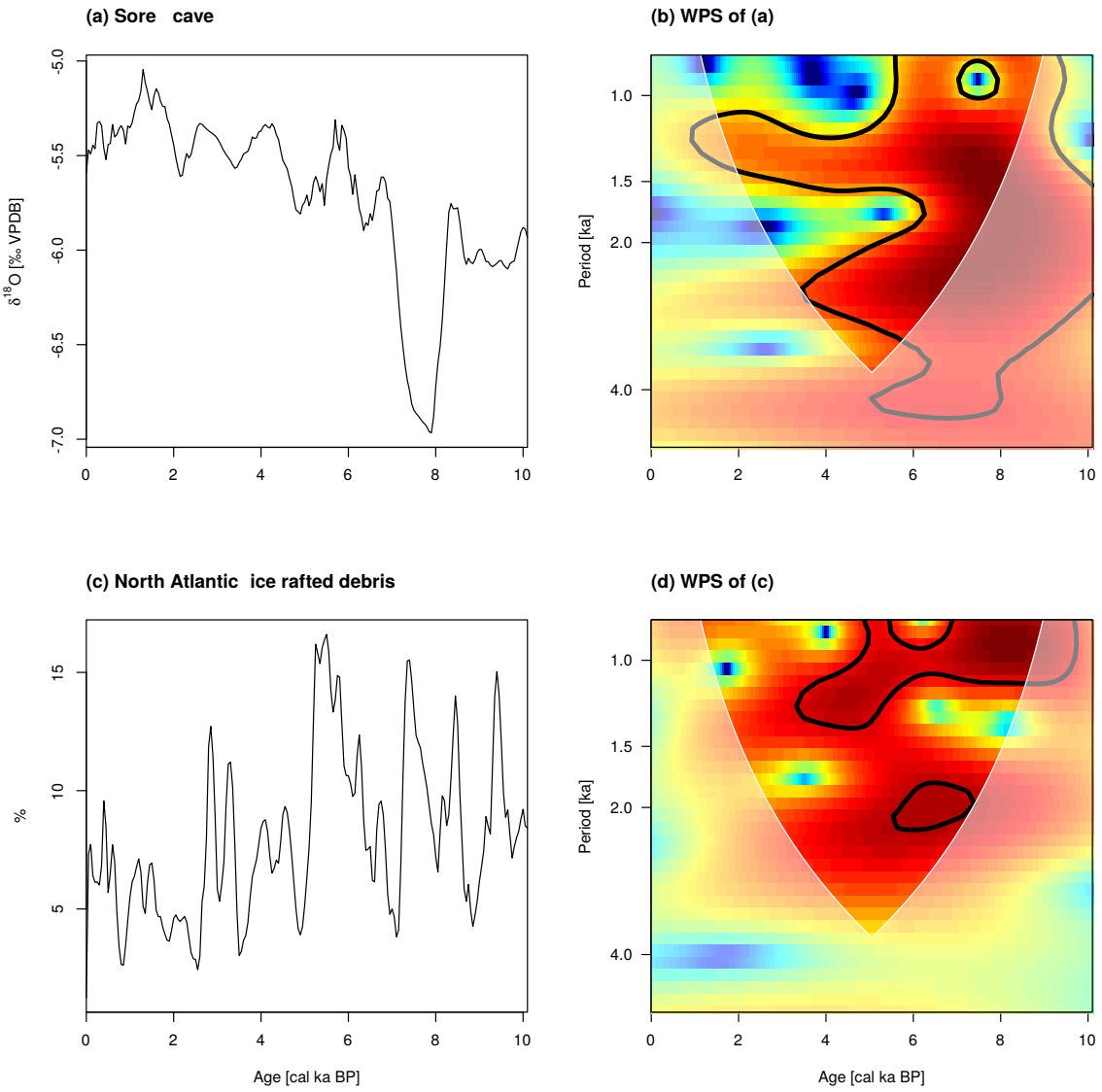


Figure 2.7: The top row shows in (a) the time series of δ^{18}O from a speleothem of Soreq Cave and in (b) the corresponding wavelet power spectrum. The lower row is analogous to the top, except that the time series is based on the percentage of ice rafted debris from the North Atlantic.

detect Bond events in Germany, with a power peak at about 1500 years. In addition, Bar-Matthews and Ayalon, 2007 describe the period 8.5 – 7 ka BP shown in Fig. 2.6 as a deluge period with a decrease in precipitation. Subsequently, the isotopic information indicates a climate more similar to the present one in the Eastern Mediterranean (Bar-Matthews et al., 1999). In Bar-Matthews and Ayalon, 2011, the periods between 7 – 4 ka BP are described in more detail. They also found cycles with a period of about 1.5 ka and associate them with some Bond events. Moreover, these events coincide with some

important cultural changes caused by major droughts in the Eastern Mediterranean. A famous example is the local maximum around 4.2 ka BP in Fig. 2.6, which Bar-Matthews et al., 1998 associate with a dry period that possibly caused the total desertification of the Sahara and the collapse of the Akkadian Empire. They also describe a transition to drier and more stable conditions during $3 - 1$ ka BP. In particular, Orland et al., 2009 recognize from these isotopic data a decrease in precipitation during the period $1.9 - 1.3$ ka BP. They suggest that this climate change weakened the economic system of the Roman and Byzantine Empires, which contributed to the decline of their rule in the Levant.

Overall, it can be concluded that many known climatic events are detectable in speleothems distributed throughout Europe. These can be identified relatively easily by spectral analysis. In addition, misinterpretations of speleothem isotope curves due to the source effect can be minimized when viewed in spectral space (Cheng et al., 2015). Therefore, this work will also focus on spectral characteristics of these proxy information as shown in Fig. 2.7. In order to correctly interpret the spectral properties of the proxy information in panels (a) and (c), some technical details must be taken into account. Fairchild et al., 2006 point out that smoothing, aliasing of the respective time series, and the total period covered must be treated with caution. For these reasons, only time series longer than 2000 years are considered in this work. In addition, the time range is chosen to coincide with that of the corresponding sediment cores (see Sect. 2.1.2). For example, the Soreq Cave data in Fig. 2.7 (a) are congruent with those from the Dead Sea pollen record of Litt et al., 2012 (nearly the last 10,000 years). The WPS in (b) reveals within the COI an important branch of significant power between 1500 and 1000 years. It is noticeable that in the early Holocene the highest power is just over 1000 years and this shifts over time to about 1500 years. Debret et al., 2007 can also show a similar trend using wavelet analysis for the IRD time series of Bond et al., 2002 with respect to the North Atlantic. In Fig. 2.7 (c) and (d), we see the time series of the IRD and the respective WPS. Again, the bifurcation between 1 and 1.5 ka period is striking, consistent with the patterns of Bond events. Debret et al., 2007 find a correlation between the fast (1 ka period) and slower (2.4 ka period) dominant fluctuations and solar activity. In contrast, periods around 1.5 ka from the middle to late Holocene could be due to internal oceanic forcing. In our work, this idea of pattern recognition is used to detect and compare spatiotemporal similarities between the respective proxies. Furthermore, Cheng et al., 2015 describe a power peak at 500 years in signals from Jeita Cave, which they hypothesize is due to solar variations. Based on this research, we will filter out signals smaller than 400 years using a low-pass filter. This way we can avoid the problem of aliasing as much as possible. The procedures of smoothing and how irregular time grids are transformed regularly are described in Sect. 2.3.

2.1.2 Pollen and macrofossils

The first studies of pollen grains began in the mid-seventeenth century (Manten, 1967), when the development of microscopes was increasingly advanced. Göppert, 1836 was the first to describe fossil pollen grains found in coal mines, compare them with recent pollen, and study their fossilization process. Edwards et al., 2017 summarize the work of Lennart von Post, who presented the first systematic pollen analysis of bog sediments at a meeting in Oslo in 1916. His work included several pollen diagrams from different sites. In these he showed percentage values of tree pollen on the x-axis and depth (time) on the y-axis. This systematic palynology, developed more than 100 years ago, has been preserved in its basic features. A few decades later, Iversen, 1944 reconstructed January and July temperatures in Scandinavia based on various plants, showing that it is possible to calculate quantitative climate reconstructions based on palynology. His basic idea was further developed in the following years and is also used in this work. A detailed description of this can be found in Sect. 2.4.

Fig. 2.8 shows a conceptual model for the deposition of pollen information in lake sediments.

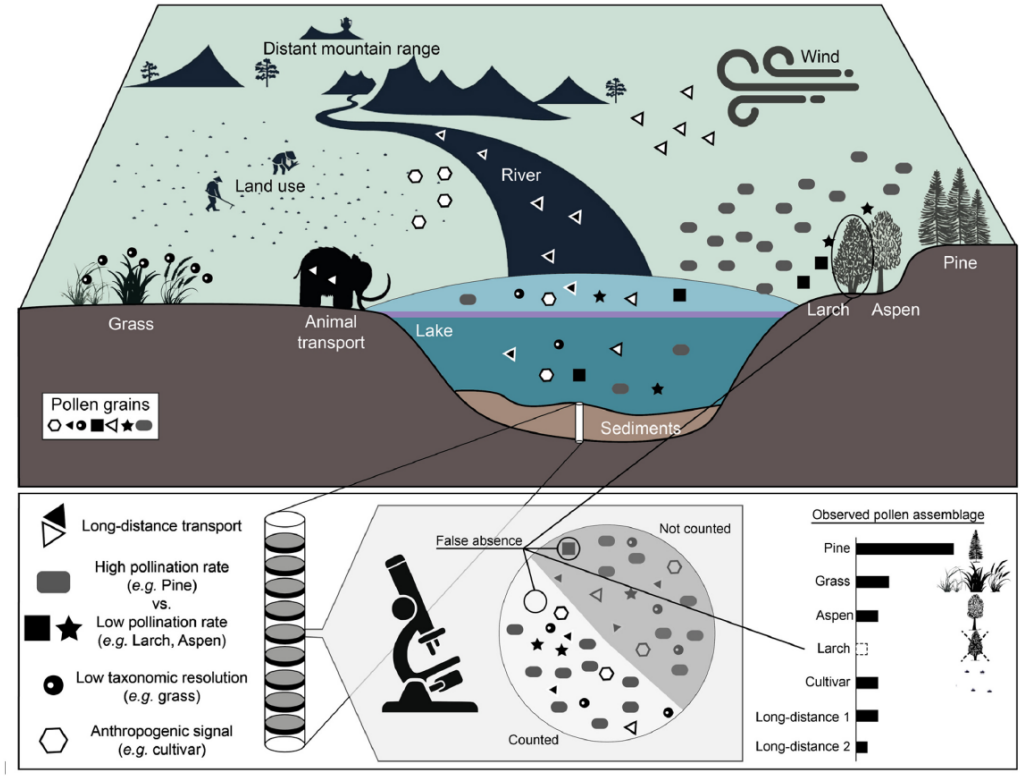


Figure 2.8: This figure is from Chevalier et al., 2020 and describes the influences that affect pollen composition in lake sediments. These include, as shown by the different symbols of pollen grains, long-distance transport (by wind, animals, and rivers), high and low pollination rates of different taxa, lower taxonomic resolution, e.g., in grasses, and human influence through land use. All of these factors ultimately play a role in lake sediment analysis.

Various pollen assemblages based on different taxa are created there. For further illustration, an example of this is shown in Fig. 2.9, where the percentage occurrence of some pollen grains is plotted against depth. The data are from a drill core from the Sea of Galilee in the Jordan Valley. A major advantage of palynology is that the envelope of pollen grains is made of the chemically stable polymer sporopollenin (Havinga, 1967). As a result, their structure is preserved over very long periods of time in depositional sites such as lake sediments. A number of factors play an important role in the collection of pollen information that must be considered in climate reconstructions. First, different plant species have different pollination rates. The larch, aspen, and pine shown in Fig. 2.8 are representative of this. The latter has a higher pollination rate than the first two. These differences make it difficult to determine if a particular plant species is present at a particular site, and if so, at what intensity. Another factor is pollen transport. These include first and foremost the wind, the rivers and the animals (Chevalier et al., 2020). As a result, pollen assemblages do not necessarily reflect local vegetation (Huntley, 2012). This is also the case in the Sea of Galilee, where the Jordan River is the largest inflow and thus mainly determines the catchment size (Schiebel and Litt, 2018). For these reasons, information about local and regional vegetation may be present in lake sediments. Another factor is the cultivation of plants by humans. If, for example, the occurrence of Olea europaea in Fig. 2.9 is subject to strong fluctuations at a depth of 1150-1300 cm, this is due to human influences rather

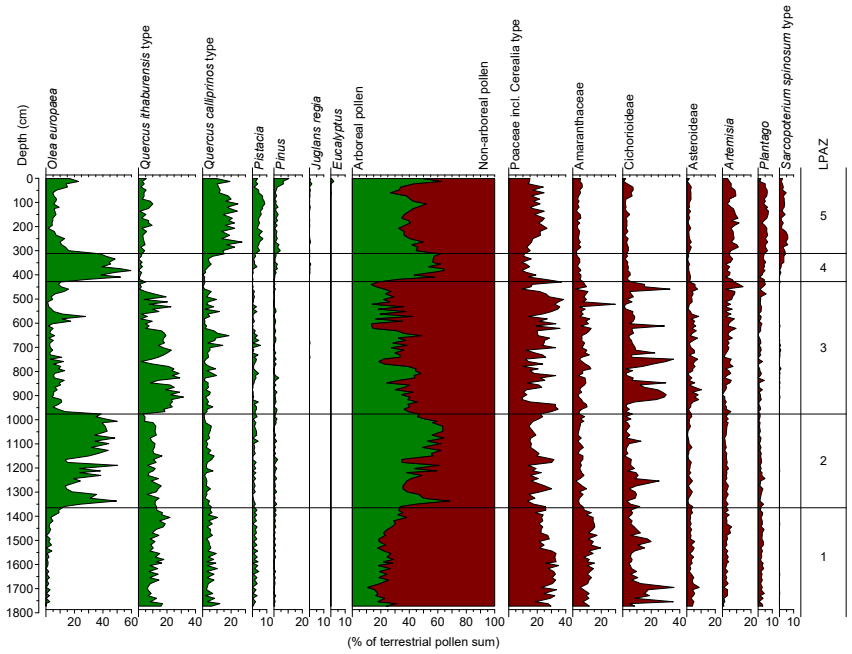


Figure 2.9: Percent distribution of terrestrial pollen sums as a function of depth for some important taxa from the Sea of Galilee. In the center, the summarized arboreal pollen (AP) (green) is contrasted with the non-arboreal pollen (NAP) (red). On the right, we see 5 local pollen assemblage zones (LPAZs) of this sediment core subdividing different pollen compositions.

than climatic changes (Schiebel and Litt, 2018). Further uncertainties in the examination of sediment cores arise from the assignment of pollen to plants. Overall, pollen analyses can be used to determine the taxonomic rank such as family or genus of a plant (Bradley, 2014). In contrast, macrofossils found in the core, for example small leaf or stem remains, can be distinguished down to species. In addition, such plant remains offer the possibility of dating without a reservoir effect, i.e., the carbon is in equilibrium with the atmosphere, allowing a more meaningful relationship between age and depth (Giesecke et al., 2013).

One goal of this work is to compute spatial climate reconstructions for the European region, North Africa, and the Levant (see Sect. 2.6) based on at least these two proxies. For this purpose, different pollen diagrams like the one in Fig. 2.9 are collected. The European Pollen Database (EPD) provides such an accumulation of data (Fyfe et al., 2009). It was established in the late 1980s and is constantly updated. Today it contains several thousand pollen records from natural archives such as lake and mire sediments. The database can be accessed with the help of the R package EPDr (Nieto-Lugilde et al., 2019), which provides access to the EPD's PostgreSQL database structure via R and some other useful functions. These provide, for example, the possibility to filter out entities for which no age dating is available or whose use is not permitted. In addition, specific taxa groups can be selected and taxa names can be changed to those accepted by the EPD. As shown in the pollen diagram Fig. 2.9 from the Sea of Galilee, it is often useful to convert pollen counts into percentages. In this regard, a corresponding function is also included in this R package. In particular, the ability to select the

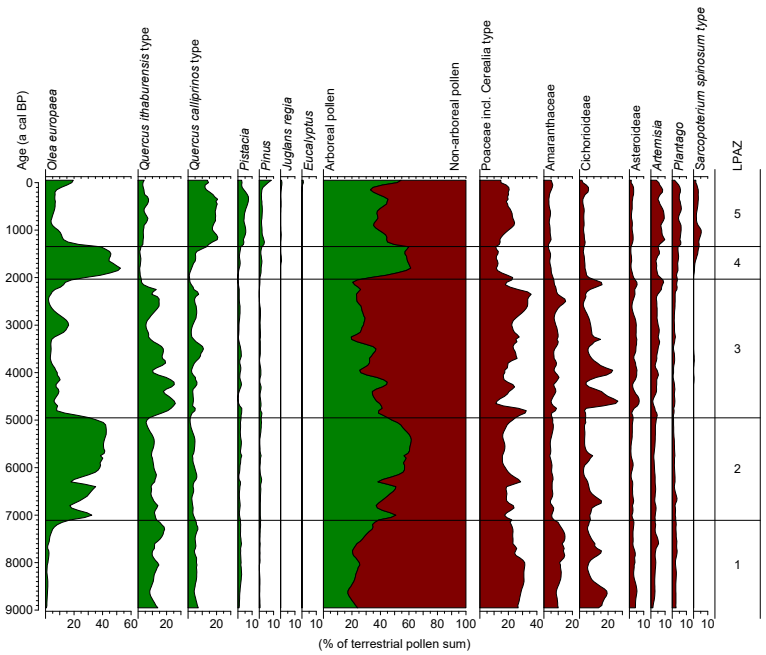


Figure 2.10: As Fig. 2.9, but with a transformation from depth to age of sediment core.

entities to which Giesecke et al., 2013 assigned new chronologies is of special interest to this work. In doing so, they included additional control data points with ages and uncertainties for some cores from the EPD. These include, for example, the onset of the Younger Dryas, the tephra of the Laacher See eruption, or the estimated age of the top of the core. With such additional data points, they calculated new age-depth relationships for 1036 sites. For some of them, the R package clam (Blaauw, 2019) was used, which applies a smoothing spline regression to compute such a model along with the uncertainties. Based on these datings, we build additional age-depth relationships using a **B**ayesian **ac**cumulati**on** model (Bacon) (Blaauw and Christen, 2011). This can be seen as a successor to clam and is described in more detail in Sect. 2.3. Ultimately, we obtain about 600 pollen diagrams with corresponding age-depth models in this way.

Each depth from Fig. 2.9 contains vegetation information from a specific time interval. This depends on the sedimentation rate and the thickness of the sampling interval of the core. Overall, this results in an irregular grid in depth. To compare this information with that from other lake and mire sediments, from speleothems, and from marine and ice cores, a conversion from depth or distance to age is required. For this purpose, a new method is applied in this work, which generates, for example, Fig. 2.10. The depth information from Fig. 2.9 is used here as input data. Together with the age uncertainties, the output data are smoothed percentage curves for each plant, which now have a regular temporal grid. This procedure is fully described in Sect. 2.3.

2.2 Bayesian statistics

As we have already seen in the Introduction, the high-dimensional climate system must be described with the help of probabilities. Moreover, in paleoclimatology we are dealing with more or less large age uncertainties, and the information from the climate archives described above are not only influenced by climate. For these reasons we have to consider a probabilistic approach to describe the variables needed in this work.

Bayesian statistics deals with uncertainties using conditional probabilities and has been successfully used for climate reconstructions for several decades (e.g. Kühl and Litt, 2003; Gebhardt et al., 2008; Neumann et al., 2007; Simonis et al., 2012; Parnell et al., 2014; Weitzel et al., 2019). The original idea goes back to Sir Thomas Bayes (Bayes and Price, 1763), who interpreted probabilities as a measure of the degree of belief that an event will occur given additional information. Bayes' theorem describes this situation by relating conditional probability distributions:

$$\underbrace{\mathbb{P}(\boldsymbol{X} \mid \boldsymbol{Y})}_{\text{Posterior}} = \frac{\mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{X}) \cdot \mathbb{P}(\boldsymbol{X})}{\mathbb{P}(\boldsymbol{Y})} \propto \underbrace{\mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{X})}_{\text{Likelihood}} \cdot \underbrace{\mathbb{P}(\boldsymbol{X})}_{\text{Prior}}. \tag{2.7}$$

The posterior probability distribution describes the quantity $\boldsymbol{X}$ inferred from $\boldsymbol{Y}$. Because $\mathbb{P}(\boldsymbol{Y})$ is a fixed normalization constant, the posterior is proportional to prior knowledge and the so-called

likelihood. The latter is used in this thesis in two different versions. First, as a transfer function with respect to local climate reconstructions and second, as a data operator with respect to spatial climate reconstructions. In addition, two common methods are used to determine the prior information. This includes the use of the so-called Jeffreys prior, which is considered non-informative due to its parameterization invariance (Gelman et al., 2013). This means that an attempt is made to minimize the information content of the prior distribution so that the posterior distribution reflects as closely as possible the information about the parameters obtained from the data. Moreover, we use so-called conjugate priors. These are called such if the posterior distribution belongs to the same probability distribution family as the prior probability distribution (Gelman et al., 2013). This has the advantage of a much more efficient determination of the posterior distribution.

As we have already seen, we intend to use a variety of different proxy information to reconstruct paleoclimate. This is ensured by the multiple application of Bayes' theorem (see Eq. 2.7), which is finally called Bayesian Hierarchical Model (BHM), taking into account the uncertainties of the different variables. This makes it possible to use several sources and combine them into a joint model. All in all, there are several ways to calculate the posterior distribution of a BHM. Two main methods

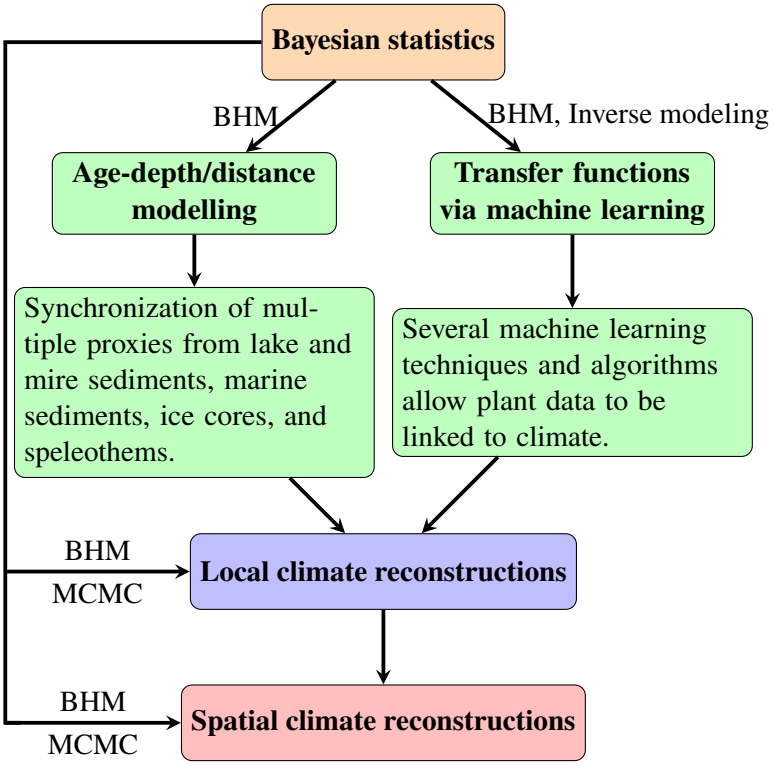


Figure 2.11: Flowchart of the applications of Bayesian statistics. The arrows indicate the reading direction of the diagram and the different work steps are marked by various colors. The descriptions along the arrows show in more detail how the ideas of Bayesian statistics are used and solved.

are used in this thesis. First, the different probability distributions are determined and discretized and finally multiplied together. This has the disadvantage of some inaccuracy, but the advantage of less computational cost. Second, Markov chain Monte Carlo (MCMC) methods are used. In this process, the various probability distributions are sampled such that the Markov chain converges to the correct posterior distribution. In Sect. 2.4, two MCMC methods are described in more detail, namely the Gibbs and Metropolis-Hastings sampling methods (Betancourt, 2017).

The flowchart in Fig. 2.11 shows how the basic principles of Bayesian statistics just described are used in the following sections. This includes Bayesian hierarchical modeling with respect to age-depth/distance transformations solved with matrix multiplications. The purpose of this step is to allow different climate proxies to be synchronized with respect to age. In parallel, transfer functions are inversely determined using various machine learning techniques to establish a probabilistic relationship between plants and climate. The next step is a BHM of local climate reconstructions based on the data just described and solved with MCMC. All this leads to the possibility of spatial climate reconstructions, also based on a BHM and solved using MCMC.

2.3 Age-depth/distance modelling

This section deals with the relationship between age dating and depths in lake and mire sediments and distances in speleothems. The various methods of age dating can be found in Bradley, 2014. For the relationship between age and depth/distance, we use a Bayesian age-depth/distance model in order to account for age uncertainties. In contrast, most previous studies are based solely on the mean values. For example, depths of lake sediments (Litt et al., 2012; Schiebel and Litt, 2018; Torfstein et al., 2015; Neumann et al., 2007; Miebach et al., 2019; Seppä et al., 2005) or distances (from a reference point) with respect to speleothems (Drăguşin et al., 2014; Waltgenbach et al., 2020; Akers et al., 2016; Psomiadis et al., 2018). For the latter, the values are based on models such as StalAge (Scholz and Hoffmann, 2011) and Bacon (Blaauw and Christen, 2011). In our work, we instead consider the entire uncertainty domain. The Bayesian age-depth model Bacon just mentioned, which is implemented in R, is used for this purpose (R Core Team, 2018; Blaauw et al., 2020). We will only briefly explain the model in this thesis. A detailed description can be found in Blaauw and Christen, 2011.

The Bacon model uses a self-adjusting MCMC simulation to infer the posterior probability distribution $\mathbb{P}(\theta, x, \omega|y)$, where θ is the model parameter, x is the accumulation rate, ω is the memory, and y are measurements such as ^{14}C data. These are the same variable declarations as in Blaauw and Christen, 2011 and apply only to this section. The middle and right panels in the top row of Fig. 2.12 show an example of the prior distributions of x and ω, which are indicated as green lines. Based on the MCMC samples, the gray areas represent the corresponding posterior probability densities. From such a posterior accumulation rate x, we can obtain the probability distributions $\mathbb{P}(\boldsymbol{A}|\boldsymbol{D})$ and $\mathbb{P}(\boldsymbol{D}|\boldsymbol{A})$. $\boldsymbol{A}$ is the age and $\boldsymbol{D}$ denotes the depth or distance. The Bacon.Age.d function in the Bacon

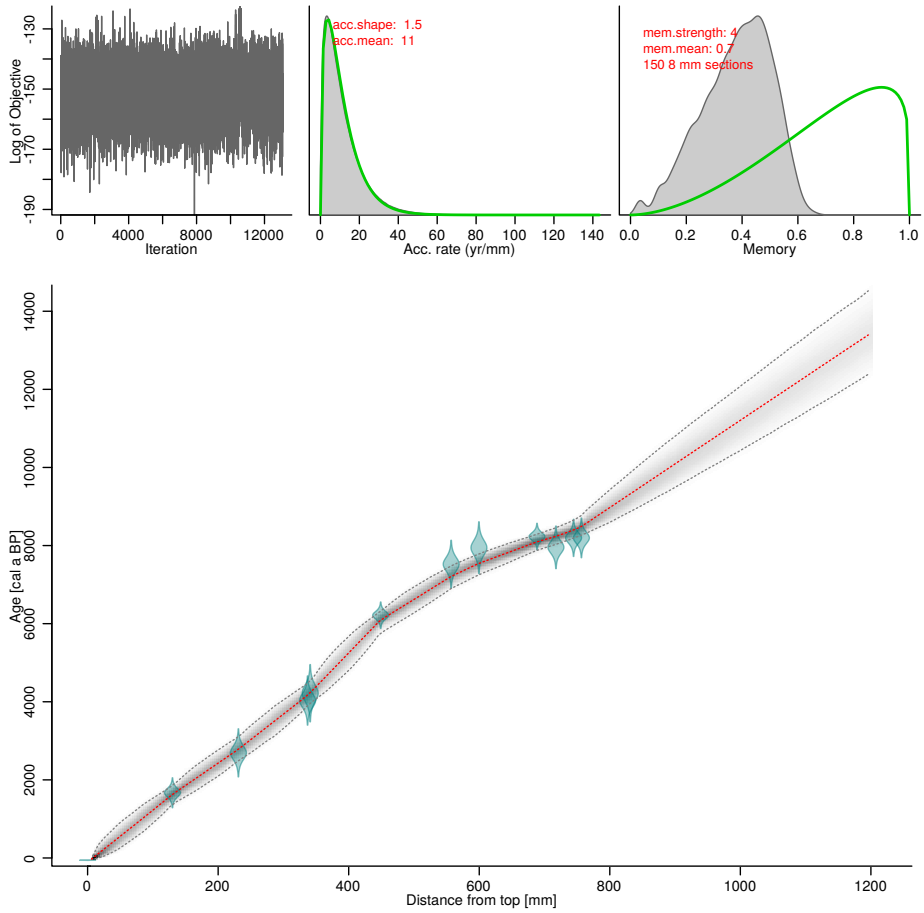


Figure 2.12: The lower panel shows an age-distance model of a speleothem from the Ascunsă Cave in Romania calculated with Bacon. The top panels depict, from left to right, a trace plot of the posterior MCMC samples, prior (green lines) and posterior (gray area) distributions of accumulation rate and memory.

package returns $\mathbb{P}(\boldsymbol{A}|\boldsymbol{D})$. In our study, however, we need a tool for $\mathbb{P}(\boldsymbol{D}|\boldsymbol{A})$ (see Eq. 2.10) which has been implemented since package version 2.5 and can be found under Bacon.d.Age. Examples of both probability densities are shown in Fig. 2.13, namely $\mathbb{P}(\boldsymbol{A}|440\,\text{mm})$ and $\mathbb{P}(\boldsymbol{D}|6000\,\text{cal a BP})$. The corresponding age-distance model can be seen in Fig. 2.12. It is based on the dating of a speleothem from the Ascunsă Cave in Romania (Drăguşin et al., 2014). Originally, the Bacon model was created for the calculation of age-depth models based on ^{14}C dating from cores of, e.g., peat bogs or lake sediments. That is, Bacon performs an appropriate calibration from ^{14}C a BP to cal a BP. However, this option can be disabled and the model assumes that the calibrated dates are already available, which is the case here. Such datings and their uncertainties can be taken from the speleothem database SISAL. Roesch and Rehfeld (2019) tested 88 entities for robustness to various age-distance models. Among them are Bchron (Haslett and Parnell, 2008) and linear interpolations as well as Bacon and StalAge. Linear interpolations, Bchron and Bacon provide robust results for these entities and the

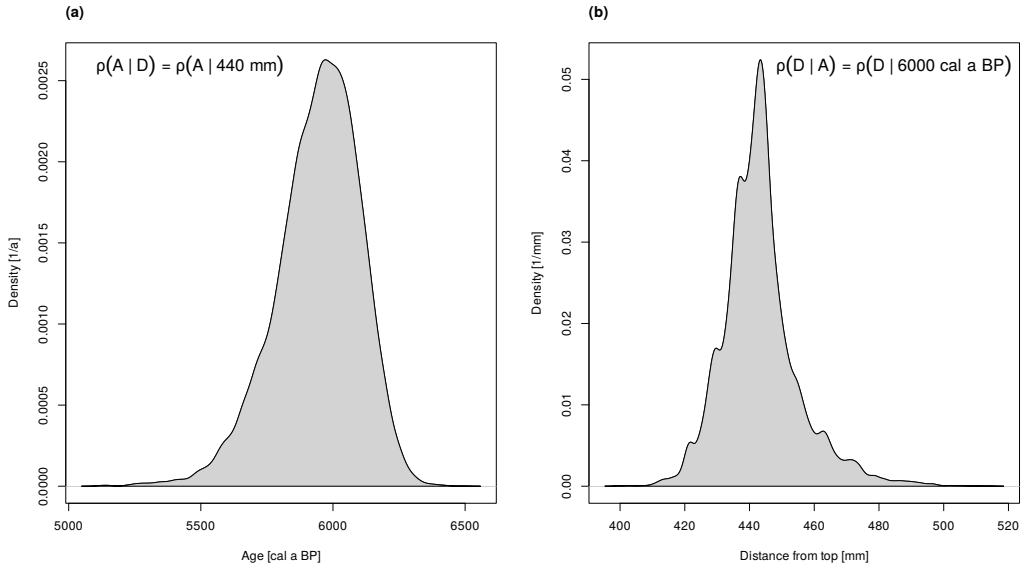


Figure 2.13: Probability density distributions based on the age-distance model in Fig. 2.12. Panel (a) depicts $\mathbb{P}(\boldsymbol{A}|440\,\text{mm})$ calculated with the R function Bacon.Age.d and (b) $\mathbb{P}(\boldsymbol{D}|6000\,\text{cal a BP})$ calculated with the R function Bacon.d.Age.

latter is the only one that can account for hiatus. Therefore, Bacon is also suitable as an age-distance model for speleothems. In the upper middle panel of Fig. 2.12, some prior parameters defining the gamma-distributed accumulation rate are also shown. These values of x are set to a mean accumulation rate of 11 mm/year and a of shape 1.5. This mean is calculated from the assumption of a uniform slope over the entire age-distance range, and the shape is the default value, which provides a flatter curve and thus more variability. As the growth rate of this speleothem is relatively constant (except the 8.2 event) during the Holocene (Drăguşin et al., 2014), the default settings of the memory are chosen, namely mean memory of 0.7 and memory strength of 4 (upper right panel). Overall, the trace plot of $\log(\mathbb{P}(\theta, x, \omega|y))$ in the upper left panel shows good mixing, indicating stable model performance with these prior parameters.

Now we need to determine how to use the information from this model to make an appropriate conversion from distances to ages. As mentioned above, the probability distributions $\mathbb{P}(\boldsymbol{D}|\boldsymbol{A})$ are crucial for this, which is shown by the following derivation using the technique of Bayesian Hierarchical Models:

$$\mathbb{P}(\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{P}) = \int_{\mathcal{D}} \mathbb{P}(\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{P}, \boldsymbol{D})\, d\,\boldsymbol{D} \tag{2.8}$$

$$= \int_{\mathcal{D}} \mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{P}, \boldsymbol{D}) \cdot \mathbb{P}(\boldsymbol{D} \mid \boldsymbol{A}, \boldsymbol{P}) \cdot \mathbb{P}(\boldsymbol{A}, \boldsymbol{P})\, d\,\boldsymbol{D}. \tag{2.9}$$

$\boldsymbol{Y}$ contains variables such as annual precipitation or δ^{18}O for which an age-depth/distance transformation

is to be performed. $\boldsymbol{P}$ means proxy information such as pollen, macrofossils, or any type of isotopic data. In the following, we assume that $\boldsymbol{D}$ is conditionally independent of $\boldsymbol{P}$ and thus completely dependent on $\boldsymbol{A}$. This is exactly the information we get from the age-depth/distance model. Moreover, the variables $\boldsymbol{Y}$ should not depend on age given $\boldsymbol{D}$. This assumption stems from the fact that all the proxy information we consider initially refers to a depth or distance. Then, we obtain the final age-depth/distance transformation:

$$\Rightarrow \mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{P}) = \int_{\mathcal{D}} \mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{D}, \boldsymbol{P}) \cdot \mathbb{P}(\boldsymbol{D} \mid \boldsymbol{A})\, d\,\boldsymbol{D}. \tag{2.10}$$

In this thesis, the conditional probability distributions are discretized and stored and evaluated in predefined matrices.

Examples of a transformation using Eq. 2.10 can be seen in Fig. 2.14. Blue represents δ^{18}O, red represents δ^{13}C, and black represents the respective smoothings of these values. Each of these 6 different curves can be understood as Y, except that their calculations are different. The colored curves are based on the mean values of the Bacon model from Fig. 2.12 (dashed red line). The dashed black curves are analogous to those in Drăguşin et al. (2014) using a 9-point moving average. In contrast, the solid black lines are calculated with the use of Eq. 2.10.

Here, $\mathbb{P}(\boldsymbol{Y}|\boldsymbol{D}, \boldsymbol{P})$ represents the relationship of δ^{18}O and δ^{13}C to their distance (with respect to the

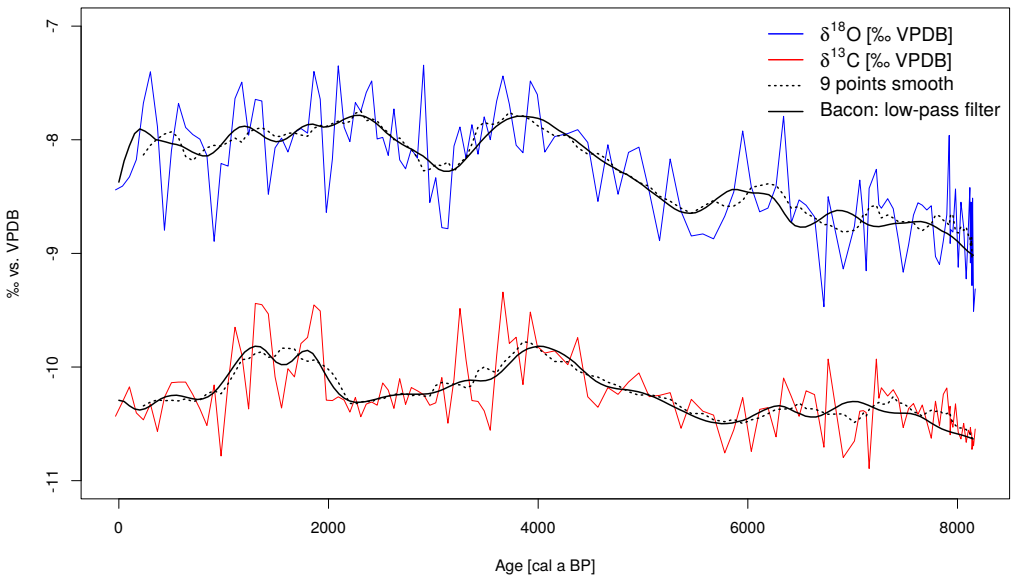


Figure 2.14: δ^{18}O (blue line) and δ^{13}C (red line) data from Drăguşin et al., 2014. The dashed black lines represent their moving average, and the solid black lines show our data-driven smoothing using the probabilistic relationship between age and distance.

top of the speleothem) and to $\boldsymbol{P}$, which can be interpreted as the total isotopic information of ^{18}O and ^{13}C, respectively. Thus, the blue and red curves can be defined as $\mathbb{P}(\delta^{18}\mathrm{O}|\boldsymbol{D},\boldsymbol{P})$ and $\mathbb{P}(\delta^{13}\mathrm{C}|\boldsymbol{D},\boldsymbol{P})$. The temporal resolution of $\mathbb{P}(\boldsymbol{D}|\boldsymbol{A})$ can be chosen arbitrarily high. In this thesis, this value is set at 50 years, which is a compromise between all speleothems and sediment cores. If it is specified evenly, this transformation results in a regular time grid. A regular grid is one requirement for the application of the wavelet transform introduced above. Nevertheless, there must be an age-depth/distance relationship per entity. Accordingly, such transformation cannot be performed for entities such as LR04, Medstack, and NGRIP (see. Fig. 2.2). Mathematically, Eq. 2.10 behaves like a convolution with different kernel smoother (cf. Eq. 2.2). For this reason, a smoothed pattern is obtained when age uncertainties are taken into account. Unlike traditional smoothing methods, this one is data-driven and shows what

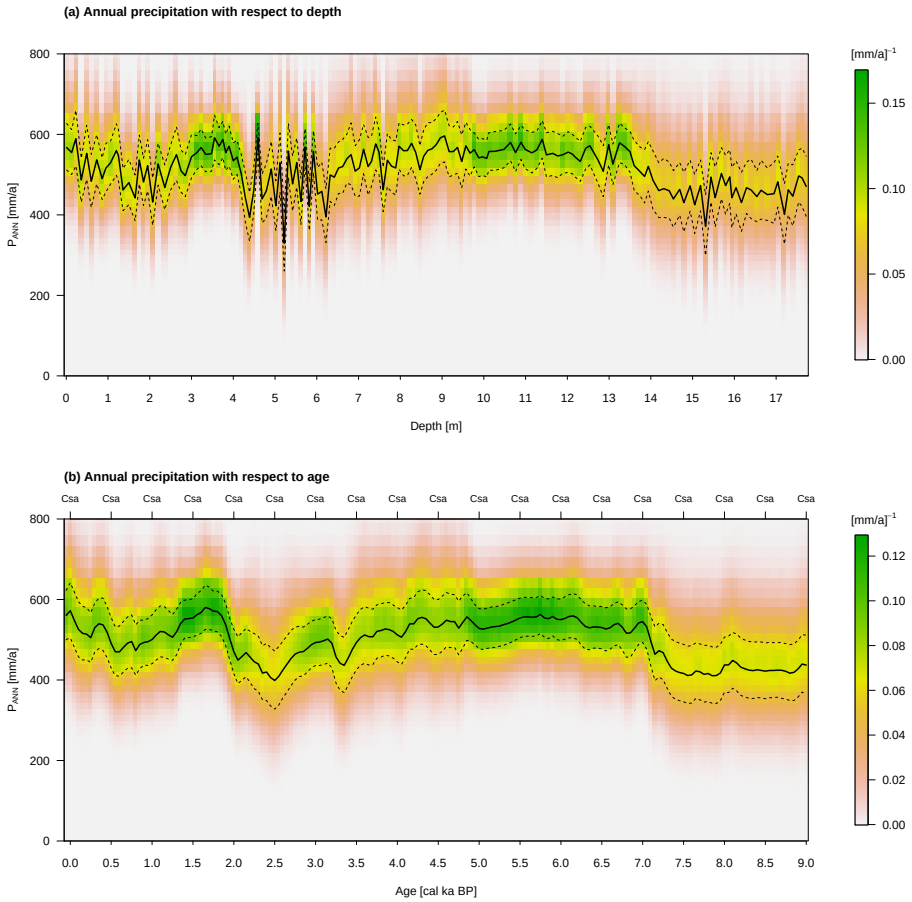


Figure 2.15: Panel (a) shows the climate reconstruction of annual precipitation as a function of the depth of the lake sediment of the Sea of Galilee. The colored contours represent the uncertainties in terms of probability densities, and the dashed black lines depict the interquartile range based on these data. In it, the median is shown in the form of a solid black line. Panel (b) corresponds to (a) with the difference that the x-axis is transformed from depth to age.

information $\boldsymbol{Y}$ actually provides in terms of age.

Another advantage of this method is its simple implementation and low computational cost. Previously, age uncertainties were based on sampling algorithms as in Parnell et al. (2014) or on ensembles of age depth/distance models (McKay et al., 2020). In our case, we only need to create $\mathbb{P}(\boldsymbol{D}|\boldsymbol{A})$ and store it in a matrix with dimensions $m \times n$. m is the number of points with respect to $\boldsymbol{D}$ and n is the user-dependent number of data points related to age. Pragmatically, $\mathbb{P}(\boldsymbol{Y}|\boldsymbol{D}, \boldsymbol{P})$ has dimension $1 \times m$ in the case of Fig. 2.14. Thus, Eq. 2.10 can be calculated simply by matrix multiplication (denoted here by $\otimes$):

$$\mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{P}) \in^{1\times n} = \mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{D}, \boldsymbol{P}) \in^{1\times m} \otimes \mathbb{P}(\boldsymbol{D}|\boldsymbol{A}) \in^{m\times n} .$$

This fact is taken up again in Sect. 2.4, where the advantage of the low computational cost of this multiplication is exploited.

$\mathbb{P}(\boldsymbol{Y} \mid \boldsymbol{D}, \boldsymbol{P})$ can also have the dimension $k \times m$. $k > 1$ denotes the number of grid points with respect to $\boldsymbol{Y}$. In Fig. 2.15, $\boldsymbol{Y}$ is the mean annual precipitation ($\mathrm{P_{ANN}}$). Probability density values are available for this variable at each depth (colored contour in (a)). The dimension for the discretized probability density function (PDF) is also user-dependent and is set to $k = 300$ in this thesis. The solid black line represents the median of each PDF. All in all, the transformation from depth (a) to age (b) follows the same procedure as the one described above. In total, the following is calculated:

$$\mathbb{P}(\mathrm{P_{ANN}} \mid \boldsymbol{A}, \boldsymbol{P}) \in^{k\times n} = \mathbb{P}(\mathrm{P_{ANN}} \mid \boldsymbol{D}, \boldsymbol{P}) \in^{k\times m} \otimes \mathbb{P}(\boldsymbol{D}|\boldsymbol{A}) \in^{m\times n} . \tag{2.11}$$

This example serves only to illustrate the transformation from depth (a) to age (b). A corresponding description from a paleoclimatological perspective can be found in Sect. 3.3.2. How such a reconstruction is calculated in detail is described in the following.

2.4 Local climate reconstructions

The climate reconstruction model used in this thesis is based, among other things, on the idea of the indicator species approach. In this method, pollen or macrofossils from deposits, such as lake and mire sediments, are considered as climate indicators. Using this approach, Iversen, 1944 reconstructed January and July temperatures during the Holocene in Scandinavia based on *Ilex*, *Viscum*, and *Hedera*. Building on these basics, the Mutual Climate Range (MCR) method has evolved (Grichuk, 1969). Here, plant distribution maps are linked to recent climate data to define a climate range where the corresponding taxon occurs. Finally, when considering multiple taxa, these climatic ranges can be combined to determine the mutual climatic range. This method brings a number of challenges. First, there is no weighting within the individual climate ranges. This means, for example, that each temperature range occurs with the same probability. These sharp boundaries in transfer functions

result from overfitting (Friedman, 2001). This is a general problem in determining such functions and will be taken up again in more detail in the next section. On the other hand, combining too many taxa can result in more than one MCR. In contrast, when too few taxa occur in a certain sediment layer, reconstruction can lead to a relatively broad climatic range (Klotz et al., 2004). In addition, only the occurrence of the plants is considered. In this method, the percentage information is included only indirectly, by determining a priori whether a plant is present or not. There are some other climate reconstruction methods that incorporate such percentage distributions. These include regression methods in which linear or nonlinear transfer functions are calculated based on modern climate and plant data (Webb and Bryson, 1972). In addition, there are classification methods, such as Modern Analogue Technique (MAT) (Overpeck et al., 1985). The same machine learning (ML) algorithms can be used in both regression and classification (Jed Wing et al., 2019). These are among others Artificial Neural Network (ANN) or Random Forest (RF) (see Sect. 2.5). Chevalier et al., 2020 present a table describing the advantages and disadvantages of the above methods. One advantage of the MCR method is that it can be used globally. In contrast, the regression and classification approaches are more spatially constrained and often have poorer estimates of uncertainty. In addition to these methods, Chevalier et al., 2020 mention two more: Bayesian approaches and inverse modeling. Both are also used in this thesis, so that an attempt is made to balance the disadvantages of all methods with their respective advantages. In the following, we describe how we incorporate a Bayesian approach into our reconstruction model.

To do this, we follow the idea of Kühl et al., 2002, who developed a probabilistic interpretation of MCR. Fig. 2.16 shows the concept of this method. This addresses the problem of overfitting by calculating uncertainty ranges for each taxon. These are shown here for taxa A, B, and C for different colored contours (blue, brown, and green). The MCR follows the colors of the color bar. The uncertainty ranges were initially based on normally distributed PDFs (with respect to temperature), which is why this is called the PDF method. Later, more complicated models were also used, which will be discussed in more detail in Sect. 2.5. The PDF method was applied for both local and spatial climate reconstructions. For example, Kühl and Litt, 2003 calculated January and July temperatures for 3 sites in Central Europe during the Last Interglacial (LIG) period. Subsequently, spatial reconstructions of Europe were performed in Gebhardt et al., 2008 for the Eemian, in Simonis et al., 2012 for the Late Glacial and Holocene, and in Weitzel et al., 2019 for the MH. Schölzel, 2006 describes the PDF method in the context of a BHM and calls it Bayesian Indicator Taxa Model (BITM). This has the advantage that additional prior information can regulate the transfer functions. Among others, the BITM was applied in Neumann et al., 2007 for Birkat Ram in Israel, in Stolzenberger, 2011 for Meerfelder Maar and Holzmaar in Germany, and in Thoma, 2017 for Lake Prespa in Greece. In all BITM-based reconstructions, thresholds must be set to define at which point a taxon occurs or does not occur. Another method presented by Schölzel, 2006 that is based on BHM is the so-called Bayesian Biome Model (BBM). In this process, certain plant taxa are assigned to different biomes. These are

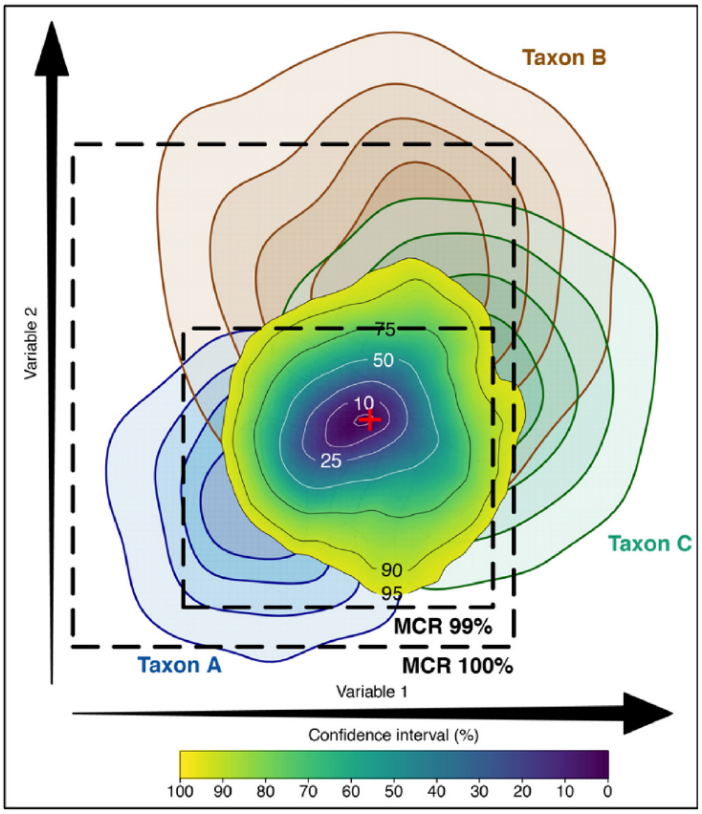


Figure 2.16: Illustration of a Mutual Climate Range from Chevalier et al., 2020. Shown are 3 taxa as a function of two climate variables and their MCR, represented by the yellow/blue contours. The red cross represents the mean of the MCR, and the dashed rectangles demonstrate the sensitivity of this technique when only a subset of the MCR data are considered. That is, if the input data contain 99 %, there may be a relatively large difference from the full input data set (MCR 100 %).

groups that have similar climatic conditions and predominant vegetation, among other things (Prentice et al., 1992). One advantage of the BBM is that one does not need recent distribution maps for every plant occurring in the core, but only for the biomes used. The computational assignment of taxa to biomes is based on the affinity score (Prentice et al., 1996). This is used to calculate the percentage distributions of the biomes with respect to the depth of the core. Nevertheless, the information of a threshold value of the individual taxa is included in the affinity score. Applications of the BBM can be found in Litt et al., 2012, in Stolzenberger, 2017, and Litt et al., 2021 for the Dead Sea in Israel. Based on the above summary, the first part of the reconstruction method used here is now described. The inverse modeling is done in the second part, which is discussed below. The goal of the new method is to provide an algorithm that performs the calculations regardless of location across Europe. This is intended to circumvent expert knowledge for setting thresholds for taxa occurrence, which is a major effort for the approximately 600 sites. Instead, the new method should independently determine which taxa are important for each reconstruction based on specific constraints. All variables that are relevant

Table 2.2: List of variables.

Variable shortcut	Description
$\mathbb{P}$	Probability distribution
$\boldsymbol{C}$	Climate: contains modern $\boldsymbol{C_m}$ and past $\boldsymbol{C_p}$ climate information
$\boldsymbol{P}$	Proxy: contains modern $\boldsymbol{P_m}$ and past $\boldsymbol{P_p}$ proxy information
$\boldsymbol{PP}$	Proxy pool: explained variances between additional proxies and $\boldsymbol{C}$
$\boldsymbol{P_s}$	Selected plant proxy information
$\boldsymbol{A}$	Age of sediments or speleothems
$\boldsymbol{D}$	Depth of sediments or distance in speleothems
$\boldsymbol{\Theta}$	Contains the following parameters:
$\boldsymbol{\psi}$	Link between climate $\boldsymbol{C}$ and taxa distributions $\boldsymbol{P}$
$\boldsymbol{\phi}$	Contains all information about the taxa weights:
ω	Taxa weights
z	Selection of taxa based on taxa weights ω

for the mathematical formulation of this model can be found in Table 2.2.

Using Bayes' theorem, we can express the probability distribution for climate $\boldsymbol{C}$ given pollen and macrofossils $\boldsymbol{P}$, depth $\boldsymbol{D}$, and parameter $\boldsymbol{\Theta}$. In the process, we also introduce the selected plant proxies $\boldsymbol{P_s}$:

$$\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\Theta}) = \int_{\mathcal{P}_s} \mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{P_s}, \boldsymbol{D}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\Theta})\, d\boldsymbol{P_s}. \tag{2.12}$$

In the case of a finite number of taxa, the integral is a corresponding sum. Consider $\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{P_s}, \boldsymbol{D}, \boldsymbol{\Theta})$ in more detail:

$$\begin{aligned}\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{P_s}, \boldsymbol{D}, \boldsymbol{\Theta}) &\overset{1.}{\approx} \mathbb{P}(\boldsymbol{C} \mid, \boldsymbol{P_s}, \boldsymbol{D}, \boldsymbol{\Theta}) \overset{2.}{\approx} \mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P_s}, \boldsymbol{\Theta}) \overset{3.}{\approx} \mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P_s}, \boldsymbol{\psi}) \\ &\overset{4.}{\approx} \frac{\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{C} \mid \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\psi})}{\mathbb{P}(\boldsymbol{P_s}, \boldsymbol{\psi})} \\ &\overset{5.}{\approx} \frac{\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\psi})}{\mathbb{P}(\boldsymbol{P_s}, \boldsymbol{\psi})}.\end{aligned} \tag{2.13}$$

With the following assumptions and applications:

1. $\boldsymbol{C}$ is conditionally independent of $\boldsymbol{P}$ if $\boldsymbol{P_s}$ is given. This assumes that $\boldsymbol{P_s}$ explains enough variability of the core.

2. The link between $\boldsymbol{C}$ and $\boldsymbol{P_s}$ is conditionally independent of depth. This means that the relationship between these quantities is assumed to be unchanged for any depth and thus any age of the core. The assumption that this relationship has not changed over time is an important part of our reconstruction method. When we look at older time periods, we need to keep this in

mind, as the relationship may well have changed due to evolutionary processes.

3. The connection between $\boldsymbol{C}$ and $\boldsymbol{P_s}$ is described only by the parameter $\boldsymbol{\psi}$. Furthermore, $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$ are a priori independent: $\mathbb{P}(\boldsymbol{\Theta}) = \mathbb{P}(\boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\phi})$.

4. Application of Bayes' theorem.

5. If $\boldsymbol{C}$ depends only conditionally on $\boldsymbol{\psi}$, it is assumed that within the considered climate ranges each value occurs with the same probability. Thus, $\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{\psi})$ follows a uniform distribution.

If we now substitute Eq. 2.13 into equation Eq. 2.12, we get:

$$\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\Theta}) \approx \int_{\mathcal{P}_s} \frac{\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\psi})}{\mathbb{P}(\boldsymbol{P_s}, \boldsymbol{\psi})} \cdot \mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\phi})\, d\boldsymbol{P_s}. \tag{2.14}$$

Combined with Eq. 2.10, the following expressions can be found:

$$\begin{aligned}
\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) &= \int_{\mathcal{D}} \mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{D} \mid \boldsymbol{A})\, d\boldsymbol{D} \\
&\overset{\text{Eq.2.14}}{\approx} \int_{\mathcal{D}} \int_{\mathcal{P}_s} \frac{\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\psi})}{\mathbb{P}(\boldsymbol{P_s}, \boldsymbol{\psi})} \cdot \mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\phi})\, d\boldsymbol{P_s} \cdot \mathbb{P}(\boldsymbol{D} \mid \boldsymbol{A})\, d\boldsymbol{D} \\
&\approx \int_{\mathcal{P}_s} \frac{\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\psi})}{\mathbb{P}(\boldsymbol{P_s}, \boldsymbol{\psi})} \cdot \int_{\mathcal{D}} \mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{P}, \boldsymbol{D}, \boldsymbol{\phi}) \cdot \mathbb{P}(\boldsymbol{D} \mid \boldsymbol{A})\, d\boldsymbol{D}\, d\boldsymbol{P_s} \\
&\overset{\text{Eq.2.10}}{\approx} \int_{\mathcal{P}_s} \frac{\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\psi})}{\mathbb{P}(\boldsymbol{P_s}, \boldsymbol{\psi})} \cdot \mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\phi})\, d\boldsymbol{P_s}.
\end{aligned} \tag{2.15}$$

This is the basic model calculated several thousand times for different $\boldsymbol{\phi}$ by systematically sampling from the pools of plant information of the respective cores using MCMC techniques. In order to be able to describe this in more detail, certain framework conditions must first be introduced. To this end, we will now look at reference curves based on pollen and isotope information. As described in Sect. 2.1.1 and Sect. 2.1.2, many studies use such reference curves. These are shown, for example, in Fig. 2.2, Fig. 2.6, and the AP/NAP curve in Fig. 2.10. If a reconstruction according to Eq. 2.15 is performed for certain taxa weights $\boldsymbol{\omega}$, the expectation and/or median values of $\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta})$ can be compared with these reference curves. Here, the explained variances R^2 are used as a similarity measure and stored in a proxy pool $\boldsymbol{PP}$. Based on this idea, an extended BHM can be constructed (the weighting term is omitted for convenience):

$$\mathbb{P}(\boldsymbol{C}, \boldsymbol{\Theta} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{PP}) \propto \mathbb{P}(\boldsymbol{PP} \mid \boldsymbol{C}, \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{A}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{A}, \boldsymbol{\Theta}). \tag{2.16}$$

Note that the posterior distribution includes not only the climate variables as in Eq. 2.15, but also $\boldsymbol{\Theta}$. Finally, this model provides information on which taxa are important with respect to the selected reference curves. In this model, one could add a variety of additional reference curves based on, for

example, solar radiation or CO_2 (see Fig. 1.1). However, this work is limited to the consideration of proxies derived from isotopes and plant information. We know that some sections of the AP/NAP curves have fluctuations that are not due to climatological changes (e.g. Panagiotopoulos et al., 2013; Miebach et al., 2016; Schiebel, 2013; Neumann et al., 2007; Litt et al., 2012). In particular, the influence of humans on the vegetation around the lakes studied during the middle to late Holocene, complicates the interpretation of these curves. To account for such uncertainties, we specify a priori that the climate reconstructions should explain about 50 % of the variance of the respective reference curves. This offers on the one hand the possibility to follow the trends of the references and supports assumption 1 above that $\boldsymbol{P_s}$ provides sufficient variability. On the other hand, additional factors that influence the curves are taken into account. Thus, we specify the same prior distribution for each reference, namely beta distributions with a mean of 0.5 and a standard deviation of about 0.19. The latter leads to a better consideration of the wide range of lakes considered in this thesis. This results in a beta distribution with shape parameters 3. Assuming that all reference curves N_{proxy} are independent, the probability distributions of each component i of the vector $\boldsymbol{PP}$ can be defined as follows:

$$\mathbb{P}(PP_i \mid \cdot) = \text{Beta}(PP_i \mid 3, 3). \tag{2.17}$$

The parameters of the beta distributions can be changed individually depending on the lake studied and the reference proxy considered. However, for the purposes of this work, they represent a good compromise for the automatic calculation of all climate reconstructions for the entire European region, Northern Africa, and the Levant. At this point, instead of using multiple beta distributions, one might consider using the appropriate multivariate generalization. This is the Dirichlet distribution, which in our case has the disadvantage of being somewhat slower in MCMC sampling. In addition, other related metrics could be used in place of explained variance in future work. For example, the target redundance described in Glowienka-Hense et al., 2020 could be an alternative. This is the information that both reference curves provide equally, allowing them to be evaluated simultaneously.

Furthermore, we split the local reconstruction module as follows:

$$\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) = \mathbb{P}(\boldsymbol{C_p} \mid \boldsymbol{C_m}, \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{C_m} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}). \tag{2.18}$$

The first term on the right gives us the ability to constrain the reconstructions based on additional climate information from the past. These may be, for example, other local climate reconstructions from the same lake or from nearby lakes. In addition, prior distributions for specific time periods of the past could also be created from grid points of the particular lake from spatial reconstructions included in PMIP. It also allows reconstructions to be supplemented with information about past permafrost. Other information could come from experts who determine relatively accurate climate conditions for a given age based on vegetation studies of sediments. The last example of possible $\mathbb{P}(\boldsymbol{C_p} \mid \cdot)$ is presented and discussed in Sect. 3.3.5. The second term allows us to insert constraints

on the reconstructed modern climate. We define the transition from modern times to the past at 0 cal a BP, i.e. 1950 AD. This is because the temporal resolution of 50 years limits us, as we can only define the years 2000 AD or 1950 AD as the most recent period. However, since most of the cores of the EPD were recovered exactly in this interval, a compromise can thus be found for the definition of the youngest time interval. For modern climate, we have measurements from which we can use probability distributions as anchors for the reconstructions. More specifically, we use the Climate Research Unit (CRU) data from the Koeppen-Geiger classification rather than just one grid point from each site. This gives us a higher climatologically based variability that should simulate the climate of the catchment area of the lake or mire in question. We use the normal distribution as an approximation for the annual temperature and the gamma distribution for the annual precipitation. Again, it can be decided individually what to use as the basis for $\mathbb{P}(\boldsymbol{C_m} \mid \cdot)$, with our choice leading to a variety of useful climate reconstructions (see Sect. 3.4.1). All in all, we refer to the above as the prior climate module, which can be summarized as follows:

$$\mathbb{P}(\boldsymbol{C_p} \mid \boldsymbol{C_m}, \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) = \begin{cases} \mathbb{P}(\boldsymbol{C_p} \mid \cdot) & \text{if } \boldsymbol{C_p} \text{ is available,} \\ \text{Unif}(1, ..., \text{N}_{\text{age}}) & \text{otherwise,} \end{cases} \tag{2.19}$$

$$\mathbb{P}(\boldsymbol{C_m} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) = \begin{cases} \Gamma(\text{P}_{\text{ANN},m}) \text{ and } \mathcal{N}(\text{T}_{\text{ANN},m}) & \text{if } \boldsymbol{A}(C_1) \leq 0 \text{ cal BP,} \\ \text{Unif}(1, ..., \text{N}_{\text{age}}) & \text{otherwise.} \end{cases} \tag{2.20}$$

This means that reconstructions can be carried out with fewer restrictions even without prior climate information. This is made possible by the use of uniform distributions that encompass the reconstruction period and thus all time slices N_{age}. How useful such reconstructions ultimately are varies from case to case. Examples of this are discussed in Sect. 3.3.1 and Sect. 3.3.6.

Finally, we consider the third term of Eq. 2.16 in detail:

$$\mathbb{P}(\boldsymbol{P} \mid \boldsymbol{A}, \boldsymbol{\Theta}) = \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{A}, \boldsymbol{\phi}, \boldsymbol{\psi}) \approx \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{A}, \boldsymbol{\phi}) \approx \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\phi}). \tag{2.21}$$

First, we use assumption 3 above and split the parameter component. Then we state that $\boldsymbol{P}$ is independent of $\boldsymbol{\psi}$ when no $\boldsymbol{C}$ is given. Finally, the updated taxa weights $\mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\phi})$ are determined under the assumption that they are conditionally independent of A and thus hold for the entire reconstruction period. At this point, taxa weights could be split temporally based on additional prior information, so that they differ for specific time periods (e.g. glacials/interglacials). This approach is not explored further in this thesis and could be included in future work. The last term of Eq. 2.16, which is the joint distribution of $\boldsymbol{A}$ and $\boldsymbol{\phi}$, is treated as follows. First, we consistently follow the above reasoning and assume that all parameters $\boldsymbol{\Theta}$ are a priori independent of $\boldsymbol{A}$. Thus, this distribution can be formulated as follows:

$$\mathbb{P}(\boldsymbol{A}, \boldsymbol{\Theta}) = \mathbb{P}(\boldsymbol{A}) \cdot \mathbb{P}(\boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{\phi}). \tag{2.22}$$

Here, the last probability distribution describes the prior taxa weights, which are directly related to equation 2.21. The second term contains the parameters of the transfer functions and $\boldsymbol{A}$ is assumed to be uniform distributed if no depth or distance information are available. We see that already in the local reconstruction module in Eq. 2.15, the relations between $\boldsymbol{A}$ and $\boldsymbol{D}$ are inserted into our reconstruction scheme. With all the reformulations and simplifications listed above, Eq. 2.16 can be summarized as follows:

$$\begin{aligned}\mathbb{P}(\boldsymbol{C}, \boldsymbol{\Theta} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{PP}) \propto \mathbb{P}(\boldsymbol{PP} \mid \boldsymbol{C}, \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{C_p} \mid \boldsymbol{C_m}, \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) \\ \cdot \mathbb{P}(\boldsymbol{C_m} \mid \boldsymbol{P}, \boldsymbol{A}, \boldsymbol{\Theta}) \cdot \mathbb{P}(\boldsymbol{\psi}) \cdot \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\phi}) \cdot \mathbb{P}(\boldsymbol{\phi}).\end{aligned} \tag{2.23}$$

Overall, the taxa should be weighted in such a way that those that better fit the constraints of the prior climate and proxy pool modules are weighted higher. How this is done in detail is described in the following. In the context of MCMC sampling, there are two ways to determine $\mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\phi})$. One of them is based on updating the weights through the random walk Metropolis-Hastings (rwMH). In the other case, an additional parameter allows the weights to be determined using Gibbs sampling. In the first case, $\boldsymbol{\phi}$ contains only the prior taxa weights $\boldsymbol{\omega}$. Without further prior information, we assume a uniform distribution across all taxa K at the beginning of the MCMC simulations:

$$\mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\phi}) = \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\omega}) = \text{Unif}(1, ..., \text{K}). \tag{2.24}$$

Such a Dirichlet distribution allows us to determine the taxa weights as we have requested above. This means that the taxa weights have values between 0 and 1 and add up to 1. The Jeffreys prior hyperparameters $\frac{1}{2}$ of this distribution give each taxon equal prior weight. Furthermore, these values provide a weaker constraint for determining the posterior taxa weights. This property follows directly from the characteristics of the Jeffreys prior (Gelman et al., 2013). Since a corresponding full conditional $\mathbb{P}(\boldsymbol{\omega} \mid \boldsymbol{P})$ does not follow a probability distribution from which we can sample directly, the taxa weights in this case are updated by a random walk Metropolis-Hastings.

In contrast, we proceed as follows with Gibbs sampling:

$$\mathbb{P}(\boldsymbol{P} \mid \boldsymbol{\phi}) = \mathbb{P}(\boldsymbol{P} \mid \boldsymbol{z}, \boldsymbol{\omega}) = \text{Unif}(1, ..., \text{K}). \tag{2.25}$$

The additionally introduced parameter $\boldsymbol{z}$ follows a categorical distribution, which leads to the following:

$$\mathbb{P}(\boldsymbol{z} \mid \boldsymbol{\omega}) = \text{Cat}(z_1, ..., z_\text{K} \mid \omega_1, ..., \omega_\text{K}). \tag{2.26}$$

This distribution requires additional information indicating how many taxa are drawn at once. We set this to t = K/2, which gives us the maximum number of distinct taxa combinations $\binom{\text{K}}{\text{t}}$. Here, too, the weights are ultimately determined by Eq. 2.30. Since the Dirichlet distribution is the conjugate prior to the categorical distribution, we can determine the full conditionals of $\boldsymbol{\omega}$ and $\boldsymbol{z}$. Specifically, $\boldsymbol{\omega}$

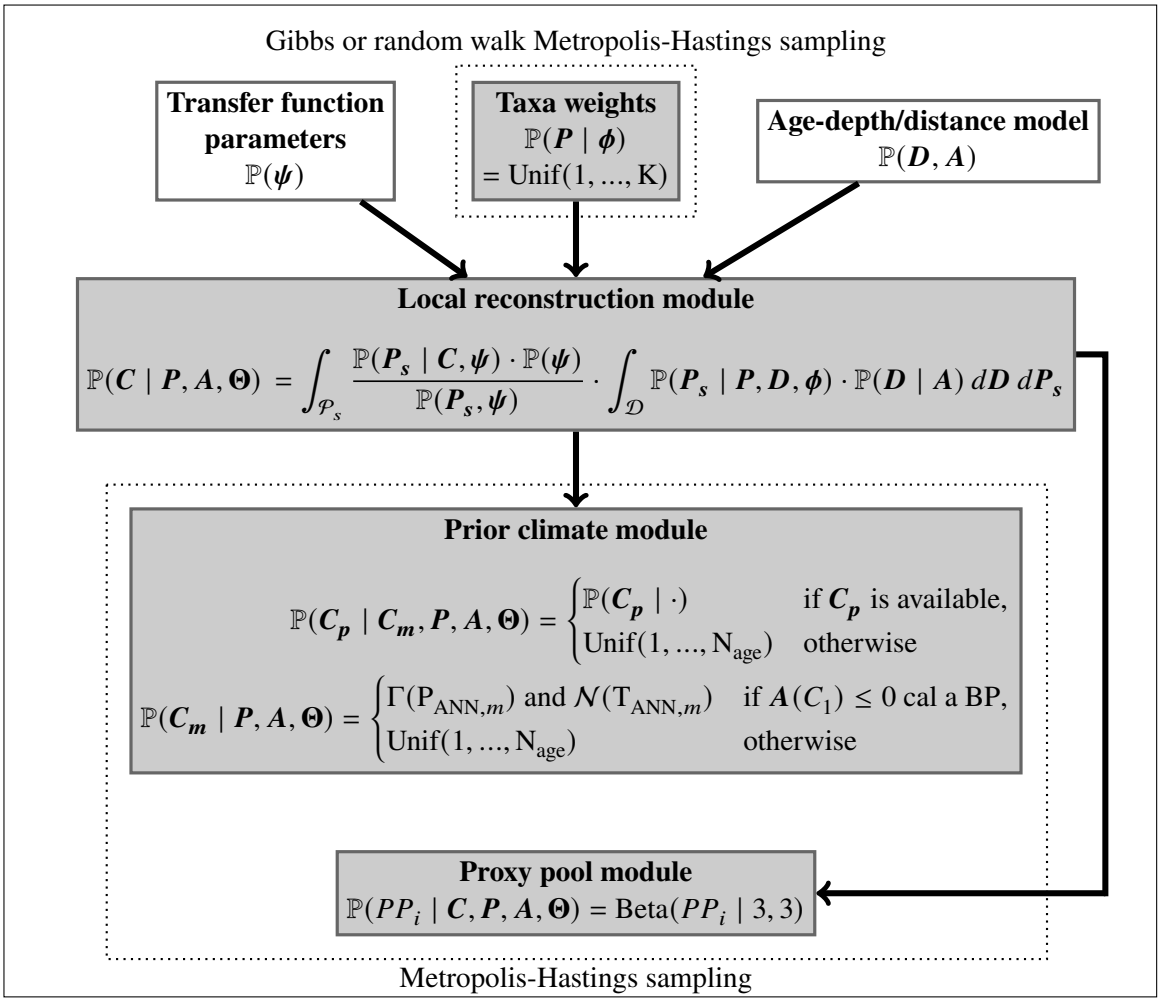


Figure 2.17: Directed acyclic graph of the Bayesian framework in Eq. 2.23. The gray boxes represent the quantities that will be inferred during MCMC sampling, and the white boxes contain fixed quantities. The corresponding arrows represent the mutual dependencies, with their direction pointing to the ascending hierarchy levels and the dashed boxes indicate the respective sampling procedures of the modules contained therein.

given $\boldsymbol{z}$ can be described as follows:

$$\mathbb{P}(\boldsymbol{\omega} \mid \boldsymbol{z}) = \mathrm{Dir}(\frac{1}{2} + z_1, ..., \frac{1}{2} + z_\mathrm{K}). \tag{2.27}$$

Moreover, we can determine the categorically distributed $\boldsymbol{z}$ given $\boldsymbol{\omega}$ and $\boldsymbol{P}$:

$$\mathbb{P}(\boldsymbol{z} \mid \boldsymbol{\omega}, \boldsymbol{P}) = \mathrm{Cat}(\alpha_1, ...\alpha_\mathrm{K}), \tag{2.28}$$

where, in our case, the following applies to each component α_i:

$$\alpha_{\mathrm{i}} \sim \omega_{\mathrm{i}}. \tag{2.29}$$

The respective weights are determined with the help of an additional prior distribution:

$$\mathbb{P}(\boldsymbol{\phi}) = \mathbb{P}(\boldsymbol{\omega}) = \mathrm{Dir}(\omega_1, ..., \omega_{\mathrm{K}} \mid \frac{1}{2}, ..., \frac{1}{2}). \tag{2.30}$$

Since we have determined a priori equally distributed weights, this relationship simplifies. However, if certain taxa are given a higher or lower weight at the beginning, this must be taken into account accordingly at this point.

Fig. 2.17 summarizes graphically how our local reconstruction framework works in general. The boxes in the first row contain the input variables, while the variables in the white boxes are not inferred during the MCMC simulation. For example, the parameters of the transfer functions are defined in Sect. 2.5. The middle upper gray box describing the inference of the taxa weights can be optionally determined via Gibbs or rwMH sampling. This is done by comparing the sampled climate reconstructions with other climate data and/or reference curves and constraining them accordingly. These comparisons are made using the Metropolis-Hastings sampling, represented by the lower dashed area.

2.5 Transfer functions via machine learning

This section describes the development of transfer functions. The main attention is paid to the probability distribution $\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi})$ from Eq. 2.15. We start with the idea of (Kühl et al., 2002), who uses multivariate Gaussian distributions for the presence of plants. For this purpose, analogous to this work, they took plant distribution maps as for the European olive tree in Fig. 2.18 (c) and assigned certain climate variables to them. In our case, these are the distributions of $\mathrm{P_{ANN}}$ and $\mathrm{T_{ANN}}$ over the years 1961 – 1990 AD from the CRU 4.01 dataset (Harris and Jones, 2017) and are shown in Fig 2.18 (a) and (b). These data, along with those from Fig. 2.18 (c), are presented in the climate feature space in Fig. 2.19. Sect. 2.1.1 already compiles the relationship between climate variability such as Heinrich, Bond, and DO events and isotopic information from speleothems. These fluctuations occurred within several decades to centuries and affected temperature and precipitation in regions of Eurasia and North America (Asmerom et al., 2010; Wagner et al., 2010; Genty, 2003; Benson et al., 1996). The assumed relationship between speleothems and these two climate variables differs from study to study. In the Levant, these are often used as proxies for the mean annual precipitation (Bar-Matthews et al., 2003; Cheng et al., 2015; Sánchez-Fernández et al., 2018; Bradley, 2014). However, in Central and Northern Europe it alternates between $\mathrm{T_{ANN}}$ and $\mathrm{P_{ANN}}$ (Mauri et al., 2015; Lauritzen and Lundberg, 1999). Since plants are also sensitive to such climatic variations (Gams,

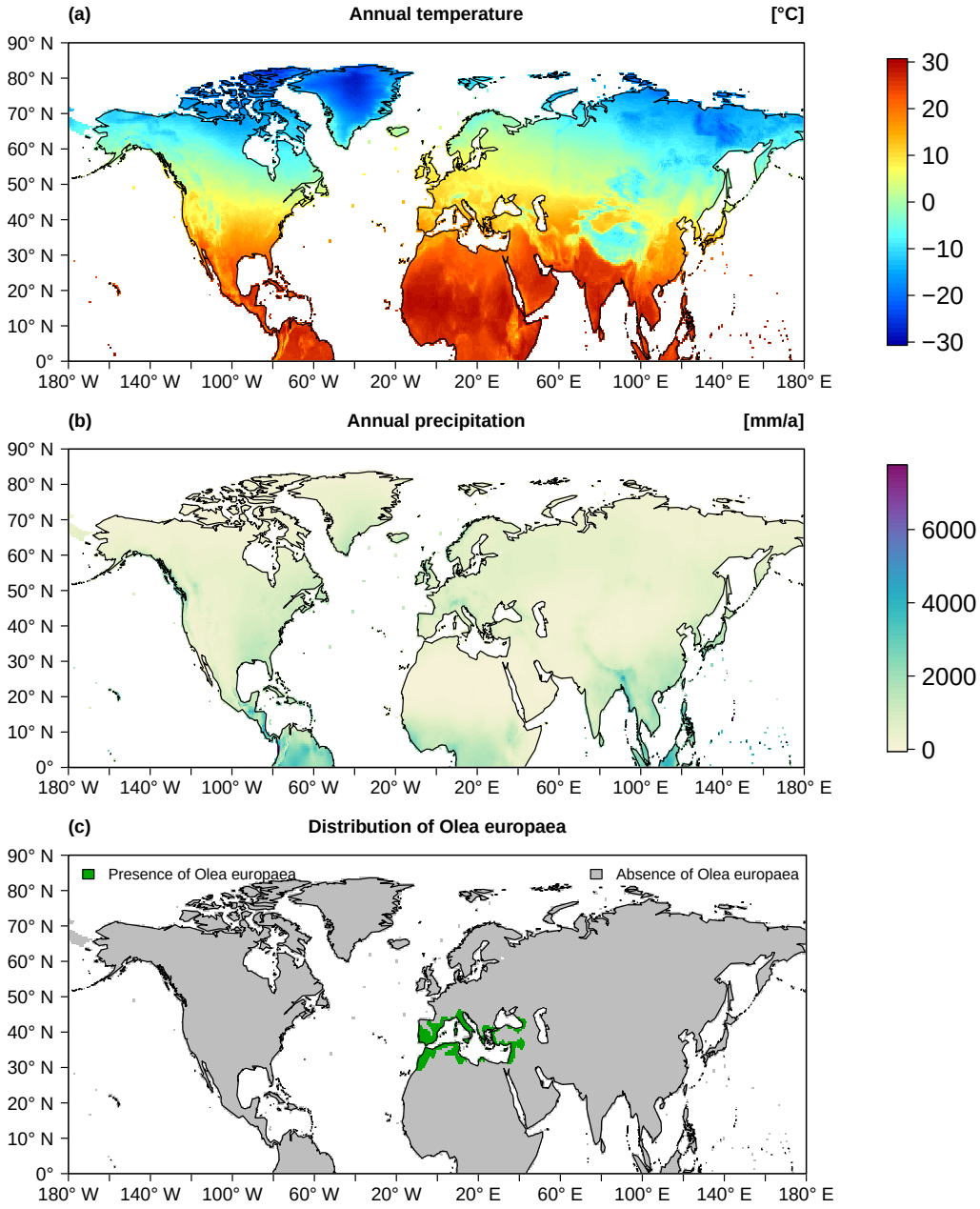


Figure 2.18: Maps of the Northern Hemisphere with annual precipitation (a), annual temperature (b), and the geographical distribution of Olea europaea (c). The climate data are based on those from CRU 4.01 and the plant distribution is according to Meusel et al., 1974.

1964), a relationship between plant distributions, temperature and precipitation variables is assumed in this work. This is done in some studies based on the PDF method (Kühl et al., 2002; Kühl and Litt, 2003; Gebhardt et al., 2008). Thus, the climate variables considered are intended to represent a compromise between the various proxies.

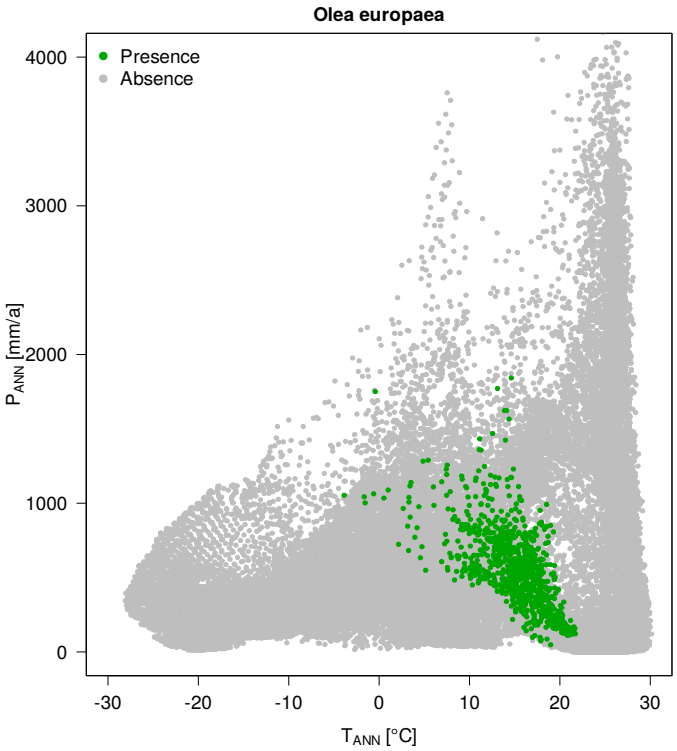


Figure 2.19: Summary of the data points from Fig. 2.18 (a) - (c) into a climate space that represents the feature space of the machine learning algorithms.

After Kühl et al., 2002 successfully used multivariate Gaussian distributions as transfer functions, Gebhardt et al., 2004 compared normal distributions, kernel densities, and Gaussian mixture models. Daniel Simonis, 2010 was the first to use precipitation in addition to temperature. Using the inverse cumulative distribution function (CDF) method (Gentle, 2004), these can be transformed into a standardized normal distribution. Standardization and scaling of input variables is generally an important concept in the computations of machine learning models (Butcher and Smith, 2020). Therefore, we also use the inverse CDF method before passing the climate data to the ML algorithms. To calculate the best number of components of the mixture model, Daniel Simonis, 2010 uses the Bayesian information criterion (BIC), in contrast to Gebhardt et al., 2004, who take the log-likelihood. Also in Chevalier et al., 2014, PDFs are applied as transfer functions that consider only the presence information in the climate space (feature space). Stolzenberger, 2011, Stolzenberger, 2017, Thoma, 2017, and Weitzel et al., 2019 instead used logistic regression that accounts for not only the presence but also the absence of plants. In Litt et al., 2012, Stolzenberger, 2017 and in Thoma, 2017, Quadratic Discriminant Analysis (QDA) is applied with respect to three biomes. The absence of these biomes (or unknown biomes) has not been used as an additional feature in previous work.

The objective of this work is to systematically test a variety of possible methods and select the most

appropriate algorithm for the task at hand. For this purpose we use the R package caret (Jed Wing et al., 2019). This stands for **c**lassification **a**nd **re**gression **t**echniques and provides a variety of models that can be used to solve corresponding problems. In our case, it is a binary (presence or absence of a taxon) classification problem. The package provides a simple way to compare the selected models via cross-validation. In this process, the provided data (cf. Fig. 2.18) are split into a training and a validation dataset. Cross-validation is performed on the training data (James et al., 2013), which in our case accounts for 70 % of all data. Accordingly, this training data is split into a certain number (k-folds) of training and validation datasets. This partitioning process can be repeated as often as desired and leads to statistical verification distributions from which the performance of the models is derived. Cross-validation is also performed for a certain number of different parameters for the respective machine learning algorithms (model tuning). The entire process is very easily accessible in caret and runs completely automatically after the initial parameters have been defined. The remaining 30 % (hold-out set) are used to validate the models obtained by cross-validation on the remaining 70 %. This has the advantage that they can be tested on an independent data set, further minimizing the risk of overfitting. More precisely, the winner is determined from the first half and its final score from the second half of the hold-out set. Thus, it is quite possible that the final selected transfer function differs from the winner of the cross-validation.

As shown in Fig. 2.18 (c), the groups for presence (minority class) and absence (majority class) are unbalanced. One could reduce the size of the entire map section so that the groups are balanced. However, the models then deliver significantly worse and sometimes more unrealistic results. This problem is discussed for example in Thoma, 2017 or in Weitzel et al., 2019. Thus, a model could provide higher probabilities of occurrence on the one hand where the plant does not occur in the feature space, and on the other hand where the climate values are biologically unrealistic. When the map section is enlarged, this problem recedes, especially if the absence values can serve as a boundary. This is the case when the occurrence domain is enclosed by the absence domain in the feature space. The reduction of the map section is analogous to the techniques of random under-sampling (Hoens and Chawla, 2013), wherein the majority class is randomly reduced to the size of the minority class, potentially losing important information. In contrast, random oversampling of the minority class risks overfitting. To solve this problem, the **S**ynthetic **M**inority **O**versampling **Te**chnique (SMOTE) is used (Bowyer et al., 2011). Here, a minority class instance a is first randomly selected and its k-nearest minority class neighbors are determined. A line segment is then formed between one randomly selected k-nearest neighbor b in feature space. Here, a synthetic instance of the minority class is created by selecting a random point along this line (Hoens and Chawla, 2013). SMOTE can only do this with one minority class at a time. The new synthetic data thus generated are the input data for the calculation of the transfer function in the ML competition. It should be noted that only the training data is processed with SMOTE. For the model verification on the hold-out set, the original data are used. The following is a brief description of the ML algorithms used in this work.

Generalized Linear Model (GLM)

Logistic regression, listed under the GLM, is similar to linear regression (Agresti, 2003). However, here the outcome variable is categorical and not continuous (Fahrmeir and Tutz, 1994). In logistic regression, it is assumed that the outcome variable follows a binomial distribution and the logistic probability (logit) is accordingly described by a linear function of the logistic regression coefficients $\boldsymbol{\psi}$. The inverse function of the logit is also called the logistic function and is:

$$\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi}) = \frac{\exp(\boldsymbol{C\psi})}{1 + \exp(\boldsymbol{C\psi})}. \tag{2.31}$$

Regression coefficients are obtained using maximum likelihood estimation for each taxon (Agresti, 2003). The values of the GLM can be interpreted as probabilities for the presence or absence of the taxa. In this thesis, quadratic logistic regression is used analogously to Stolzenberger, 2017, Thoma, 2017 and Weitzel et al., 2019. The resulting bilinear terms are crucial for additional flexibility in the feature space and can be written as follows for the first selected taxon $P_{1,s}$:

$$\mathbb{P}(P_{1,s} \mid \boldsymbol{C}, \boldsymbol{\psi}) = \text{logit}(\psi_1^T + \psi_2^T C_1 + \psi_3^T C_2 + \psi_4^T C_1 C_2 + \psi_5^T C_1^2 + \psi_6^T C_2^2). \tag{2.32}$$

When these quadratic terms are not used, the GLM performs significantly worse than the other ML algorithms (not shown), even though they do not use the additional information. They are deprived of these, as this hardly leads to any improvement (not shown). The GLM therefore always has an advantage, but should still be compared in order to be able to establish a reference to the earlier studies.

Artificial Neural Network (ANN)

In the broadest sense, the idea of Artificial Neural Network dates back to the 18th century (Schmidhuber, 2015). The structures described there are variants of logistic regression. Such an ANN is also called a "logistic perceptron" (Spackman, 1991). The input values (here: climate variables) are transferred to the output via a simple interface. Analogous to Eq. 2.31, the input is provided with weights (regression coefficients) and processed via a so-called activation function (here: logistic function). The counterpart to the coefficients of the axis intercepts is called bias in ANNs. The further development of ANNs began in the middle of the 20th century (Mcculloch and Pitts, 1943). The architecture of these networks was modeled on the biological nervous system. The different layers consist of artificial neurons (interfaces) that "fire" at certain values using the activation functions and thus pass on their signal. In addition to the input and output layer, a so-called hidden layer is also used. This allows, in contrast to the logistic perceptron, the modeling of more complex relationships between input and output (Ayer et al., 2010). Over the course of time, the network structure has been continuously developed (Schmidhuber, 2015). Among the best known networks is the recurrent neural network (Rumelhart et al., 1986). In contrast to the feedforward neural network (FNN), the

connections of neurons of one layer to neurons of the same or a previous layer are guaranteed. This opened up the possibility that ANNs could also be used for speech and character recognition. In image recognition the so-called Convolutional Neuronal Networks are used (LeCun et al., 1989). For a detailed description of the history of ANNs, the reader is referred to the work of Schmidhuber, 2015.

Comparatively simple classification problems arise in this work, so relatively simple network structures can be used. These give similarly good results with much less computational cost, and the risk of overfitting is generally lower with simpler structures. The caret package offers a wide range of different ANNs, including those based on Python. After initial tests, the ANN from the R package nnet is chosen in this thesis. It is an FNN that allows one hidden layer with an arbitrary number of hidden neurons (Venables and Ripley, 2002). This enables much more complex correlations to be found than with the GLM described above. Moreover, for our classification problem, it gives partially better results than the multilayer perceptrons provided by the caret package and the training time is also lower.

Discriminant analysis

Discriminant analysis involves the development of discriminants, i.e. linear combinations of independent variables that discriminate the categories of the dependent variable (James et al., 2013). Linear Discriminant Analysis (LDA), for example, first extracts linear constructs (discriminants) that maximize separation between groups and then uses them to perform a Gaussian classification. This step can also be used as data reduction. At baseline, LDA assumes normality as well as variance-covariance homogeneity of the groups. Then Bayes' theorem (see Eq. 2.7) is used to assign the observations to groups based on the discriminants. In our case, a corresponding equation could look like this:

$$\mathbb{P}(\boldsymbol{P_{1,s}} \mid \boldsymbol{C}, \boldsymbol{\psi}) = \frac{\mathcal{N}(\boldsymbol{C} \mid \boldsymbol{\mu_1}, \boldsymbol{\Sigma}) \cdot \mathbb{P}(\boldsymbol{P_{1,s}})}{\sum_{i=1}^{2} \mathcal{N}(\boldsymbol{C} \mid \boldsymbol{\mu_i}, \boldsymbol{\Sigma}) \cdot \mathbb{P}(\boldsymbol{P_{i,s}})}.$$

$\boldsymbol{\mu_1}$ is the estimated mean value of group 1, which is assumed to be multivariate normally distributed, e.g. the occurrence of plants. Group 2 is interpreted accordingly as absence. $\boldsymbol{\Sigma}$ is the joint covariance matrix and $\boldsymbol{P_{i,s}}$ the respective population group proportions (Johnson and Wichern, 2007) of the groups. The Bayesian approach, based on discriminants, allows the use of separate group discriminant matrices in addition to the standard use of a covariance matrix. This brings us to Quadratic Discriminant Analysis (QDA). QDA is a modification of LDA that accounts for heterogeneity in the covariance matrices of the groups. Therefore, the covariance matrices $\boldsymbol{\Sigma_i}$ estimated separately for each group are used here:

$$\mathbb{P}(\boldsymbol{P_{1,s}} \mid \boldsymbol{C}, \boldsymbol{\psi}) = \frac{\mathcal{N}(\boldsymbol{C} \mid \boldsymbol{\mu_1}, \boldsymbol{\Sigma_1}) \cdot \mathbb{P}(\boldsymbol{P_{1,s}})}{\sum_{i=1}^{2} \mathcal{N}(\boldsymbol{C} \mid \boldsymbol{\mu_i}, \boldsymbol{\Sigma_i}) \cdot \mathbb{P}(\boldsymbol{P_{i,s}})}.$$

In this thesis, QDA provides much better results compared to LDA due to the higher flexibility of the respective plant distributions (not shown). The latter is therefore removed from the competition. Next,

we consider Mixture Discriminant Analysis (MDA). MDA can be seen as an extension of LDA and thus of the work of Kühl et al., 2002. It modifies the within-group multivariate density of predictors by a mixture (i.e., a weighted sum) of multivariate normal distributions (Rausch and Kelley, 2009):

$$\mathbb{P}(\boldsymbol{P_{1,s}} \mid \boldsymbol{C}, \boldsymbol{\psi}) = \frac{\sum_{j=1}^{G_1} \gamma_{j,1} \mathcal{N}(\boldsymbol{C} \mid \boldsymbol{\mu_{j,1}}, \boldsymbol{\Sigma_{j,1}}) \cdot \mathbb{P}(\boldsymbol{P_{1,s}})}{\sum_{i=1}^{2} \sum_{j=1}^{G_i} \gamma_{j,i} \mathcal{N}(\boldsymbol{C} \mid \boldsymbol{\mu_{j,i}}, \boldsymbol{\Sigma_{j,i}}) \cdot \mathbb{P}(\boldsymbol{P_{i,s}})}.$$

G represents the clusters and $\gamma_{j,i}$ the mixture weights within each group. In principle, this approach brings two further advantages over LDA and QDA: (1) more accurate classification by modeling multivariate nonnormality or nonlinear relationships between variables within each group, and (2) the possible detection of subgroups in each group (Rausch and Kelley, 2009).

Decision tree-based algorithms

Algorithms based on decision trees are often used for classification of ecological data (De'ath, 2007). They are divided into so-called branches and leaves, which subdivide the feature space by yes-no queries. The disadvantage of simple decision trees is their inaccuracy. Often their prediction accuracy is significantly worse than the achievable (Hastie et al., 2009). This can be increased dramatically in some cases if bagged or boosted trees are used. The random forest, for example, uses bagging based on decision trees (Breiman, 2001). This typically involves determining several hundred bootstrap datasets from the feature space. If multiple features are available, subsets are drawn randomly from these datasets and the decision trees are finally calculated on the basis of these subsets. This abundance of different trees forms the random forest. Bootstrapping the data and using the aggregate to make a decision is called bagging. Now, when a prediction is made with a new data point, the votes of all decision trees are collected together, from which a probability for that data point can be determined. However, in the present case, the random forest is not the most appropriate algorithm because the dataset consists of the two features T_{ANN} and P_{ANN}. If more features were available, it would make sense to include the random forest in such a machine learning competition.

Boosted trees are similar to random forests in that a number of decision trees are computed and considered. However, only shallow trees are considered here. For example, in Adaptive Boosting (Freund and Schapire, 1996), so-called stumps (trees with two leaves) are almost always used as "weak" learners (Friedman et al., 2000) and in the Gradient Boosting Machine (GBM) (Friedman, 2001), the "depth" of the trees is also specified at the beginning. To reduce bias and variance, each decision tree is determined as a function of its predecessor and weighted according to its accuracy. This combining of many model results can decrease bias and, to some extent, variance. However, boosting also loses some advantages over simple decision trees, such as speed, interpretability, and, in the case of Adaptive Boosting, robustness to overlapping group distributions and, most importantly, mislabeling of the training data. A GBM is a generalization of tree-boosting that attempts to mitigate

these problems to create an accurate and effective standard procedure (Hastie et al., 2009). For this reason, GBM is chosen in this thesis to introduce a machine learning algorithm based on decision trees.

Support Vector Machines (SVMs)

Support Vector Machines determine a function that divides the input data into classes (Jergensen et al., 2020). In our case, the support vectors are the correctly classified data points closest to the separation line between the two classes. The distance between this line and the support vectors is called margin, which is maximized. Although these algorithms provide the best results in some cases, they consistently failed in comparison with the other algorithms because their computation time is disproportionately long. Similar difficulties with SVMs are also found in Jergensen et al., 2020, where a ML competition for forecast models of convective storms is presented. For this reason, we exclude SVMs from the competition.

2.6 Spatial climate reconstructions

Finally, in this work we aim to create climate field reconstructions (CFRs) based on our local climate reconstructions. Let us first briefly review the history of spatial reconstructions in the paleoclimatological context.

For reconstructions within the last two millennia, there are a number of methods based on regression techniques (Mann et al., 1998). Other examples include Cook et al., 1999 with point-to-point regression, Ammann and Wahl, 2007 with principal component regression, Arrigo et al., 2006 with univariate linear regression, Hegerl and Zwiers, 2011 with total least squares, and Mann et al., 2008 with regularized expectation maximization. A different approach is used by Gebhardt et al., 2008. They combine pollen and macrofossil samples with an advection-diffusion model, which is driven by insolation changes between the present and past time slices. A corresponding cost function is minimized and thus the climate of the early Eemian (ca. 125 cal ka BP) is reconstructed. Simonis et al., 2012 built on this idea and extended the model by importance sampling and reconstructed the periods 6 cal ka, 8 cal ka, 12 cal ka, and 13 cal ka. Davis et al., 2003 and Mauri et al., 2015, in turn, calculate CFRs for different time periods using pollen samples primarily from EPD. To spatiotemporally interpolate these, they use 4D thin plate splines.

Stolzenberger, 2017 and Weitzel et al., 2019 combine the pollen and macrofossil dataset of Simonis et al., 2012 with a MH climate simulation ensemble from the Paleoclimate Modelling Intercomparison Project 3 (PMIP3; Braconnot et al., 2011). Each individual model setup is considered a separate ensemble member. The entire procedure aims to constrain the spatial structures of physically possible climate states for a given external forcing. In particular, Weitzel et al., 2019 systematically investigate this procedure using a Bayesian framework. In doing so, they present several ways in which spatial

Table 2.3: Basic information on the 7 CMIP5/PMIP3 simulations used.

Model	Institute	Atmospheric grid	Simulated years (HIST / MH / LGM)
CCSM4	NCAR	$288 \times 192 \times L26$	156 / 301 / 101
MRI-CGCM3	MRI	$320 \times 160 \times L48$	156 / 100 / 100
CNRM-CM5	CNRM-CERFACS	$256 \times 128 \times L31$	156 / 100 / 200
MPI-ESM-P	MPI-M	$192 \times 96 \times L47$	156 / 100 / 100
GISS-E2-R	NASA/GISS	$144 \times 90 \times L40$	156 / 100 / 100
IPSL CM5ALR	IPSL	$96 \times 96 \times L39$	156 / 500 / 200
FGOALS-g2	LASG-CESS	$128 \times 60 \times L30$	156 / 685 / 100

prior samples can be generated from the ensemble members. The simplest approach is the Gaussian Model (GM). Here, independent and identically Gaussian distributed simulations are assumed. A somewhat more complex model is referred to as a Regression Model (RM). This involves inferring weights for each PMIP simulation using the random walk Metropolis-Hastings sampling. As a result, simulations that better fit the proxy data are given more weight. Another approach to determine such weights is described in the so-called Kernel Model (KM). This technique allows the description of full conditionals and thus the use of Gibbs sampling. Two ways of determining the underlying covariance matrix of the ensemble are also presented. One uses a graphical lasso (Glasso) to estimate sparse inverse covariance matrices. The other way is the so-called shrinkage approach. In the latter case, a weighted average of the empirical correlation matrix of a climate simulation ensemble and a reference correlation matrix is used. This reference is based on a stationary Matérn correlation matrix and could include additional spatial modes. The results in terms of Euopean MH show that the shrinkage-based covariance matrix performs better than the Glasso technique. Also, GM and RM are less biased and more robust to KM. Ultimately, they choose RM because it uses the derivation of model weights, which provides a more flexible method of interpolating the proxy data with the PMIP ensemble.

The method from Weitzel et al., 2019 is applied in this work to calculate the CFRs for the HIST Period, the MH, and the LGM. Again, a multi-model ensemble of climate simulations within the PMIP3 is used as the climate prior. We select the models so that they are available for all 3 time slices. For the HIST period, these are from the Coupled Model Intercomparison Project 5 (CMIP5). Table 2.3 gives an overview of these 7 ensemble members and some important features. To ensure that the simulations adequately reflect the respective time periods, they are updated with the appropriate orbital settings, greenhouse gas concentrations, and land-sea and ice sheet distributions (Abe-Ouchi et al., 2015). After each model reaches the so-called equilibrium state (spin-up), the time periods given in Table 2.3 are simulated. To account for the different spatial resolutions and reconcile them with the distributions of the local climate reconstructions, we use a common grid of 2.5° by 2.5° generated by bilinear interpolation. Finally, the climatological means of T_{ANN} and P_{ANN} are extracted. In order to adequately account for precipitation values, we use the anomalies of the corresponding HIST simulations of CMIP5 with respect to the reference period 1961 – 1990 AD. A similar approach is

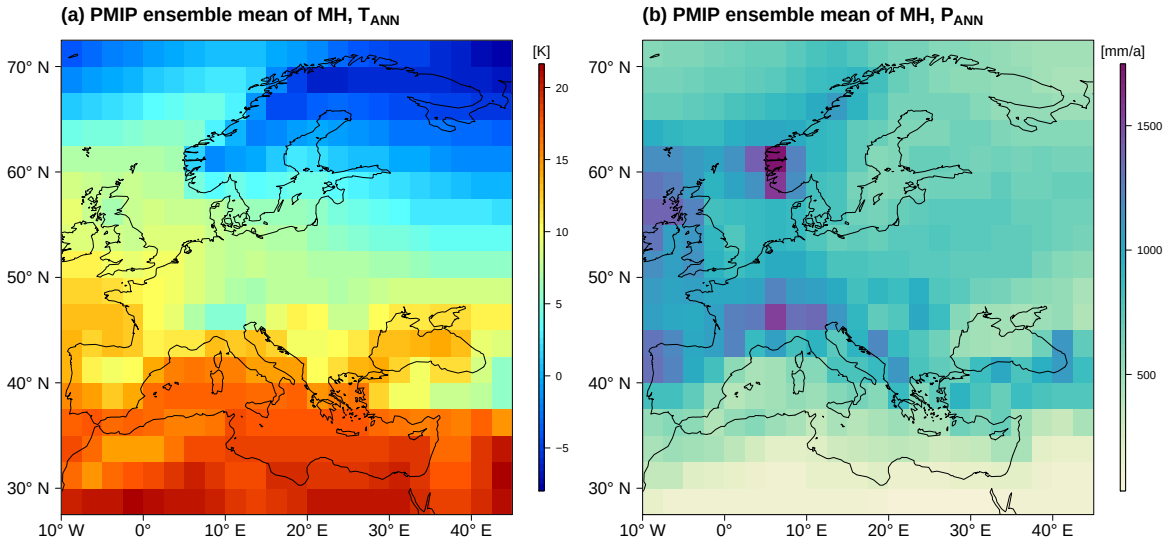


Figure 2.20: Summary of the PMIP ensemble over the area of the CFRs considered in this work. Panel (a) shows the mean annual temperature of the PMIP ensemble for the MH. In (b) the corresponding annual precipitation is depicted.

also used by Weitzel et al., 2019 to approximate the assumed normality of the simulations. Using this technique, we succeed in assuming normality on average for each multi-model ensemble distribution. To determine this, we use a one-sample Kolmogorov–Smirnov test (Simard and L'Ecuyer, 2011). Fig. 2.20 shows the study area of this work in terms of the reconstructed variables. In order to include some local climate reconstructions in the Near East, we have extended the field in this respect compared to Stolzenberger, 2017 and Weitzel et al., 2019.

The Bayesian framework in Weitzel et al., 2019 needs to be reformulated for this work, as we combine local climate reconstructions with the PMIP ensemble rather than transfer functions. We do this by analogy with the approach in Weitzel, 2020, where local climate reconstructions update a PMIP ensemble in Siberia for the LGM. First, we distinguish between spatial and local climate information with subscripts $\boldsymbol{S}$ and $\boldsymbol{L}$. So we can set up the following BHM:

$$\mathbb{P}(\boldsymbol{C_S}, \boldsymbol{\Omega} \mid \boldsymbol{C_L}) \propto \underbrace{\mathbb{P}(\boldsymbol{C_L} \mid \boldsymbol{C_S}, \boldsymbol{\Omega})}_{\text{Data level}} \cdot \underbrace{\mathbb{P}(\boldsymbol{C_S} \mid \boldsymbol{\Omega})}_{\text{Process level}} \cdot \underbrace{\mathbb{P}(\boldsymbol{\Omega})}_{\text{Prior level}} . \tag{2.33}$$

A directed acyclic graph (DAG) of this equation is shown in Fig. 2.21. The labels of the different levels follow Tingley and Huybers, 2010. The data level contains the climate information of our local reconstructions $\boldsymbol{C_L}$ for the particular time period and grid point considered. These are integrated there into the framework of the spatial reconstruction. In this work, $\boldsymbol{C_L}$ is embedded in the BHM in the form of so-called lookup tables (LUTs) to achieve a tradeoff between accuracy and sampling

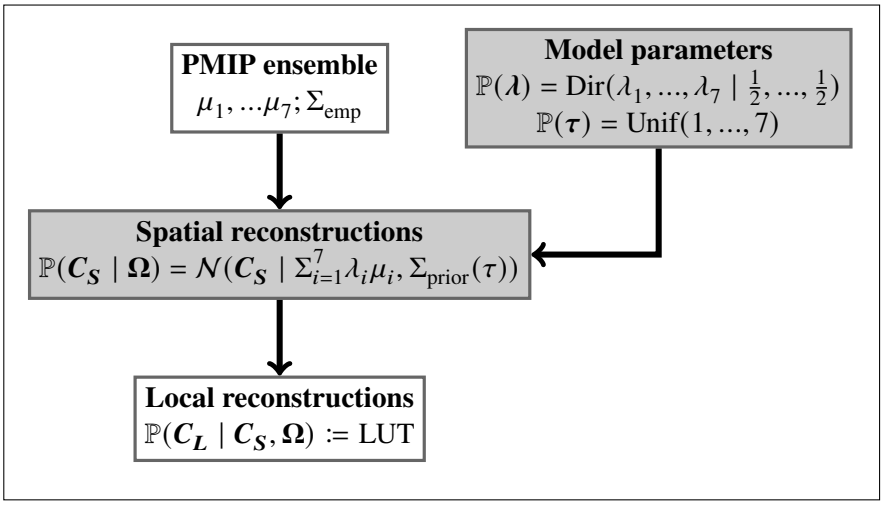


Figure 2.21: Directed acyclic graph of the Bayesian framework in Eq. 2.33. The gray boxes represent the quantities that will be inferred during MCMC sampling, and the white boxes contain fixed quantities. The corresponding arrows represent the mutual dependencies, with their direction pointing to the ascending hierarchy levels.

speed. The process level interpolates the local with the spatial climate information $\boldsymbol{C_S}$. Based on the research of Weitzel, 2020 regarding the LGM and MH, we use the Regression Model and shrinkage approach. This is because RM is the most flexible model and the shrinkage covariance matrix Σ_{prior} provides additional degrees of freedom compared to the empirical covariance matrix Σ_{emp}, which allows for better reconstruction of regional heterogeneity. To complete the BHM, the model parameters are defined at the prior level. Ω contains the a priori independent variables λ and τ. The former describes the weights of the 7 PMIP simulations via a Dirichlet distribution with Jeffreys prior hyperparameter $\frac{1}{2}$. τ parameterizes the parameter sets of the Matérn correlation model in terms of a uniform distribution. As with Weitzel et al., 2019, these are determined for each reconstruction by first applying the shrinkage model to each member of the climate simulation ensemble. In this process, all other members are taken into account, and the resulting set of parameters is passed to the reconstruction framework shown in Fig. 2.21.

CHAPTER 3

Results

3.1 Machine learning competition

In this section we look at the results of the machine learning competition in detail. Table 3.1 gives an overview of the algorithms used for this purpose. The abbreviations given are used accordingly in the following. In the evaluation, the focus is on the problem of unbalanced data sets. These are shown in Fig. 3.1 (a) for each of the classification problems considered here. More precisely, we see the percentage ratio between presence and absence of 279 plant taxa. The green vertical line corresponds to the values of Olea europea (cf. Fig 2.18 (c)) of almost 2 %, which are shown graphically in a feature space in Fig. 3.1 (b). The prediction of the final ML model is also shown with black contour lines. Such unbalanced data sets are expanded with SMOTE until all values from Fig. 3.1 (a) are at ca. 50 %. Subsequently, the models from Table 3.1 are trained on these data sets and finally evaluated with a fraction of the original data, marked with light red dots in Fig. 3.1 (b). In our work, this ranking is based on what is known as balanced accuracy (BA). How this is determined using the example of Olea europaea is shown in Table 3.2. This contingency table compares the original data approach (dots in Fig. 3.1 (b)) with the corresponding predictions of the final selected model (contours in Fig. 3.1 (b)). For the latter, values above 0.5 correspond to presence and values below 0.5 represent

Table 3.1: Machine learning algorithms used for the competition:

Algorithm:	Abbreviations:	Citation:
Artificial Neural Networks	NNET	Venables and Ripley, 2002
Stochastic Gradient Boosting	GBM	Greenwell et al., 2019
Generalized Linear Model	GLM	R Core Team, 2018
Quadratic Discriminant Analysis	QDA	Venables and Ripley, 2002
Mixture and Flexible Discriminant Analysis	MDA	Leisch et al., 2017

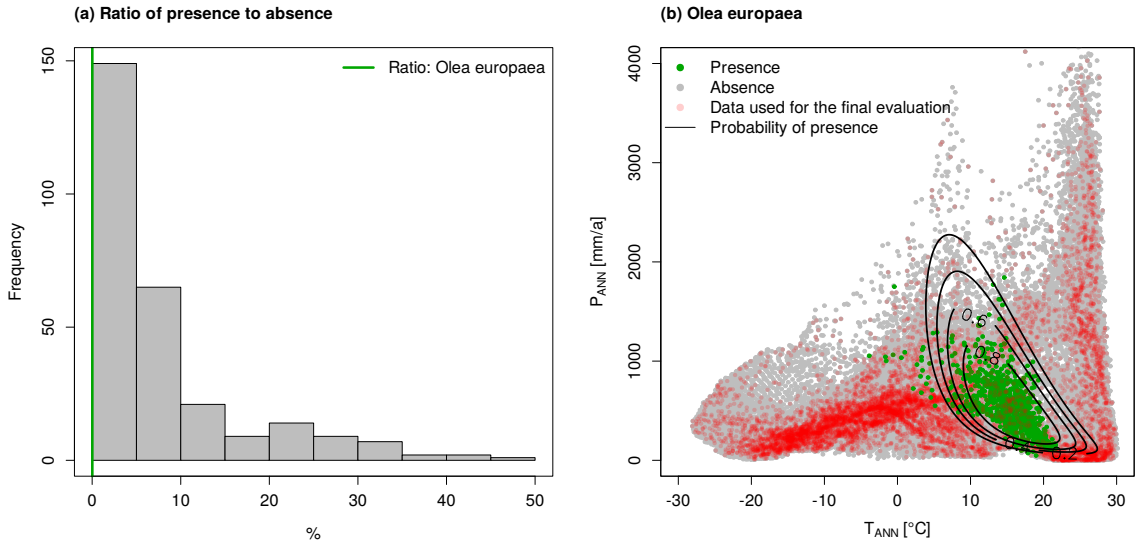


Figure 3.1: In (a), the percentage ratio of grid points corresponding to the presence or absence of all available plant distributions is shown. In (b), the respective climate feature space of Olea europaea is depicted along with the contour lines of a QDA-based transfer function. The latter is ultimately evaluated using the data points highlighted in light red.

absence. Together with the observation values, true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values are thus determined in Table 3.2. From this, in turn, the true positive rate $\frac{TP}{TP+FN}$ and true negative rate $\frac{TN}{TN+FP}$ can be calculated, which are referred to as sensitivity and specificity, respectively (Chicco et al., 2021). The arithmetic mean of these two measures is the balanced accuracy, which is an appropriate metric for trained ML models designed to describe an unbalanced data set (Brodersen et al., 2010). There is also the commonly used F1 score, which is also used for such classification problems (Chicco et al., 2021). Here, the focus is more on the positive data. However, we do not want to give preference to these, but consider the information about presence and absence to be equally important. For this reason, only the BA is used in this work.

The results of all trained models for each taxon are shown in Fig. 3.2 (a). A distinction is made between models trained on the original data set (without SMOTE) and those trained on data augmented

Table 3.2: Contingency table summarizing the information from Fig. 3.1 (b). In addition, the corresponding metrics for sensitivity, specificity, and balanced accuracy are provided.

		Observation	
		Presence	Absence
Prediction	Presence	135 (TP)	1443 (FP)
	Absence	10 (FN)	6515 (TN)
	Sensitivity/Specificity =	0.93 /	0.82
	Balancend Accuracy =	0.875	

with SMOTE. It is immediately noticeable that the results marked by the red boxplots have a BA of 0.5 in most cases. In these instances, the sensitivity is always 0 (TP = 0) and the specificity is 1 (FP = 0), which means that no presence is predicted (not shown). In contrast, the other models (green boxplots) have an average BA of about 0.9, which is a significant increase. This is not really surprising, as we define the entire Northern Hemisphere as the climate region at the beginning (see Fig. 2.18) and most taxa from our pool occur in the European region. Nevertheless, through this data augmentation and subsequent SMOTE, we can not only obtain fits with high significance (true positive rates), but also reduce the boundary effects in the feature space, resulting in more closed contour lines as in Fig. 3.1 (b). When selecting the best model for each taxon, regardless of the algorithm, we obtain BAs higher than 0.7, which are summarized in the gray boxplot. We use these as input to our local reconstruction scheme (see Fig. 2.17). Here, the great advantage of not relying on a single ML algorithm becomes immediately apparent, but rather comparing a large number of algorithms in a competition and selecting the best ones.

Finally, let us divide the models used into their different algorithms. This is shown in Fig. 3.2 (b), where we find the relative occurrence of each model. NNET is clearly in the majority with around 70 %. GBM then follows in second place with about 13 % and GLM close behind. In fourth place we find QDA with around 5 % and finally MDA with ca. 1 %. Although almost all algorithms provide at least usable trained models on their own (cf. Abb. 3.2 (a)), the direct comparison between them yields the clear result that a simple artificial neural network emerges as the clear winner. In particular, compared to GLM and QDA, which are commonly used for such classification problems,

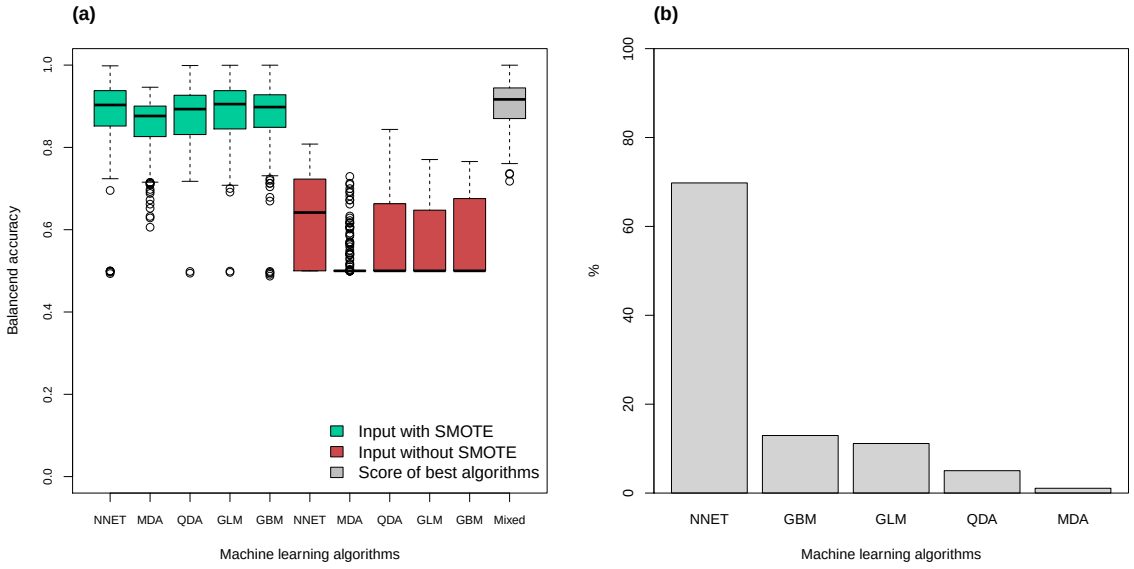


Figure 3.2: Panel (a) summarizes the balanced accuracy of all ML algorithms. In (b), the relative frequency of winning algorithms per competition is shown.

NNET can achieve a slightly higher BA. Nevertheless, the computational cost for training an NNET is significantly higher than that of QDA, which also provides good results. For example, the contour lines in Fig. 3.1 (b) are based on QDA, which achieves the highest BA for Olea europaea.

3.2 MCMC performance

In addition to the isotope data from the Mediterranean Sea (Medstack), from globally distributed ocean cores (LR04) and from Greenland (NGRIP) described above, proxy data from various lake and mire sediments and from caves are also used in this study. Specifically, these are the pollen records and the speleothem isotope data from the caves shown in Fig. 3.3. If a Bayesian age-depth model can be calculated for a sediment profile from the European Pollen Database, both MCMC methods are applied to it. These are the Metropolis-within-Gibbs (MG) and random walk Metropolis-Hastings

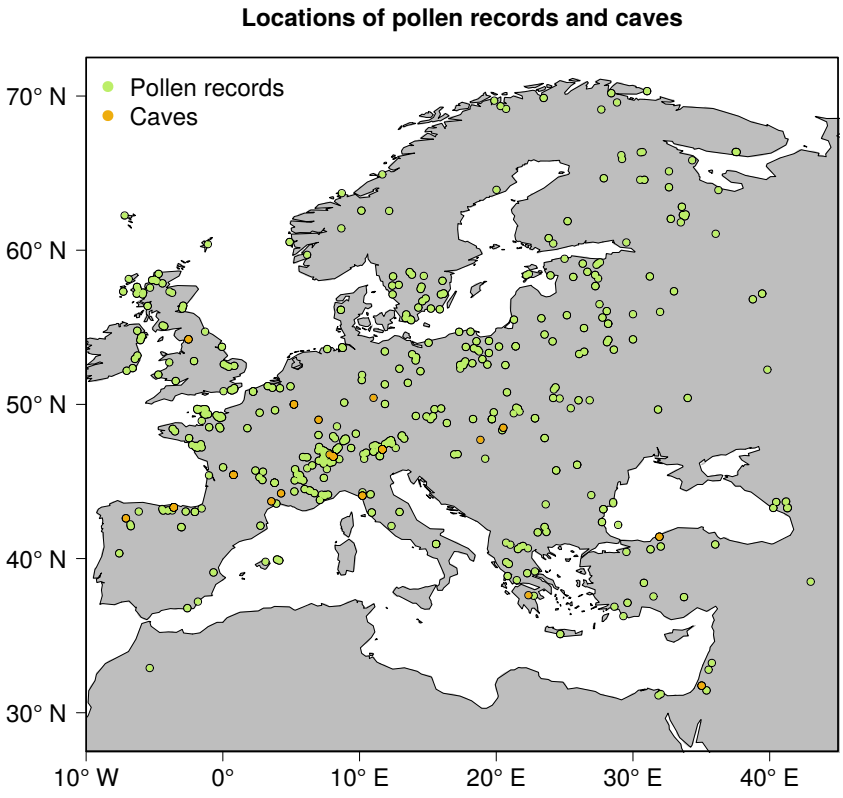


Figure 3.3: This map shows the locations marked with green dots where MCMC simulations are performed. These describe a subset of the European Pollen Database and include fossil and modern pollen records from natural archives such as lake and mire sediments. The brown dots show the caves where we take isotopic information from speleothems. These are a subset of SISAL.

algorithms. If caves are within 150 km of a given location with pollen records, our reconstruction tool compares them according to the lowest mean age resolution, the highest common age range of sediment and speleothem, and the shortest distance between them. Finally, an entity is chosen via a corresponding Pareto optimum. This describes the best possible state of all the characteristics mentioned, whereby it is not feasible to improve one of the characteristics without simultaneously worsening another (Langenbrunner and Neelin, 2017). In our case, the brown dots in Fig. 3.3 show the caves used in this work. The respective data come from the speleothem database SISAL.

Our study includes a total of more than 600 MCMC simulations. Below we summarize some important features of these final results by looking at the most common MCMC indicators. To this end, all simulations are first examined and then a reconstruction of Lake Prespa is considered in more detail.

3.2.1 All simulations

Fig. 3.4 shows some insights into all MCMC simulations. First, it is noticeable that the average computation time in (b) is higher than in (a). This is caused by an increase from 12,500 (MG) to 50,000 (rwMH) iterations and is necessary because of the higher integrated autocorrelation time (IAT) shown in (g) and (h). IAT is an estimate of the average number of iterations for drawing an independent sample in a Markov chain (Christen and Fox, 2010). In addition, two further MCMC diagnoses are given, namely the Gelman-Rubin convergence indicator (Brooks and Gelman, 1998) in panels (c) and (d) and the acceptance rate in panels (e) and (f). The first is calculated by running and comparing 10 simulations (chains) for each site and algorithm. More precisely, it measures the ratio of inter-chain to intra-chain variance. An example of a chain of 4 parameters is shown in Fig. 3.10 for Lake Prespa, which will be discussed in more detail later. Since we consider more than one parameter in the MCMCs, the multivariate extension of this type of analysis of variance (ANOVA) is calculated as described in Brooks and Gelman, 1998. The closer this ratio is to 1, the more likely it is that convergence has been achieved. Gelman et al., 2013 recommend a ratio of less than 1.1, which is guaranteed in all simulations. Moreover, both medians are quite close to one (two decimal places are shown), and only a few are above 1.04. The acceptance rate in (e) and (f) is the second MCMC diagnosis and shows higher values on average for the MG. Note that the Gibbs part of this method has an acceptance rate of 1, but the MG part can be checked in the same way as with rwMH. In panel (e) not only the median is higher, but also the variance. The reasons for these differences are explained in the following.

Since Gibbs sampling in our case uses the categorical distribution to determine the taxa weights (see Eq. 2.26), only a finite number of different taxa combinations is possible, unlike rwMH. More precisely, the categorical distribution requires, in addition to the weights, a parameter indicating the number of taxa considered. We choose half the number of taxa because $\binom{K}{t}$ is maximal when $t = K/2$. In this case, K is the number of taxa considered and t is the parameter inserted into the

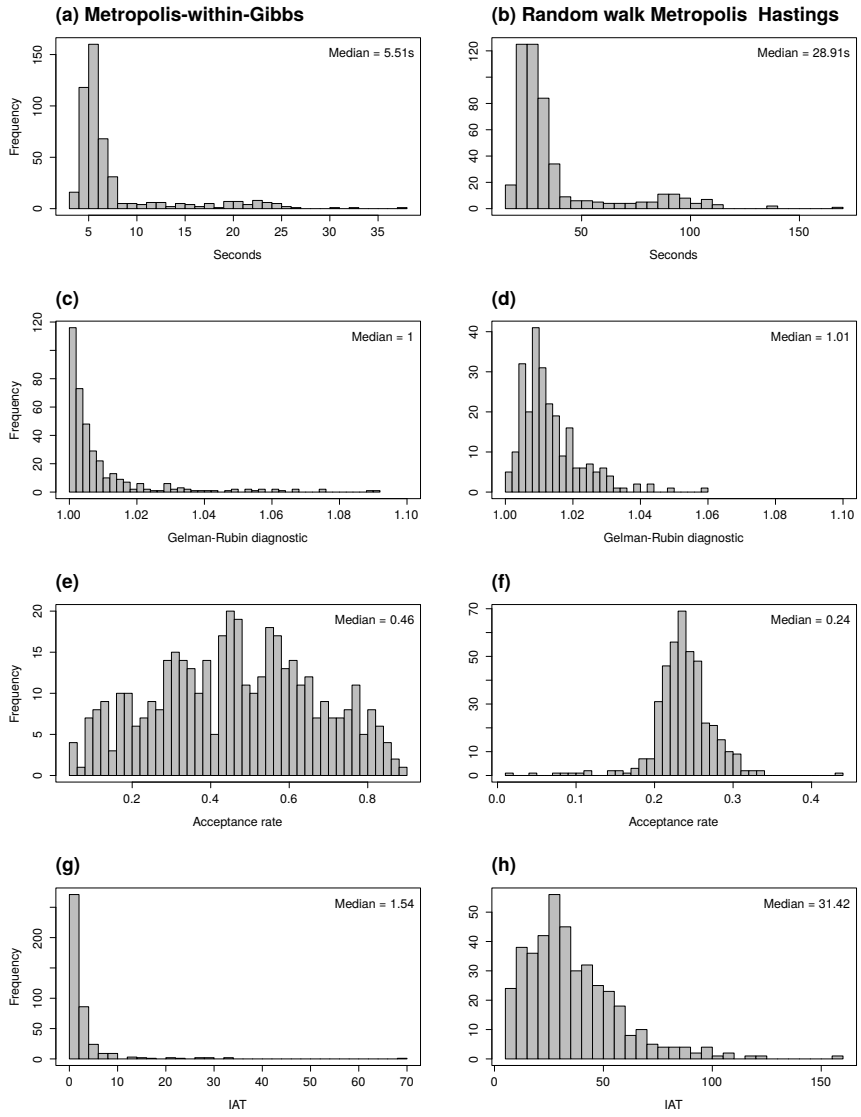


Figure 3.4: The left column shows MCMC results based on Metropolis-within-Gibbs sampling and the right column the results based on random walk Metropolis-Hastings. Panels (a) and (b) depict the computation time of each simulation and (c) and (d) a convergence indicator based on the Gelman-Rubin convergence diagnosis. In addition, panels (e) and (f) reveal the acceptance rate and (g) and (h) the integrated autocorrelation time of each MCMC simulation.

categorical distribution. On the one hand, this allows the algorithm to make larger "jumps" in the parameter space during the inference process, as omitting or adding different taxa can lead to larger changes in the information. The autocorrelation thus decreases faster than with rwMH. On the other hand, since not all taxa information are considered at the same time (as with rwMH), the step size can be relatively small. These extremes could lead to comparatively high and low acceptance rates (Lynch, 2007), which are shown in Fig. 3.4 (e). Note that each accepted reconstruction contains the

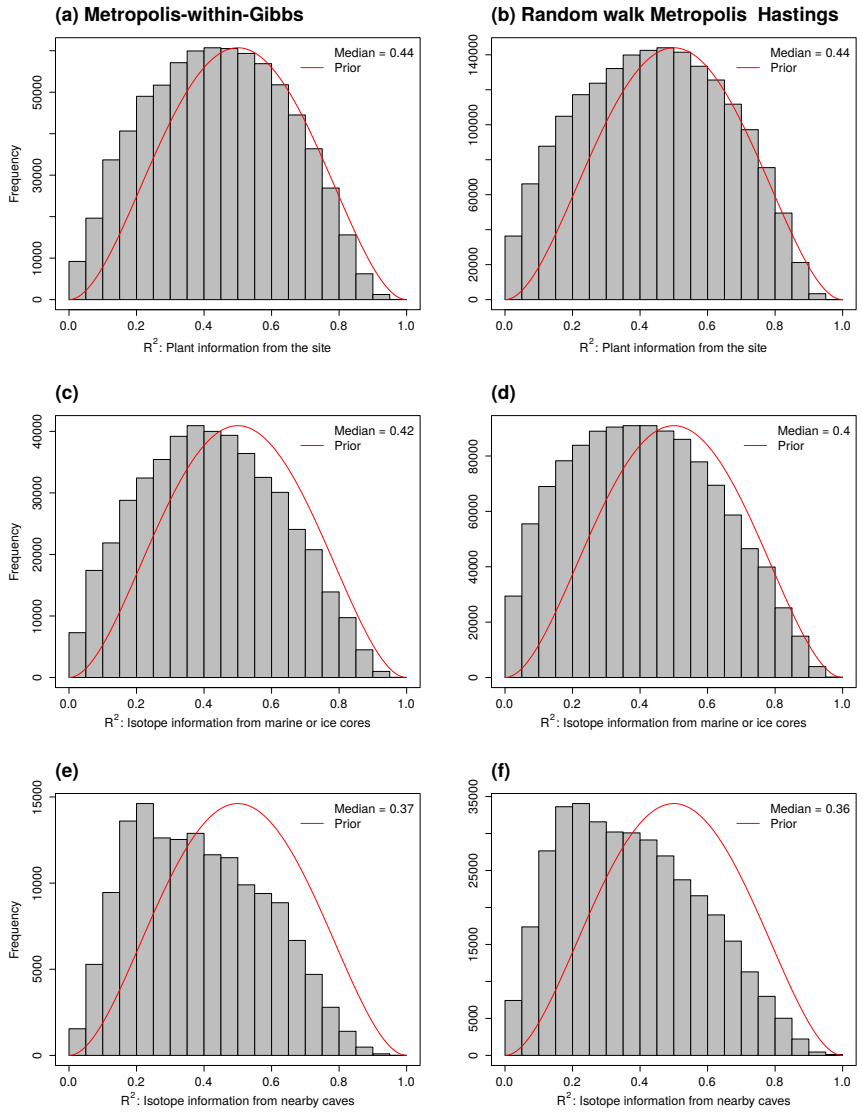


Figure 3.5: The left column shows MCMC results based on Metropolis-within-Gibbs sampling and the right column the results based on random walk Metropolis-Hastings. In summary, all explained variances of the final sample iterations are shown. Panels (a) and (b) depict R^2 with respect to plant information, (c) and (d) regarding isotopic information from marine or ice cores, and (e) and (f) concerning isotopic information from speleothems.

uncertainty information of the transfer functions. Therefore, even a single sample corresponding to our proposal distributions may be sufficient. However, with appropriate MCMC settings, we can still extract additional information from the proxies.

Fig. 3.5 and Fig. 3.6 give us an overview of all samples of all parameters. There are 5 parameters in total, namely the explained variances in terms of plant information, in relation to isotopic information from marine or ice cores, in reference to isotopic information from nearby caves, the annual temperature

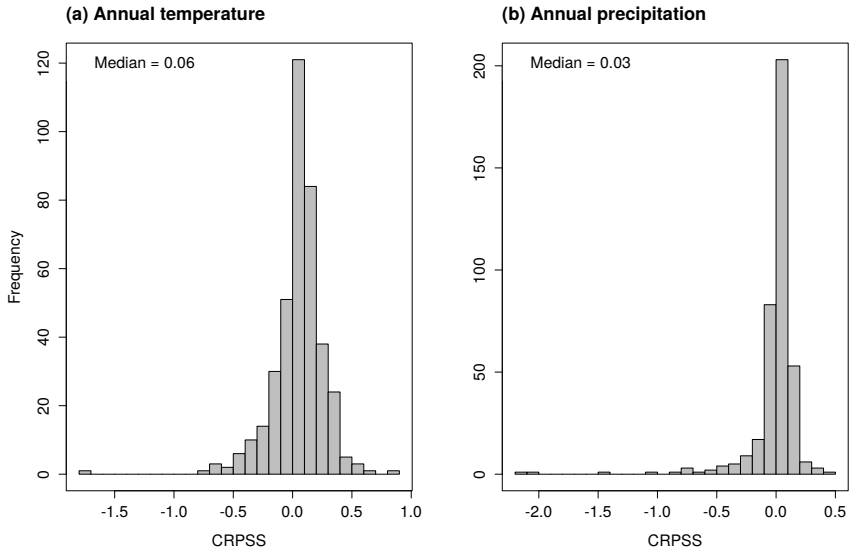


Figure 3.6: Continuous Ranked Probability Skill Score based on reconstructed time periods for ages younger than 1950 AD for annual temperature and precipitation compared to their respective current measurements. The "forecasts" are the results of the Metropolis-within-Gibbs method and the "reference forecasts" are selected from the random walk Metropolis-Hastings samples.

and the annual precipitation of the most recent time layers. Fig. 3.5 shows the explained variances for both methods and all MCMC simulations. The red lines represent the symmetric proposal prior distribution of R^2. Their means are around 0.5 and their variances are close to 0.2. These moments are used to find a compromise between the different locations and not to allow relatively high or low values. For easier comparison, the maximum values of the prior distributions shown correspond to the maximum values of the respective histograms containing the posterior samples.

In summary, the histograms for both methods look the same with respect to each parameter, which is also reflected in the medians. From top to bottom, the skewness of the posterior samples becomes more and more positive, indicating that the respective medians decrease. The explained variances in terms of the plant information are closest to the proposal distribution. This could be due to the fact that the taxa used for the reconstructions and the corresponding reference curves are based on the same sediment information (Giesecke et al., 2017). With regard to the reference curves from marine or ice cores, the values are somewhat lower. We see that the available taxa information for each core have sufficient variability so that these reference lines can be approximated in the manner presented. The lowest mean explained variance is found in panels (e) and (f), where pattern correlation of wavelet power spectra plays a role. Nevertheless, the medians indicate that the cores contain enough information to mimic the spectral patterns of nearby caves.

In Fig. 3.6 we see a comparison of the two methods in terms of annual precipitation and annual temperature. This is done by calculating the Continuous Ranked Probability Skill Score (CRPSS).

Specifically, the CDFs of the two climate variables from periods younger than 1950 AD are compared with the corresponding CRU data. Therefore, the Continuous Ranked Probability Score (CRPS) is calculated first, as described in Hersbach, 2000. To compare the two methods, we define the CRPS from MG as "forecast" and the CRPS from rwMH as "reference forecast". The CRPSS is then calculated via

$$\mathrm{CRPSS} = 1 - \frac{\mathrm{CPRS}_{\text{forecast}}}{\mathrm{CPRS}_{\text{reference forecast}}}. \tag{3.1}$$

This means that values close to 0 indicate that both methods have equal skill. We see that the medians of the two climate variables show that MG and rwMH have relatively similar posterior results. Except for some outliers, the histograms show distributions centered around 0.

All in all, both methods lead to similar results in parameter space. There are some dissimilarities in typical MCMC indicators due to differences in the inference of taxa weights. The MG has the advantage of converging much faster, but the mixing rate is sometimes a bit lower. Nevertheless, both methods lead to similar final quantitative reconstructions. An example of this is described in the next section.

3.2.2 Lake Prespa

We will now use the example of Lake Prespa to show how we determine a final reconstruction. Lake Prespa is located on the border between Greece, Albania and Northern Macedonia at an altitude of 849 m above mean sea level (amsl) (Aufgebauer et al., 2012). It is one of the oldest lakes in Europe and one of the three largest in the Balkan Peninsula. The lake has no surface runoff and drains into Lake Ohrid via karst aquifers. A quantitative paleoclimate reconstruction of Lake Ohrid is presented in Sect. 3.3.6. In Fig. 3.7 a climate diagram of the CRU grid point closest to Lake Prespa shows a maritime climate Cfb. Although drier climatic conditions prevail during the warm summer months, the Mediterranean influence is relatively low due to the orography and there is no year-round dry season. The mean annual temperature in this region is about 8.8 °C, and the average annual precipitation is around 705 mm. In this work, we use all CRU grid points of the Cfb climate classification to calculate the climate PDFs. In this way, the catchment area of Lake Prespa should be better represented than if only the nearest grid point is considered. This results in higher variance for both climate variables and provides our MCMC algorithms with higher flexibility in anchoring recent time slices within these distributions. Due to the orography, the climate is highly variable influencing the vegetation surrounding the lake, which act like an vegetation refuge during Quaternary climate cycles (Panagiotopoulos et al., 2020). Panagiotopoulos et al., 2013 describe the current vegetation as a mixture of Central European, Mediterranean, and Balkan endemic plants.

In autumn 2009, core Co1215 was drilled in the central northern part of the lake. The vegetation of the uppermost 320 cm, i.e., the last ca. 18,000 years, are examined in Panagiotopoulos et al., 2013. Later, additional samples were added to this palynological dataset to increase temporal resolution

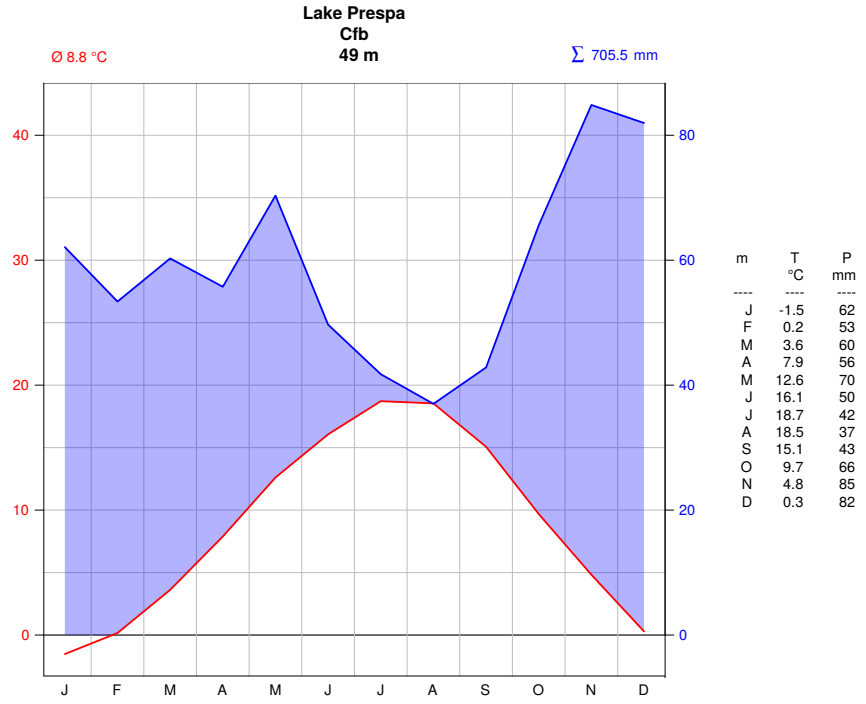


Figure 3.7: Climate diagram of the Lake Prespa region. The data are based on CRU 4.01 (Harris and Jones, 2017) and represent the grid point which is closest to the Lake Prespa over the period from 1961 – 1990 AD. The red colors and numbers represent the temperatures in °C and the blue ones the precipitation in mm. The letters on the x-axis indicate the months, which are also shown on the right side of the diagram, where the climate data are listed in a table.

(Konstantinos Panagiotopoulos, personal communucation). In Fig. 3.8 we see the two reference curves used for the climate reconstruction below. Panel (a) shows isotopic information from multiple cores in the Mediterranean Sea, referred to as Medstack (Colleoni et al., 2012). In (b) we see the percentage of arboreal pollen in relation to the percentage of non-arboreal pollen, based on the palynological data of the drill core. We correlate AP/NAP and Medstack respectively with the reconstructed annual temperature. This is based, on the one hand, on the discussion in Wagner et al., 2019, where Medstack is associated with warm/cold SST. On the other hand, Panagiotopoulos et al., 2013 link AP/NAP with warmer/cooler temperatures, respectively. As input to our Bayesian age-depth model, we use the tephra layers and radiocarbon ages described in Aufgebauer et al., 2012. The corresponding result is shown in Appendix A.

Let us first consider the acceptance rate of the two methods shown in Fig. 3.9 (a) and (b), which are generated during both MCMC simulations of Lake Prespa. The x-axes show the different number of iterations. We choose a burn-in size of 2,500 for MG and 25,000 samples for rwMH. Both curves show that the MCMC simulations seem to converge after these iterations. To account for autocorrelation (AC), every 5th iteration of the remaining samples is considered. Panels (c) and (d) contain the ACs of the final sample sizes of 2,000 and 5,000. This type of thinning is a good choice for MG because

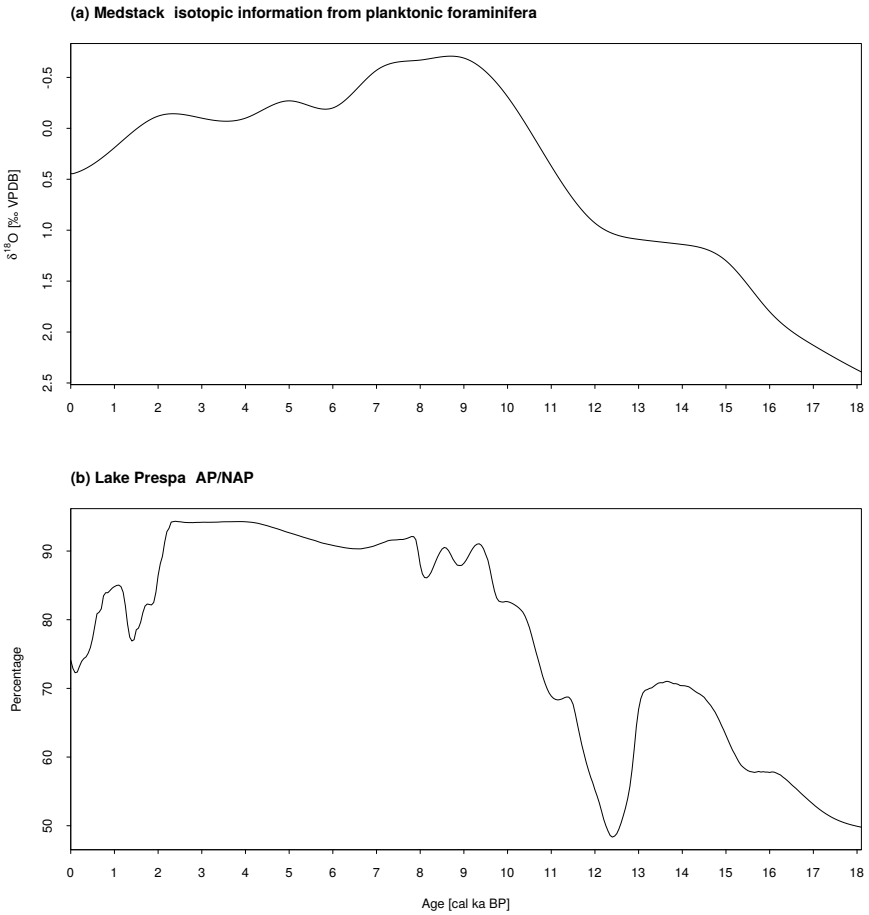


Figure 3.8: Reference curves based on (a) isotopic information from planktonic foraminifera from stacked marine cores of the Mediterranean Sea (Colleoni et al., 2012), and (b) on total terrestrial pollen sums: arboreal and non-arboreal pollen taxa from Lake Prespa (Panagiotopoulos et al., 2013).

the corresponding AC in (c) drops immediately after the first lag. In contrast, the rwMH method requires a larger thinning step. As already explained, the IAT in Fig. 3.4 (g) and (h) gives us an idea of such step sizes. Therefore, this indicator is chosen not only for further trimming of chains, but also to obtain independent posterior samples. First, however, we will look at examples based on thinning without IAT.

In Fig. 3.10 the thinned chains of the MG method are shown for 4 parameters. As SISAL does not offer a cave near Lake Prespa, this parameter is not included. In addition, panel (d) is based on isotope data from the Mediterranean Sea (Medstack). At this point it should be mentioned again that the flexible algorithm gives the users the possibility to include reference curves of their choice. On the right side we see the corresponding prior distribution and the posterior densities defined by each chain. For comparison, the posterior distribution of rwMH is also shown in the same diagrams, with all PDFs normalized to the same range. As far as the annual temperature in panel (a) is concerned,

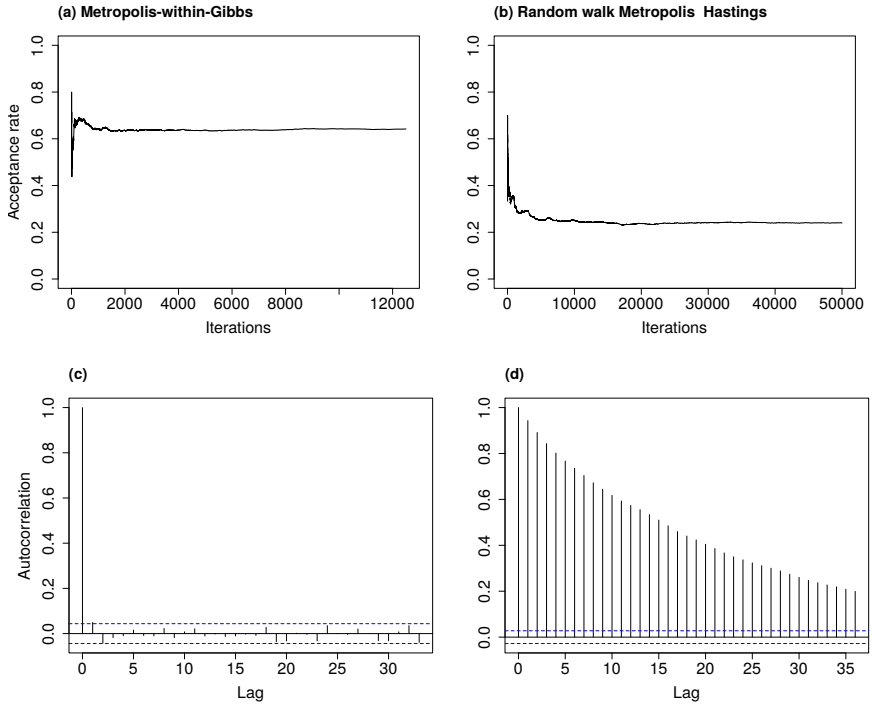


Figure 3.9: The left column shows the autocorrelation and acceptance rate of the Metropolis-within-Gibbs method for Prespa Lake. The right column depicts the same, but for the random walk Metropolis-Hastings method.

both posterior distributions look relatively similar. They have a higher mean and lower variance than the prior distribution. In (b) the output converges close to the mode of the prior gamma distributed annual precipitation. Here we see a bimodality in the MG method, indicating that the algorithm "jumps" between two convergence zones. However, the rwMH converges within the mode with higher precipitation. The explained variances with respect to the arboreal percentages from the core shown in (c) exhibit similar convergence behaviour for both methods, resulting in slightly higher posterior values. In comparison, the rwMH in (d) has a higher mean value than the MG. Both reference curves show a relatively high correlation in the case of Lake Prespa. Nevertheless, the additional computational cost is quite small and the final result contains, although not very much, more information than if only one reference curve is considered.

Even though there are slight differences in the results just explained, the overall reconstruction of the two methods is relatively similar. This is shown in Fig. 3.11. Depicted are the PDFs of the mean values derived by the MCMC parameters described above for both methods. The only obvious difference can be seen for ages older than 12 cal ka BP in panel (c), where there are two branches of annual precipitation. This is also reflected in the bimodal posterior distribution of MG in Fig. 3.10 (b). The comparison between MG and rwMH shows that the mode with higher annual precipitation for

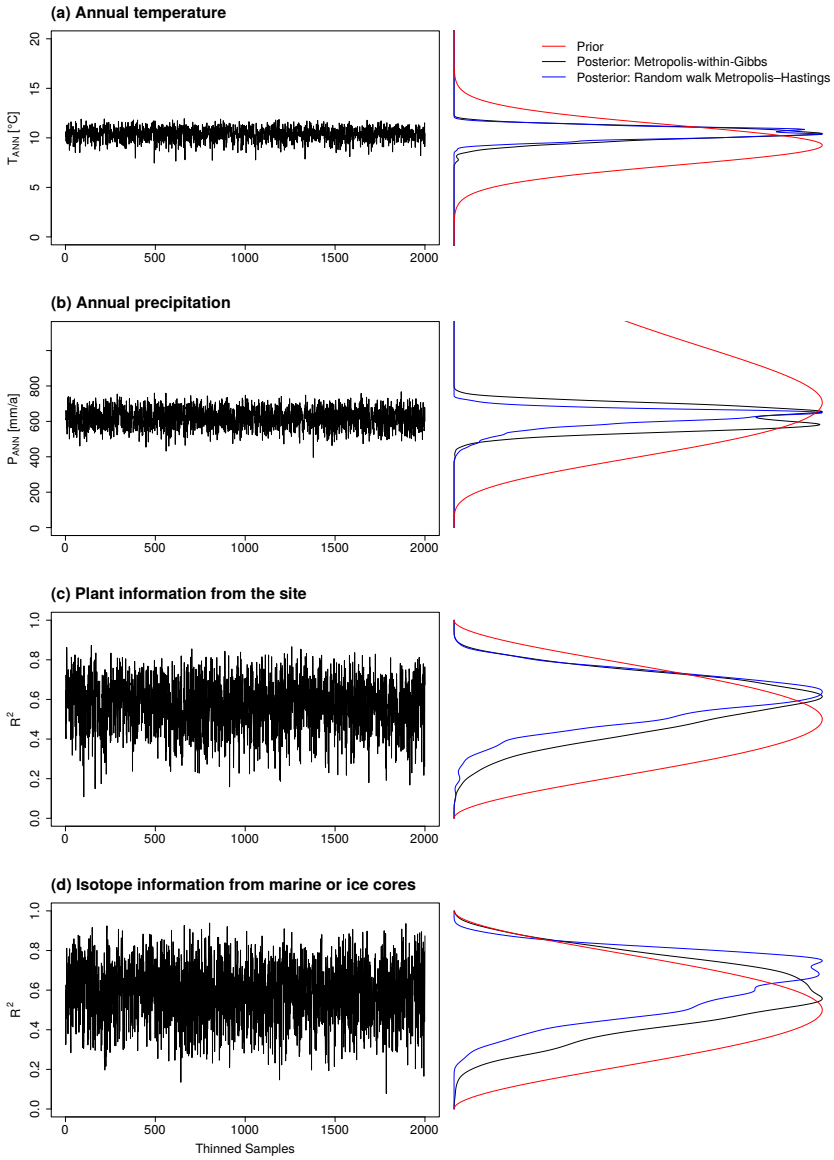


Figure 3.10: Trace plot and corresponding PDFs of the MCMC simulation of Lake Prespa. The left column shows the thinned samples from the Metropolis-within-Gibbs method for both climate variables and the explained variances with respect to the plant and isotope information. The right column depicts the PDFs based on these samples from both MCMC methods, distinguished by black and blue lines. The red lines show the respective prior distribution for each parameter.

the most recent time slice leads to the branch with lower precipitation in the late Last Glacial (not shown). Ultimately, however, this difference does not affect the common trade-off between the other parameters, as panels (a) and (b) do not show large discrepancies.

So far, we have only examined the mean values of the reconstructions. However, for a final climate reconstruction, we want to include the uncertainty information that results from the probabilistic

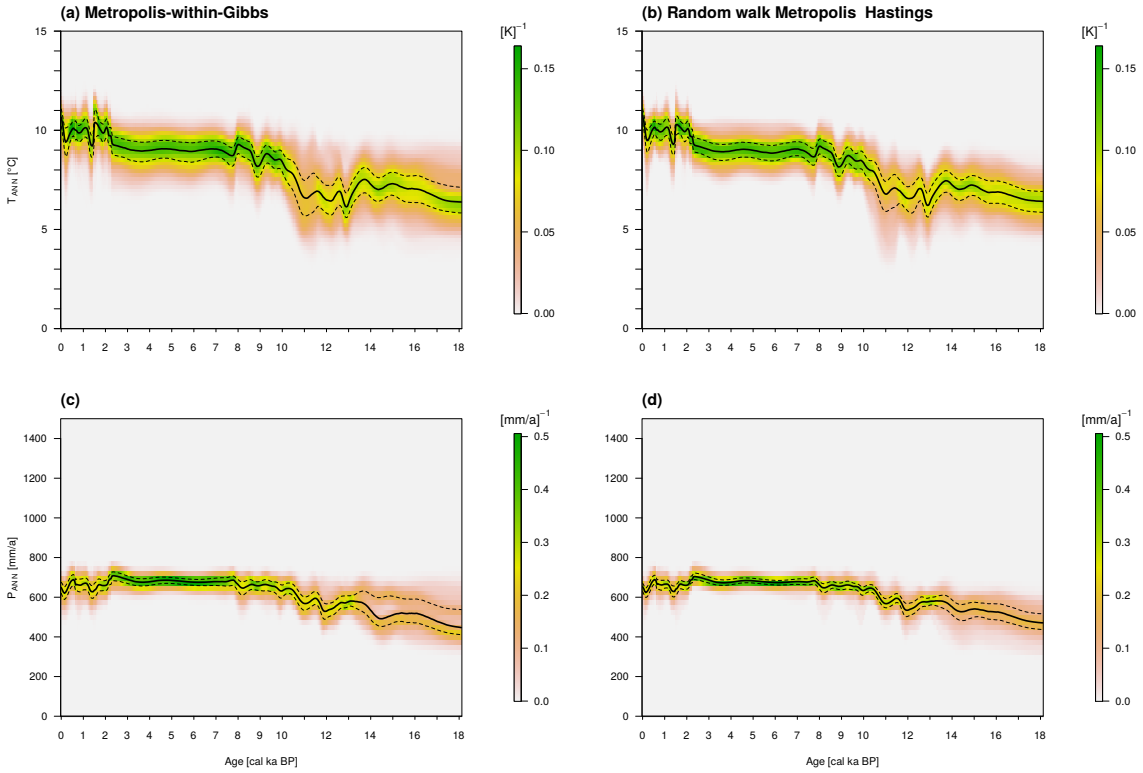


Figure 3.11: Posterior PDFs of all medians of the two climate variables (first and second row) and for both methods (first and second column).

transfer functions. There are many ways to do this, and we will briefly introduce some of them below and finally present the most appropriate one for our concern.

As a first step, the chains in Fig. 3.10 are first pruned by selecting those samples that lie within the interquartile ranges of the individual prior distributions. In a second step, trimming is continued until the IAT is below 1.3. Finally, a Pareto optimum is used to select the samples closest to the corresponding prior means. Here the user can define how many samples are to be selected at the end. Initial tests show that 10 individual reconstructions contain enough variability to describe large parts of the MCMC simulation. For this reason, 10 samples are always considered for each quantitative climate reconstruction in the following discussions.

The Figs. 3.12, 3.13, and 3.14 show different ways in which these 10 final samples can be combined. Note that each sampled reconstruction contains the uncertainty information from the transfer functions. So we have to deal with the combination of PDFs. Each of the following calculations is performed separately for the two climate components of the vector $\boldsymbol{C}$ and is also true for the corresponding random vectors $\mathcal{C}_i$. For simplicity, the two expressions are not presented decomposed into their respective components. Fig. 3.12 shows the final mixture PDF $\mathbb{P}_{\text{mixt}}(\boldsymbol{C})$. These weighted averages are

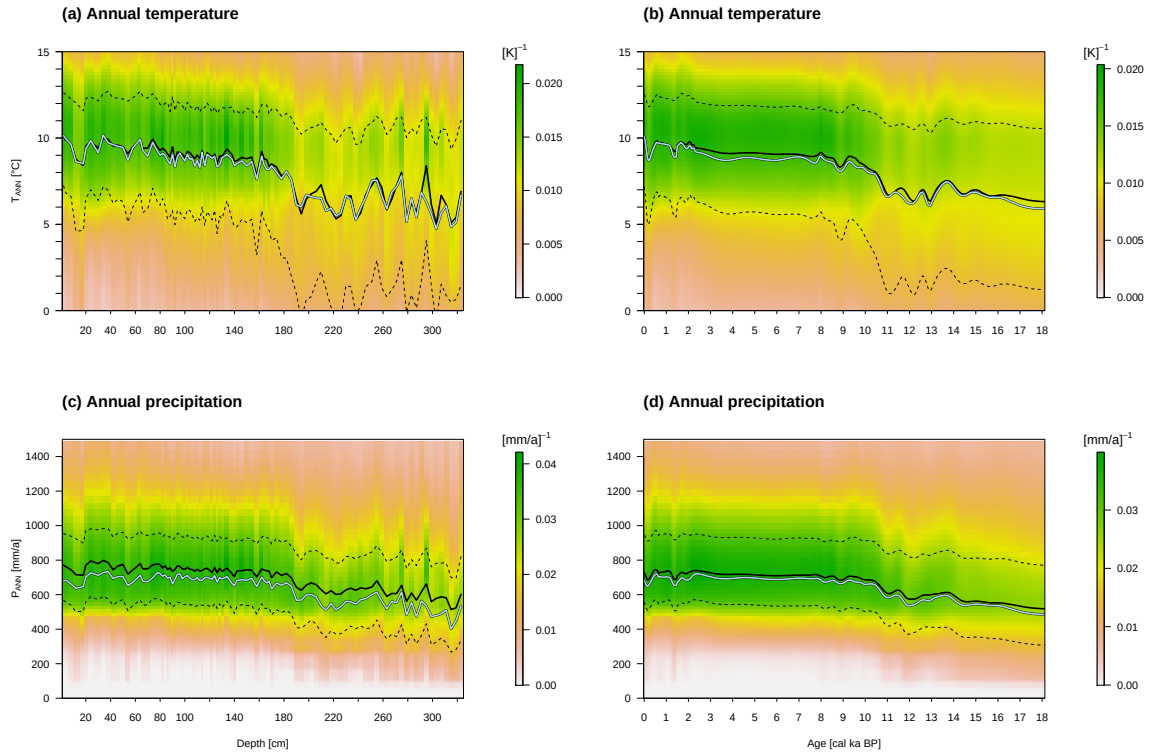


Figure 3.12: Posterior PDFs of the two climate variables (first and second row) and depth (first column) and age (second column). The basis of this result is the weighted average of all posterior PDFs. The black lines are the medians of the colored probability density values calculated using the Metropolis-within-Gibbs results. The blue lines are based on the random walk Metropolis-Hastings method.

calculated as follows:

$$\mathbb{P}_{\text{mixt}}(\boldsymbol{C}) = \frac{1}{N}\sum_{i=1}^{N} w_i \cdot \mathbb{P}_{\mathfrak{e}_i}(\boldsymbol{C}); \quad \text{with: } w_i > 0 \text{ and } \sum_{i=1}^{N} w_i = 1. \tag{3.2}$$

$N = 10$ is the number of independent climate reconstructions and w the corresponding weights. Here, each user can decide for themselves which weights they want to pass on to the algorithm. These could be defined, for example, by the lowest mean variances of the reconstructions or the highest mean probability values with respect to the prior PDFs. In this study, we use the same weights for each PDF. To demonstrate the age-depth transformation in relation to these PDFs, we plot the depths in the left column. The advantage of Eq. 3.2 is the low computational cost, that the input PDFs do not have to be independent and that the mean of $\mathbb{P}_{\text{mixt}}(\boldsymbol{C})$ (black and blue solid lines) is the average of the means of the input distributions $\mathbb{P}_{\mathfrak{e}_i}(\boldsymbol{C})$ (Hill, 2008). This method thus transfers the information

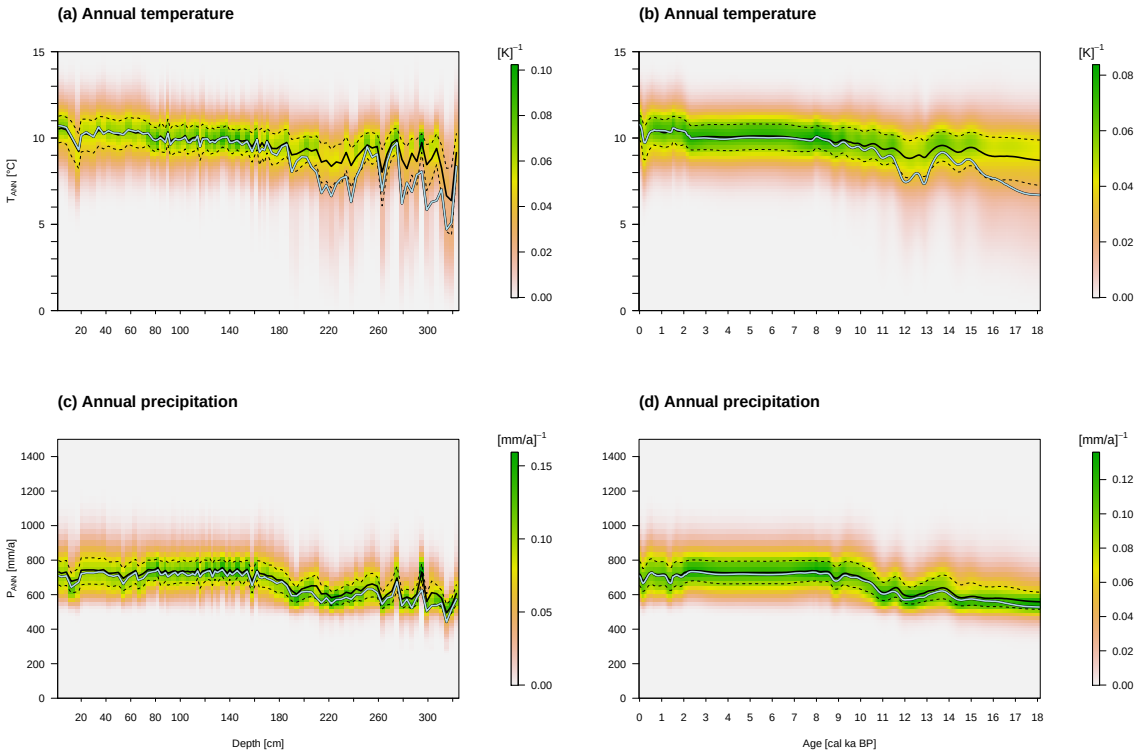


Figure 3.13: As Fig. 3.12, but the basis is the normalized product of independent posterior samples.

obtained during the MCMC simulation into the final result. Moreover, we can see the great similarity of both methods in terms of such a final result. A disadvantage of this combination is that the variance of $\mathbb{P}_{\text{mixt}}(\boldsymbol{C})$ is at least as large as the minimum of the variances of the input distributions $\mathbb{P}_{\mathfrak{C}_i}(\boldsymbol{C})$ (Hill, 2008). This leads to the high variances in Fig. 3.12, which make reasonable interpretation of paleoclimate difficult.

The next method remedies this effect, and the final result shown in Fig. 3.13 has a much lower uncertainty. This is achieved by calculating the normalized product of the individual input PDFs and is called conflation (Hill, 2008):

$$\mathbb{P}_{\text{confl}}(\boldsymbol{C}) = \frac{\prod_{i=1}^{N} \mathbb{P}_{\mathfrak{C}_i}(\boldsymbol{C})}{\int_{-\infty}^{\infty} \prod_{i=1}^{N} \mathbb{P}_{\mathfrak{C}_i}(\boldsymbol{C})\, d\boldsymbol{C}}.$$

At first glance, we see in panel (b) a big distinction between the black and the blue lines in the Last Glacial period. This is due to the fact that the conflation takes into account the respective accuracies of each $\mathbb{P}_{\mathfrak{C}_i}(\boldsymbol{C})$ by ultimately considering their overlapping regions (Hill, 2008). Therefore, this method is relatively sensitive to only comparatively small changes in the respective input PDFs, which has a corresponding effect on the final mean value.

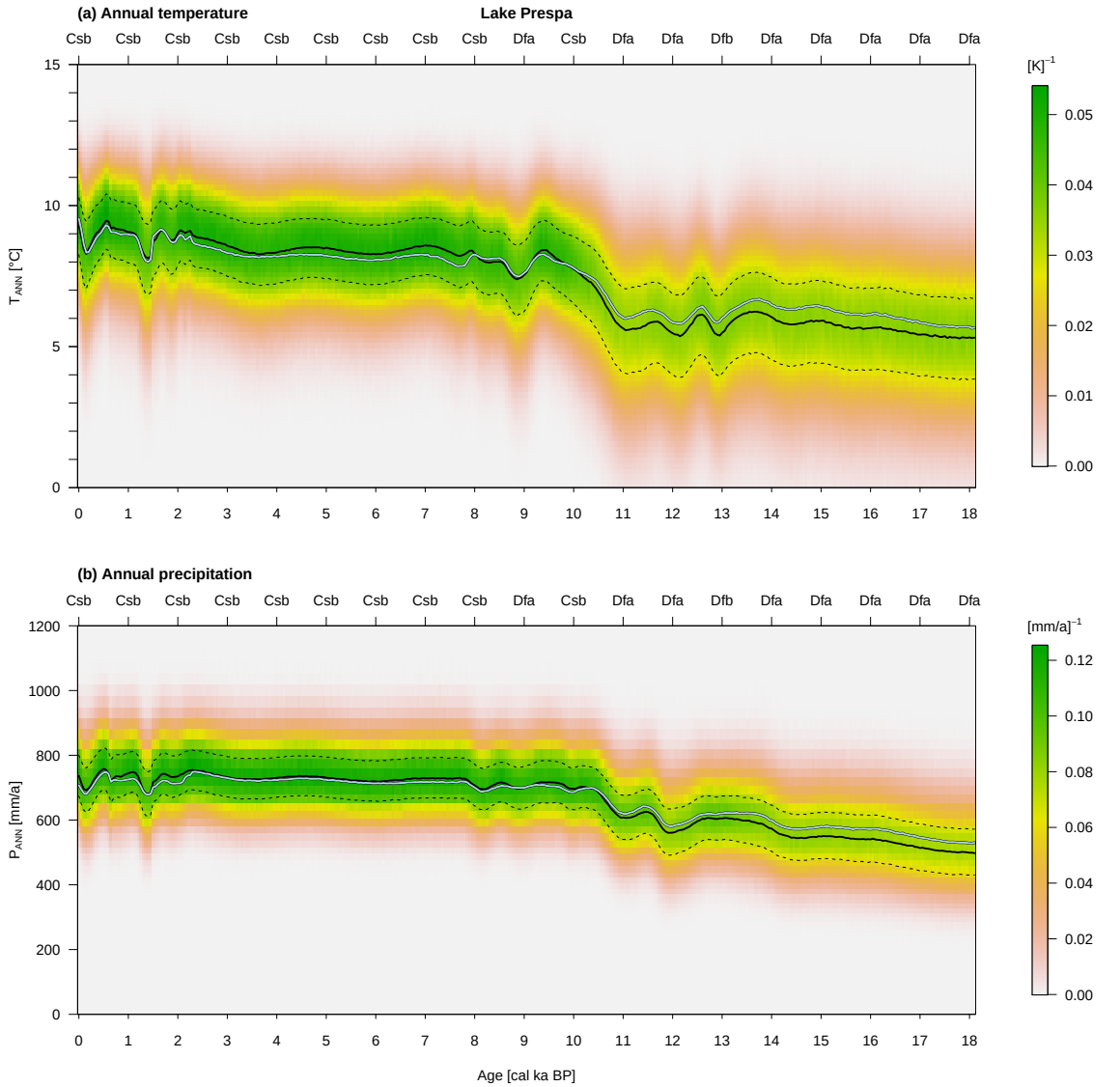


Figure 3.14: Posterior PDFs of annual temperature in (a) and annual precipitation in (b), based on the average of independent posterior random variables. In addition, the axis above the graph shows the approximated Koeppen-Geiger classifications corresponding to the PDFs for the respective periods.

The next method excludes the disadvantages mentioned above. That is, it reduces the variances due to the additional information of independent reconstructions and has the same average mean as the input distributions (Hill, 2008). A respective result is shown in Fig. 3.14 and is calculated via the mean of the input random variables: $Z_N = \frac{\mathfrak{C}_1+\mathfrak{C}_2+\ldots+\mathfrak{C}_N}{N} = \tilde{\mathfrak{C}}_1 + \tilde{\mathfrak{C}}_2 + \ldots + \tilde{\mathfrak{C}}_N$. Here the constant $1/N^{\frac{1}{\alpha}}$ is chosen with $\alpha = 1$. The latter is called the characteristic exponent of the stable density (Honerkamp, 2002). As a result, $1/N$ compresses the limit distribution around a point. This means that for larger

values of N a realization of Z_N scatters less and less around the mean of Z_N (Honerkamp, 2002), which explains the shape of the PDFs in Fig. 3.14. To calculate them, we use the fact that the PDF of the sum of the random variables can be obtained by calculating the convolution of the respective PDFs (cf. Eq. 2.2): $\mathbb{P}_{Z_N}(\boldsymbol{C}) = \mathbb{P}_{\tilde{\mathfrak{C}}_1}(\boldsymbol{C}) * \mathbb{P}_{\tilde{\mathfrak{C}}_2}(\boldsymbol{C}) * ... * \mathbb{P}_{\tilde{\mathfrak{C}}_N}(\boldsymbol{C})$. Although more time-consuming to compute, all subsequent local climate reconstructions are based on this combination method because it has exactly the advantageous properties we want for such reconstructions.

At first glance, it can be seen that the course of the annual temperature in Fig. 3.14 (a) is similar to the reference curves in Fig. 3.8. This also affects the reconstructed annual precipitation, which is linked to T_{ANN} via the transfer functions and shows a similar pattern. It is also noticeable that both reconstructions (blue and black solid lines) give a very similar final result. We obtain the Koeppen-Geiger information by comparing the PDFs of the two climate variables with those of all Koeppen-Geiger classes. For this purpose, the so-called Earth Mover's Distance (EMD) is used, which takes into account the entire uncertainty structure (Rubner et al., 2000). The Koeppen-Geiger class with the lowest EMD is then used as the most likely. It should be noted that this estimate does not include direct intra-annual climate information. Nevertheless, we will see that this method gives usable results. In the following, the reconstruction in Fig. 3.14 is explained in more detail and compared with the corresponding palynological investigations by Panagiotopoulos et al., 2013 and Aufgebauer et al., 2012. For this purpose, we also predominantly use their temporal subdivision of vegetation history.

During the Late Pleniglacial, which in our case roughly corresponds to the period from **18 – 15 cal ka BP**, Panagiotopoulos et al., 2013 describe a relatively high AP (cf. Fig. 3.8 (b)) with low pollen influx. They indicate sparse vegetation, which Aufgebauer et al., 2012 describe with cold-tolerant herbs and pinus typical of stadial conditions. On average, our reconstructed variables show a temperature of about 6 °C and a precipitation of 600 mm/a, leading to the approximated Koeppen-Geiger climate Dfa, which describes a continental climate with hot summers without dry periods. The reconstructed values are similar to those in present-day areas in Russia between the Black Sea and Moscow, where the natural vegetation is described in Bohn et al., 2007 as a transition from forest steppes to herb-rich and pure feather grass steppes.

The period from **15 – 13.2 cal ka BP** encompasses the climatic fluctuations known as Bølling-Allerød, for which Panagiotopoulos et al., 2013 and Aufgebauer et al., 2012 describe a spread of tree taxa and a decline of herbaceous steppe taxa. From this they infer an increase in temperature and moisture availability. These fluctuations are also present in the two reconstructed variables. In the process, the mean temperature rises to 7 °C and precipitation by 50 mm/a. However, the Koeppen-Geiger class remains the same, indicating a relatively broad climate spectrum in Dfa.

During the Younger Dryas from ca. **13.2 – 11 cal ka BP**, vegetation reverts to near-glacial conditions, indicating a drier and colder climate (Panagiotopoulos et al., 2013; Aufgebauer et al., 2012). This is also reflected in the climate reconstruction, where temperatures are partly colder than in the lowest part of this core. Accordingly, the approximated Koeppen-Geiger classification changes to Dfb, which

describes a continental climate with warm summers without dry periods.

The early Holocene from ca. **11 – 7.9 cal ka BP** is described by Panagiotopoulos et al., 2013 and Aufgebauer et al., 2012 with an increase in vegetation density and the formation of closed oak forests. From this they deduce a steady increase in temperature and precipitation, which is reflected in our reconstruction. Here, T_{ANN} rises to around 9 °C and P_{ANN} to ca. 750 mm/a, leading to the climate classification Csb, which characterizes a Mediterranean climate with warm summers. In the period from about 9 – 7 cal ka BP, both variables show larger fluctuations that can be attributed to the 8.2 ka event. This Bond event can be seen well in Fig. 3.8 (b) and is described with a decrease in temperature and precipitation (Panagiotopoulos et al., 2013; Aufgebauer et al., 2012). However, our reconstruction shows an opposite behaviour in terms of temperature. This is because, for example, Ulmus is weighted higher due to the approximations of the reference curves and the climate anchor points. The corresponding transfer function shows a relatively cold and humid climate, which together with a pollen fluctuation in the period around the 8.2 ka event leads to this reconstructed pattern.

In the middle and late Holocene from about **7.9 cal ka BP onwards**, the vegetation shows a rapid spread of several thermophilous trees (Panagiotopoulos et al., 2013) and thus a return to the Mediterranean climate with warm summers (Csb). From 1.9 cal ka BP we see in Fig. 3.8 (b) strong fluctuations in the pollen spectrum and a decrease in AP. This is due to the increasing influence of humans on vegetation (Panagiotopoulos et al., 2013), which is also partly reflected in our reconstruction. Both climate variables, however, do not show a steady decline as the reconstructions of the most recent period are drawn towards the climate anchor points (cf. Fig. 3.7). Furthermore, the a priori fixed parameters of the beta distributions prevent too close approximation to the reference curves. The recent Koeppen-Geiger classification of Lake Prespa does not correspond to what is seen in our reconstruction. Nevertheless, we will see in Sect. 3.3.6 that the neighbouring Lake Ohrid, which has a similar climate, is referred to as Csb. This means that this area is a transition zone between Cfb and Csb, which confirms the method for determining the approximated Koeppen-Geiger climate classification as useful.

With the exception of the period from 9 – 7.9 cal ka BP, our reconstruction agrees with Panagiotopoulos et al., 2013 and Aufgebauer et al., 2012. This becomes particularly clear when looking at the Koeppen-Geiger climate classification of the respective periods. This allows them to be compared well with present-day areas and their vegetation, giving a more comprehensive picture of climate reconstruction.

Since it is quite time-consuming to obtain the above AP/NAP curves for most sites from the EPD, we introduce another reference curve based on the results of Giesecke et al., 2017. They provide reconstructed European maps for 16 taxa based on the EPD datasets. These show the relative occurrence of these plants in relation to different thresholds in the period 0 – 15 cal ka BP. The temporal resolution is 500 years, during which we obtain information on regional occurrence (threshold 1) to regional dominance (threshold 4) of each taxa. Giesecke et al., 2017 finally interpolate these

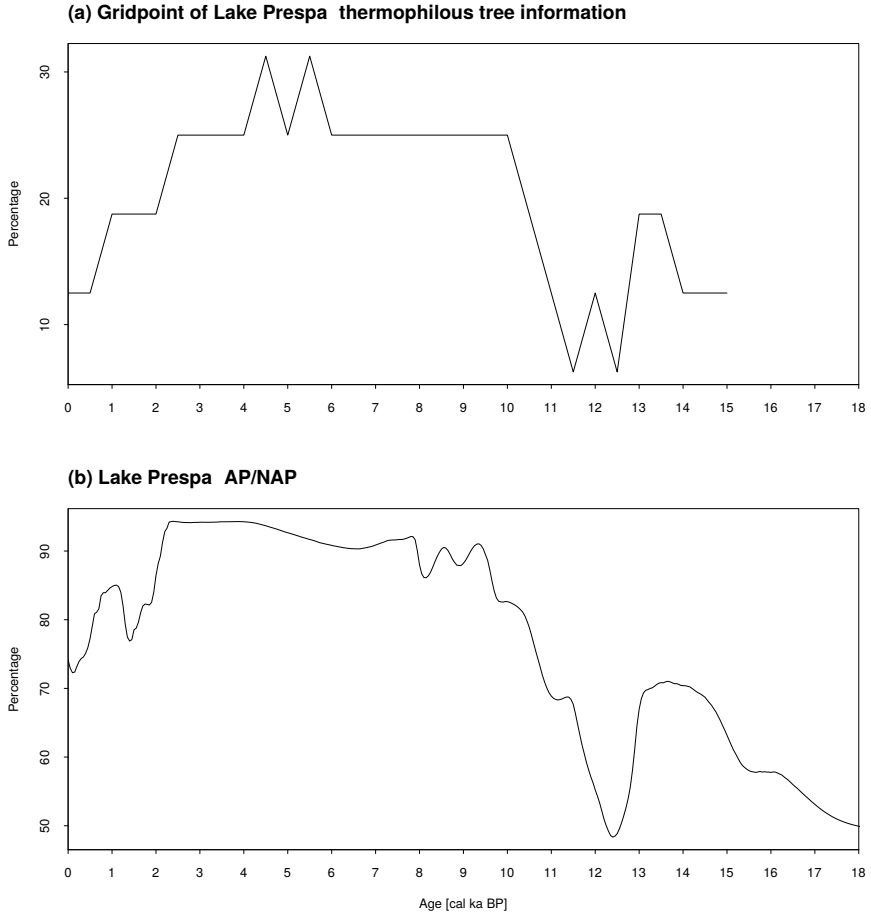


Figure 3.15: Reference curves based on (a) thermophilous tree information from Giesecke et al., 2017 and (b) on total terrestrial pollen sums: arboreal (AP) and non-arboreal (NAP) pollen taxa from Lake Prespa (Panagiotopoulos et al., 2013).

results to a $0.5^\circ \times 0.5^\circ$ grid. In this way, we can retrieve this information for the respective lakes and mires. For Lake Prespa, the grid point closest to the lake is considered and all values above the regional occurrence are taken into account. If only values above threshold 4 were considered, part of the reference curve would be equal to 0, as the regional dominance of these taxa is missing (not shown). We consider 7 different thermophilous trees, which are compared with the other 9 taxa, giving us a relative abundance of these plants that varies according to regional occurrence and regional dominance. The result for Lake Prespa is shown in Fig. 3.15 (a). When comparing with the AP/NAP curve in (b), one immediately notices a great similarity. The main difference is in the respective percentages, temporal resolution, and restriction to the period from $0 - 15$ cal ka BP. The latter has the greatest impact on climate reconstruction, as shown in Fig. 3.16. Here the blue line shows the final reconstruction based on the Giesecke reference curve. Despite the high degree of resemblance between this and the black solid line with respect to both climate variables, P_{ANN} shows

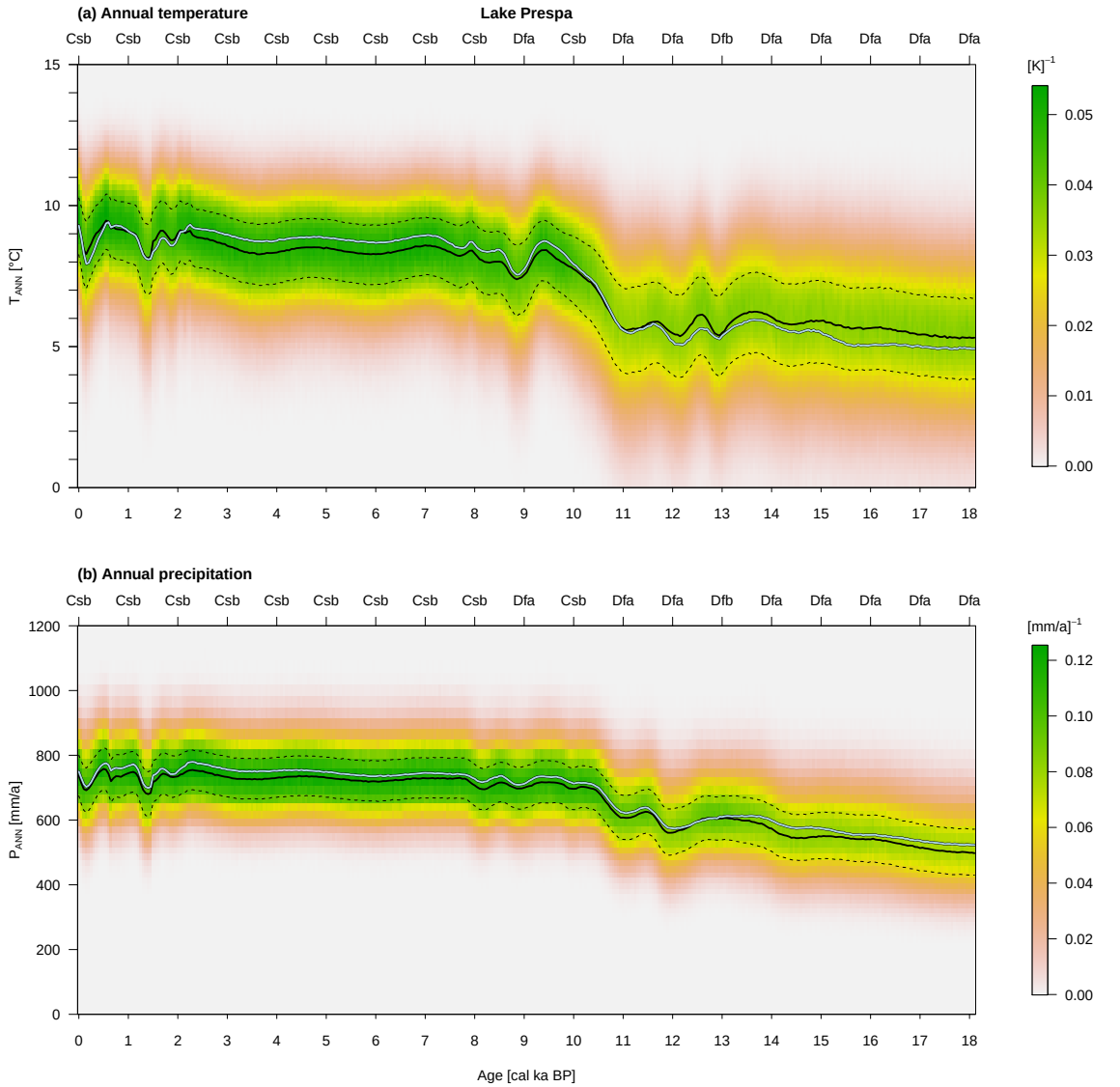
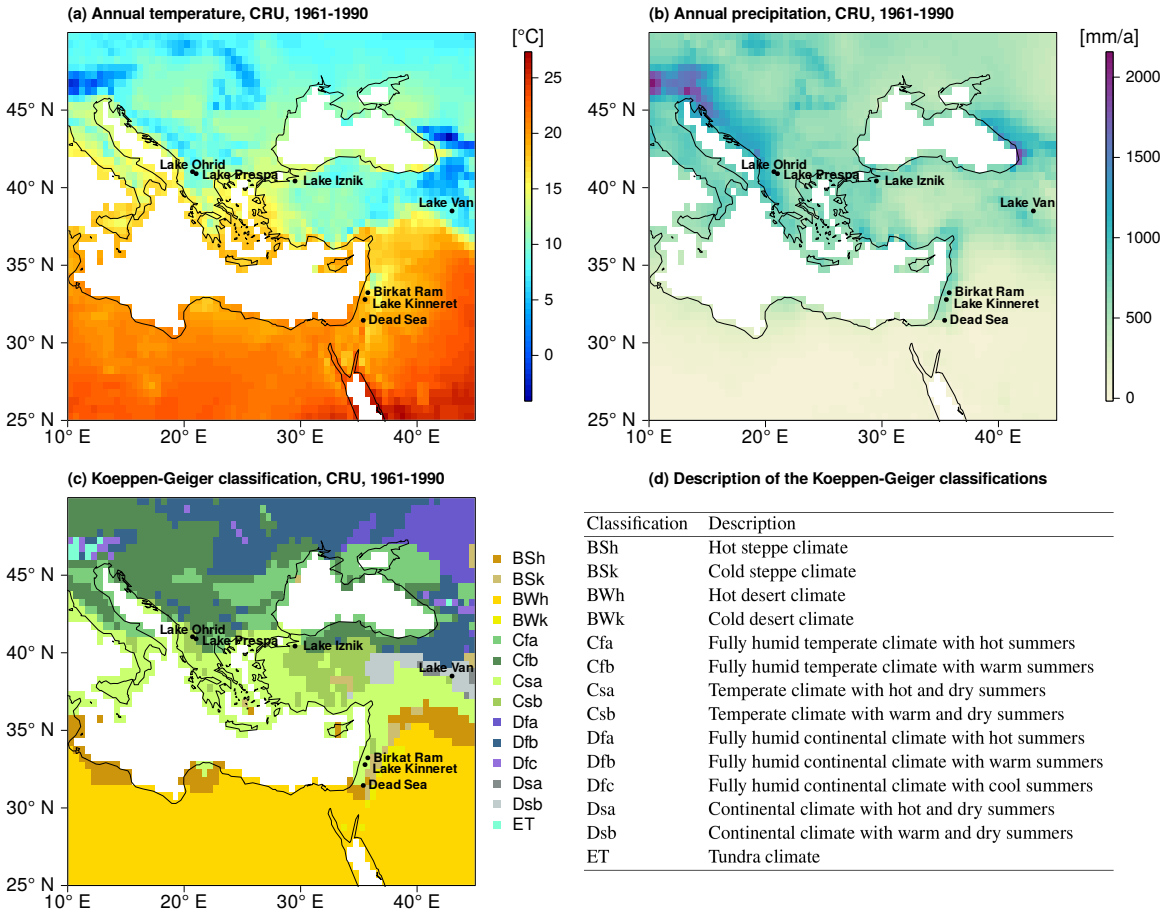


Figure 3.16: As Fig. 3.14, but the blue line shows the final reconstruction based on the Giesecke reference curve.

greater differences in the lower core, which is also observed in T_{ANN}. However, large fluctuations are also observed in (a) during the middle Holocene. These are directly related to those between 15-18 cal ka BP, as the corresponding reconstruction samples for the entire period are generated at once. So, for example, if the blue line in the Late Pleniglacial has a colder temperature, this may affect the entire subsequent course. All in all, Fig. 3.16 shows that a reasonable climate reconstruction can be determined under the aspects just described.

3.3 Local climate reconstructions

In Sect. 3.2 we explain how we compute the final local reconstructions and apply this procedure to the example of Lake Prespa. Here, we will present and discuss other local reconstructions. The respective results of the Bayesian age-depth models performed with Bacon are shown in Appendix A. First, some general information about the regional climate are given. Fig. 3.17 (a) and (b) show annual temperatures and precipitation for parts of southeastern Europe and the Near East. These climate values are from the CRU dataset with the reference period 1961 – 1990 AD. As we have already seen with the Lake Prespa example, we use estimated Koeppen-Geiger climate classifications for additional assessment of the reconstructed climate. The classifications based on our reference period are shown in Fig. 3.17 (c). Using the table in Fig. 3.17 (d), we can see that the sites presented have a wide



Classification	Description
BSh	Hot steppe climate
BSk	Cold steppe climate
BWh	Hot desert climate
BWk	Cold desert climate
Cfa	Fully humid temperate climate with hot summers
Cfb	Fully humid temperate climate with warm summers
Csa	Temperate climate with hot and dry summers
Csb	Temperate climate with warm and dry summers
Dfa	Fully humid continental climate with hot summers
Dfb	Fully humid continental climate with warm summers
Dfc	Fully humid continental climate with cool summers
Dsa	Continental climate with hot and dry summers
Dsb	Continental climate with warm and dry summers
ET	Tundra climate

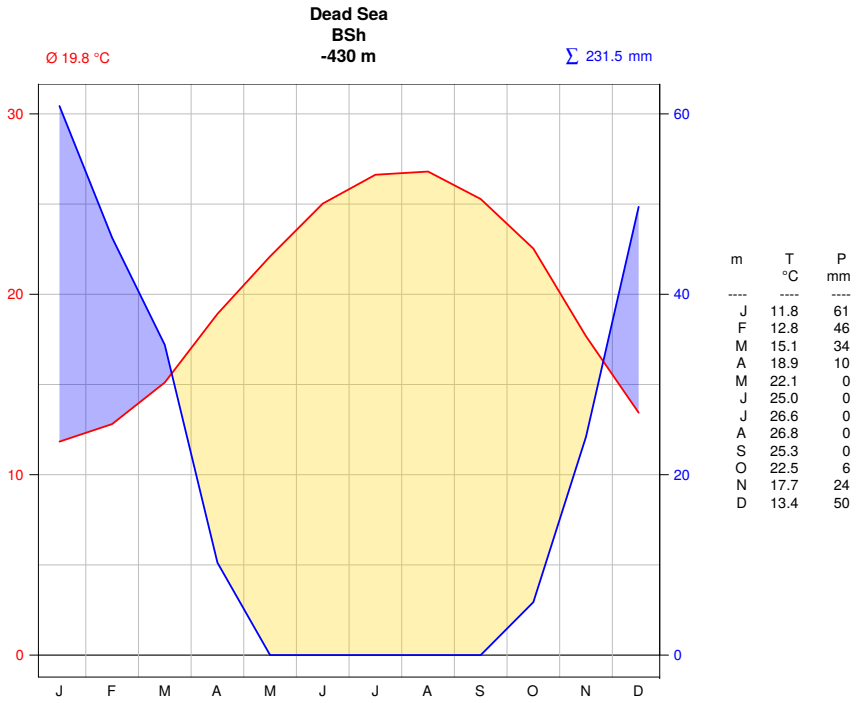
Figure 3.17: Annual temperature in (a) and precipitation in (b) based on CRU 4.01 data (Harris and Jones, 2017) and the 1961 – 1990 AD reference period. The black dots mark the locations for which the reconstructed paleoclimate is discussed in detail in this work. Panel (c) shows the corresponding Koeppen-Geiger climate classifications and the table in (d) describes them based on Chen and Chen, 2013.

variety of climate classes. This ranges from the steppe climate in the Levant to the continental climate in Eastern Anatolia to the Mediterranean and maritime climate in the Balkans. The corresponding explanations can be found in the sections listed below. These are arranged counterclockwise compared to the locations shown in Fig. 3.17.

The descriptions of the following local reconstructions always follow a similar structure. This should make it easier for the reader if they are only interested in specific reconstruction sites. In addition, the results of each site are summarized at the end of the corresponding discussion. For this reason, no additional summary is provided at the end of this Sect. 3.3.

3.3.1 Dead Sea

The Dead Sea is located at the deepest point of the continents, about 430 m below mean sea level (bmsl). This region lies on the Dead Sea Transform, a tectonic boundary between the Sinai Plate and the Arabian Plate (Miebach et al., 2019). It is a terminal lake and is fed primarily by the Jordan River, which flows year-round, but also by groundwater and several streams. The Dead Sea basin has been continuously subsiding since its formation in the early Miocene and has served as an important



m	T °C	P mm
J	11.8	61
F	12.8	46
M	15.1	34
A	18.9	10
M	22.1	0
J	25.0	0
J	26.6	0
A	26.8	0
S	25.3	0
O	22.5	6
N	17.7	24
D	13.4	50

Figure 3.18: Climate diagram of the Dead Sea region. The data are based on CRU 4.01 (Harris and Jones, 2017) and represent the grid point which is closest to the Dead Sea over the period from 1961 – 1990 AD. The red colors and numbers represent the temperatures in °C and the blue ones the precipitation in mm. The letters on the x-axis indicate the months, which are also shown on the right side of the diagram, where the climate data are listed in a table.

sediment trap. The climate diagram in Fig. 3.18 of the grid point closest to the Dead Sea shows a hot steppe climate BSh. The mean annual temperature in the Dead Sea region is approximately 20 °C, and the mean annual precipitation is around 230 mm. Summers are dry and hot, while winters have a mild and humid climate. Most of the rainfall in this region occurs between October and April. This hyper-dry climate leads to a predominance of Saharo-Arabian desert vegetation (Miebach et al., 2019). The unique orography in this region leads to a strong precipitation and temperature gradient. Therefore, the surrounding mountains are dominated by Irano-Turanian steppe and Mediterranean woodland (Miebach et al., 2019). Overall, this site is characterized not only by a transition zone of different climatic classes (cf. Fig. 3.17 (c)), but also by diverse vegetation due to the size of the Dead Sea catchment area (Litt et al., 2012).

The precursor of the Dead Sea, Lake Lisan, existed from 70 to 15 cal ka BP (Torfstein et al., 2013) or even 15 – 20 cal ka earlier (Neugebauer et al., 2016). In the following sections, we present two climate reconstructions for this region. The focus is firstly on the Dead Sea during the Holocene and secondly on the Dead Sea and Lake Lisan during the period from the Penultimate to the Last Glacial.

Holocene

Introduction In 1997 AD, a 21 m sediment core was recovered during a drilling campaign on the west coast of the Dead Sea near En Gedi (Schölzel, 2006). This contains the palynological data for the last 10,000 years, which serves as one climate proxy for the reconstruction described below. Although the top layer studied reflects the climate circa 1955 AD, we suggest that the range of the Koeppen-Geiger classification BSh for the hot steppe takes this assumption into account. We also assume that the choice of such a climate classification better reflects the catchment area of the taxa source in question than just one grid point. Therefore, this climate information is used to anchor the recent time slice in the reconstruction.

A problem arises when examining this core. It largely reflects human activity, with various cultures influencing the vegetation in the Dead Sea basin. The AP/NAP reference curve in Fig. 3.19 (a) also reveals this information. For this reason, we include another climate proxy, isotopic data of speleothems from the nearby Soreq Cave (see Fig. 3.19 (b)). In Sect. 2.1.1, the description of these data are linked to precipitation in the Eastern Mediterranean. Although the temporal space of this reference could be used for most periods of the Holocene (Bar-Matthews and Ayalon, 2007; Bar-Matthews and Ayalon, 2011), we choose to use the pattern correlation of the respective WPSs. This is due to the fact that the signal also reflects the composition of the source water of the Mediterranean Sea (Litt et al., 2012). In this case, correlation in temporal space poses some problems, but moving to the appropriate spectral space allows us to circumvent them and take advantage of the variability within these data. For all reconstructions within the Holocene and Levant, we choose the period 0 – 6.5 cal ka BP as a zone of non-negligible human impact (Schiebel, 2013). To account for this, the AP/NAP curve is used as a guide to annual precipitation only for ages prior to 6.5 cal ka BP. In

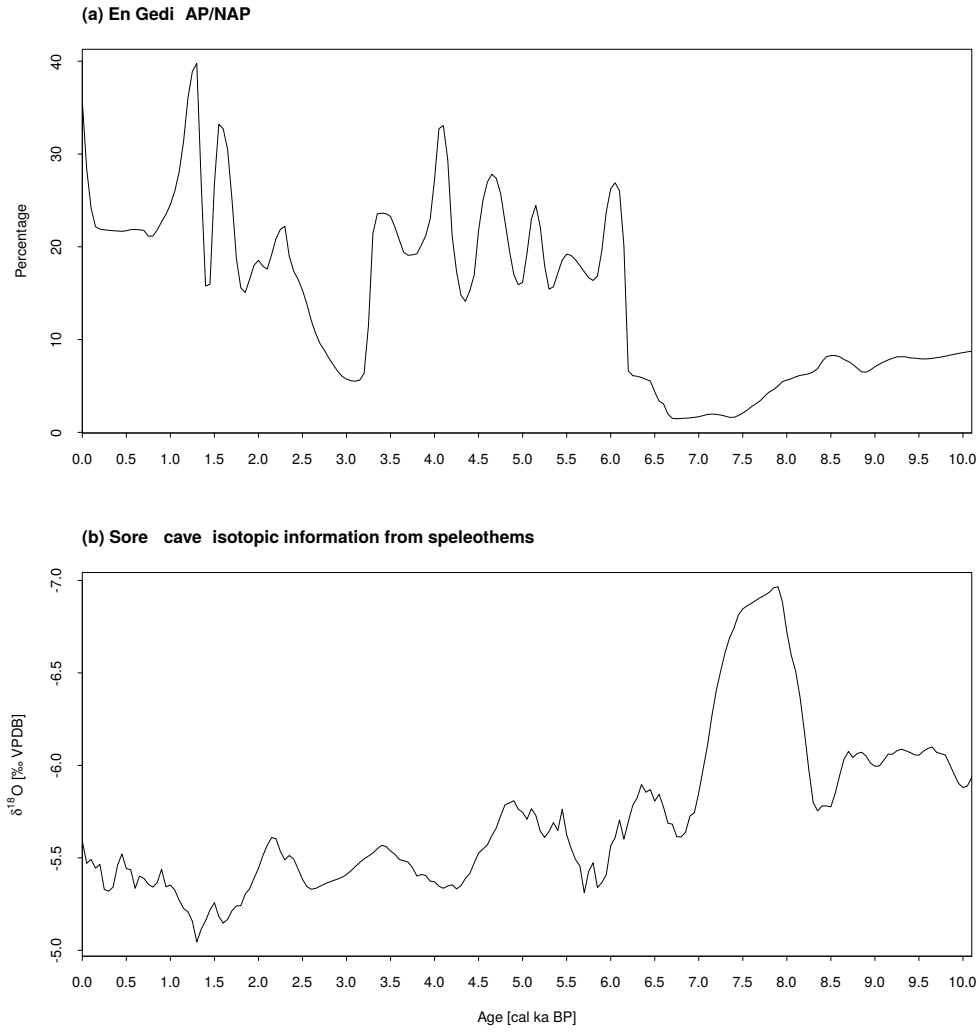


Figure 3.19: Reference curves based on (a) total terrestrial pollen sums: arboreal and non-arboreal pollen taxa from Litt et al., 2012 and (b) on isotopic information from speleothems from Soreq Cave (Bar-Matthews et al., 2003). The latter is expressed in δ^{18}O isotopic ratios in parts per thousand relative to the VPDB scale.

contrast, this climate variable is correlated with patterns of the WPS from the Soreq Cave throughout the reconstruction period.

The age-depth model data can be found in Litt et al., 2012. They are inserted into the Bayesian-based age-depth model Bacon and the results are used to convert depths to ages. These 20 radiocarbon dates indicate mostly continuous sedimentation, although they assume a hiatus at about 16 m core depth. The Bacon model is able to account for such abrupt events.

Reconstruction In the following, we describe the final reconstruction shown in Fig. 3.20. Panel (a) displays the reconstructed annual temperature T_{ANN} in °C and (b) the annual precipitation P_{ANN} in

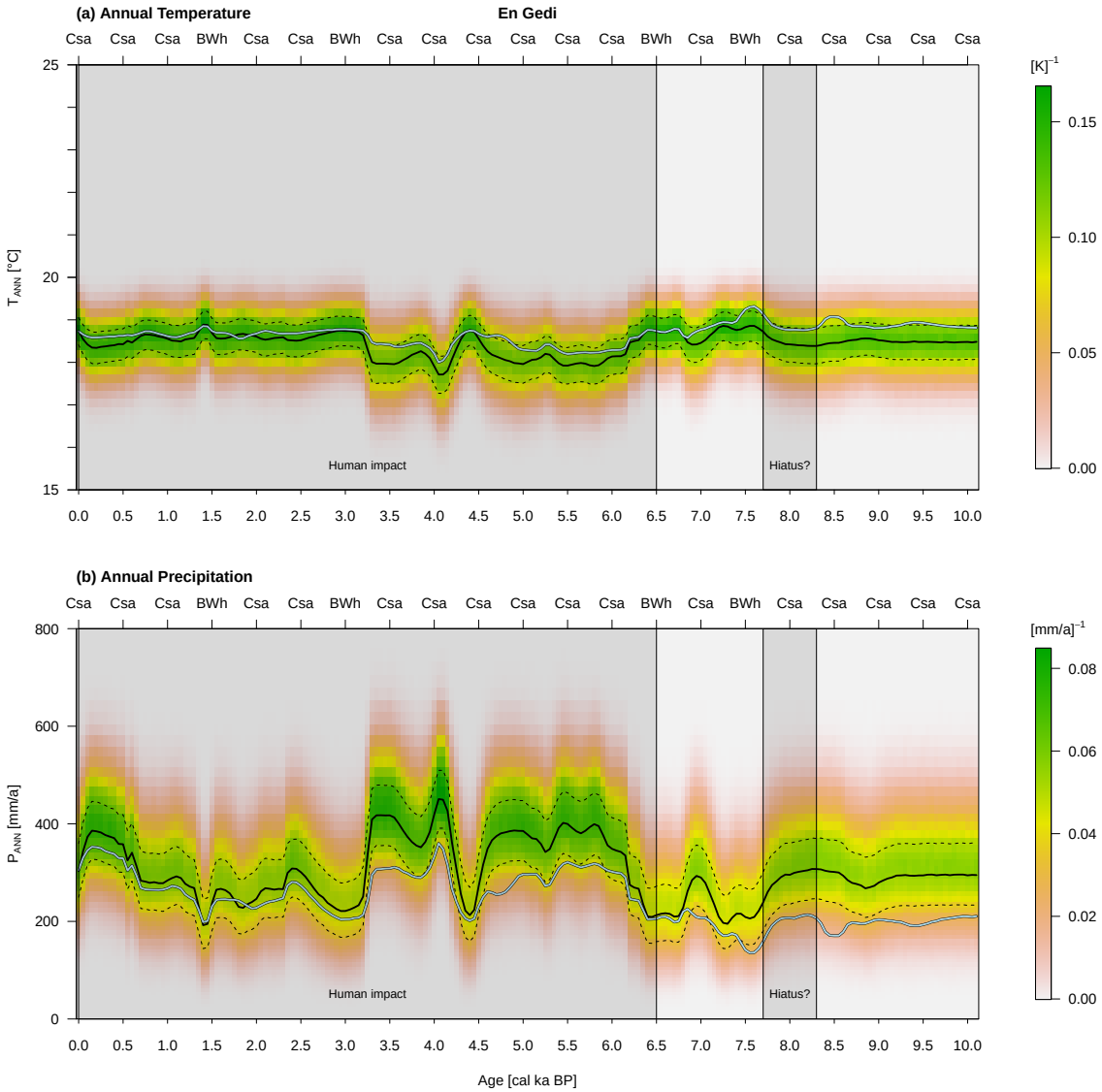


Figure 3.20: Quantitative reconstruction of the Dead Sea region based on plant and isotopic information in terms of calibrated kiloanni before 1950 AD (cal ka BP). In (a) is the reconstructed annual temperature in °C and in (b) is the annual precipitation in mm/a. The colors indicate the probability density values, the black solid lines its median, and the dashed black lines the first and third quartiles. The blue lines represent the medians of the reconstructions without additional information. The grayed out area on the left shows the zone of increased human impact and on the right the ages that contain approximately a hiatus. The axis above the graph shows the approximated Koeppen-Geiger classifications corresponding to the PDFs for the respective periods.

mm/a. At the beginning of the explanations, we focus on the solid black lines. Later, we will work out the difference between these reconstructions and others whose medians are shown as blue lines. Litt et al., 2012 divided their vegetation history into different age ranges. These are also used for this study.

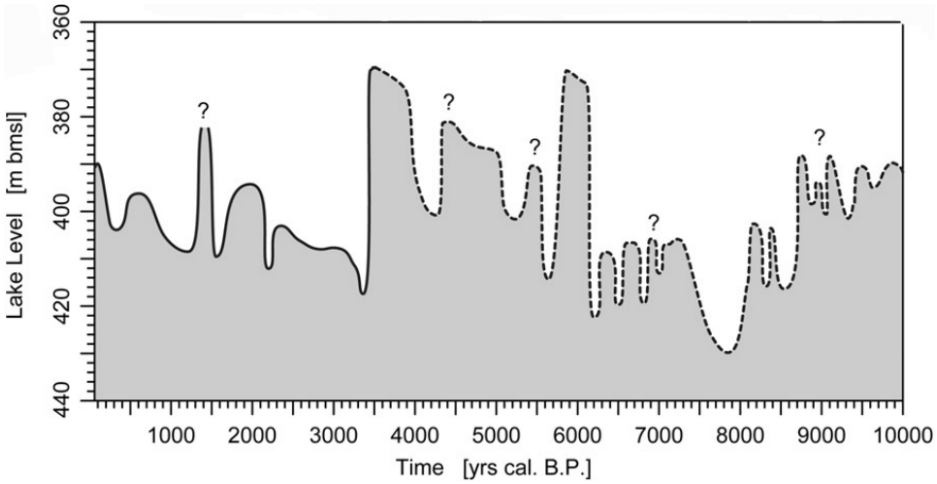


Figure 3.21: This figure is from Litt et al., 2012 and depicts the independently determined height of the Dead Sea lake level based on the work of Migowski et al., 2006.

First, some **general remarks** are made. Our reconstructed precipitation and temperature behave strongly anti-correlated with a value of about -0.95. This is due to the corresponding transfer functions, where higher temperatures are associated with lower precipitation. In the Levant, this is a typical pattern (cf. Fig. 3.17 (a) − (c)), as the hotter deserts and steppes (BWh/BSh) have on average much less precipitation than the colder Mediterranean coastal regions (Csa). When we describe a decrease in precipitation, it is always associated with an increase in annual temperature, unless otherwise noted. Although Litt et al., 2012 reconstruct winter and not annual temperature, such an anti-correlation also exists in their climate reconstruction.

The **10 − 7.6 cal ka BP** time interval spans the transition from Pre-Pottery Neolithic to Pottery Neolithic at ca. 8.4 cal ka BP (Schiebel, 2013) and Litt et al., 2012 describe it with low tree pollen values. These low percentages in Fig. 3.19 (a) correspond to the relatively low reconstructed lake levels in Fig. 3.21, including the notable decline from about 8.2 cal ka BP. The influence of humans during these ages could therefore be neglected. The authors assume a hiatus during these periods, which is shown in Fig. 3.20 with the right grayed out area. The reconstructed annual precipitation in (b) shows relatively low values around 300 mm. Together with the respective temperature around 18.5°, this section lies within the hot Mediterranean Koeppen-Geiger classification Csa. The AP/NAP reference curve in Fig. 3.19 (a) shows a very similar pattern to P_{ANN}. Bar-Matthews et al., 1999 also describe this with a decrease in precipitation. In our reconstruction, the decrease in P_{ANN} starts relatively slowly at first at the same time.

For the remainder of the Pottery Neolithic (**7.6 − 6.3 cal ka BP**), Litt et al., 2012 describe a decline in Mediterranean trees and shrubs. They conclude that the Mediterranean vegetation belt has shifted northward and human influence is relatively low due to low settlement. The reconstructed precipitation shows the lowest and the temperature the highest values of the whole reconstruction period. This is

associated with a change from the hot Mediterranean climate to the hot desert climate BWh, which underscores the conclusion of Litt et al., 2012. Both reference curves show their minimum during these periods. The AP/NAP curve increases by about 6.5 cal ka BP, while the isotope-based curve does so 1000 years earlier. This could be the cause of the change to Csa by 7 cal ka BP and the return to BWh by 6.5 cal ka BP. The quantitative climate reconstruction in Litt et al., 2012 shows a reduced Mediterranean biome prior to 6.5 cal ka BP with positive winter temperature anomalies and an expected annual precipitation value of 280 mm/a. As for the trend analysis, their reconstruction shows a negative trend in P_{ANN} for these years. Such features are also evident in our new reconstruction of En Gedi.

The next time span of **6.3 – 3.2 cal ka BP** includes the beginning of the Chalcolithic (ca. 6.5 – 5.5 cal ka BP), the Early and Middle Bronze Age (ca. 5.5 – 4.15 and ca. 4.15 – 3.5 cal ka BP), and the Late Bronze Age (ca. 3.5 – 3.15 cal ka BP) (Schiebel, 2013). Litt et al., 2012 note an increase in Mediterranean woodland, which could be due to an increased ratio of precipitation to evapotranspiration. However, human impact increases during this period. For this reason, the reconstructed P_{ANN} is no longer correlated with the AP/NAP curve. As a guideline, we are left only with the pattern correlation of this climate variable with the WPS from Soreq Cave. The reconstructed precipitation shows the highest values with a maximum around 450 mm during this period. On the other hand, the annual temperature drops to almost 17.5 °C. This change in climate is also reconstructed in Litt et al., 2012, consistent with relatively high lake levels. They found precipitation minimums at about 5.6 and 4.3 cal ka BP. In comparison, our reconstruction shows minima of P_{ANN} at about 5.4, 4.3, and 3.8 cal ka BP. Since these declines sometimes last several hundred years, the archaeological transitions described above could be related to these climate fluctuations. Bar-Matthews and Ayalon, 2007 also connect the corresponding variations in Fig. 3.22 (b) to cultural changes in the Eastern Mediterranean and to the periodic behavior of Bond events. This can be seen in the corresponding WPS in Fig. 3.22 (a), which shows a significant power branch with a period of about 1.4 ka. Comparing this with the WPS of P_{ANN} in Fig. 3.22 (c), two branches with significant power of about 1 ka and 2 ka periods stand out at the beginning of this time interval. Later, both merge to a period of about 1.5 cal ka BP. This spectral pattern is also evident in the WPS of IRD in the North Atlantic from Bond et al., 2002 in Fig. 2.7 (d), underscoring that our reconstruction has similar variations to some Bond events.

The next period of **3.2 – 2.3 cal ka BP** begins with the transition from the Late Bronze Age to the Iron Age. Litt et al., 2012 notes a dramatic decline in the Mediterranean biome, reconstructed precipitation, and lake level, that cannot be explained by human influences. Our reconstruction shows a similar pattern and a change to a hot desert climate. This climate change could be associated with the so-called 3.2 ka drought event and with the Late Bronze Age collapse (Schiebel and Litt, 2018). The corresponding WPS in Fig. 3.22 (c) shows the highest power with a period around 1.5 ka at the beginning of this period. At the end of the Iron Age (ca. 2.6 cal ka BP), a moderate increase in oak and olive pollen is accompanied by an increase in precipitation and the transition to a hot Mediterranean

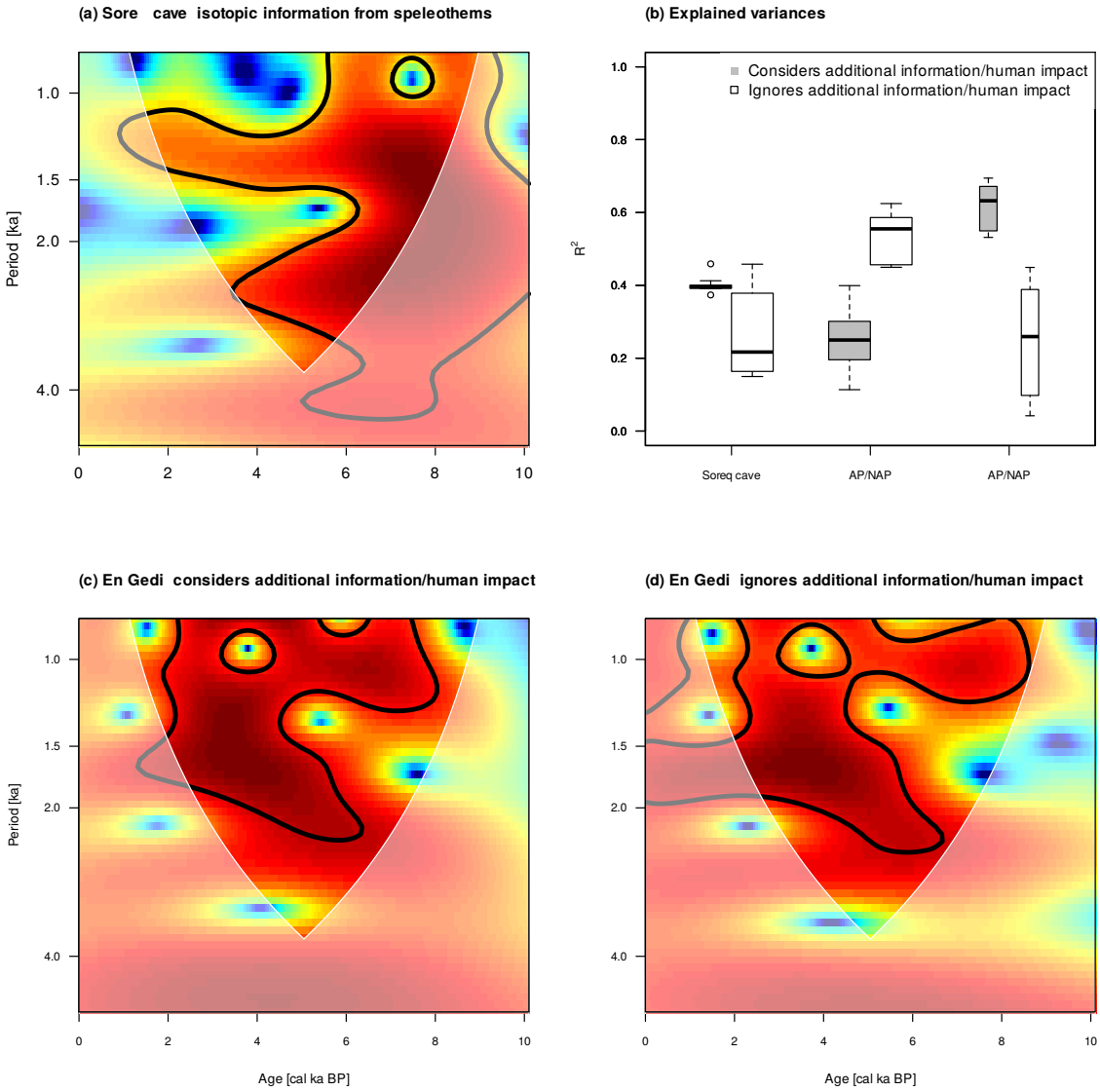


Figure 3.22: Wavelet power spectrum of (a) isotopic information from speleothems from Soreq Cave (Bar-Matthews et al., 2003), (c) the reconstructed annual precipitation with additional information, and (d) the reconstructed annual precipitation without additional information. The gray boxplots in (b) show the results with the additional information from speleothems. From left to right are plotted, first, the explained variances between the WPSs of the medians of P_{ANN} and the WPSs of the Soreq cave, and second, the medians of P_{ANN} and the entire AP/NAP curve. The thinner boxplot indicate R^2 between the medians of P_{ANN} and the section of AP/NAP where human influence is lower. The white boxplots show the same calculations, but without the additional information from the speleothems. The basis for these calculations are the selected samples of the posterior reconstruction.

climate.

In the years **after 2.3 cal ka BP**, human influence continues to increase. Orland et al., 2009, using

isotopic information from Soreq Cave, describe a decrease in precipitation during 1.9 – 1.3 cal ka BP that is also reflected in the level of the Dead Sea. They assume that this climate change weakened the economic system of the Roman and Byzantine Empires, which contributed to the decline of their rule in the Levant. This decrease in precipitation is also evident in the reconstructions of Litt et al., 2012 and in this study. At about 1.3 cal ka BP, the corresponding Koeppen-Geiger climate classification changes to BWh. Thereafter, P_{ANN} increases again, and both reconstructed climate variables approach the recent climate measurements for the youngest time interval.

Sensitive study The final reconstruction contains the pollen information from the core, including the AP/NAP, the isotopic data from Soreq cave, the recent climate measurements, and the human impact. In contrast, the reconstruction represented by the blue line in Fig. 3.20 does not take into account the additional information from the speleothems and the human influence on the vegetation.

In Fig. 3.22 (b) we can see the different explained variances. The results from the Soreq cave in the case of additional information from this site are higher (about 0.4) than the results without this isotopic information (between 0.2 and 0.4). Comparing the medians of P_{ANN} of the selected posterior

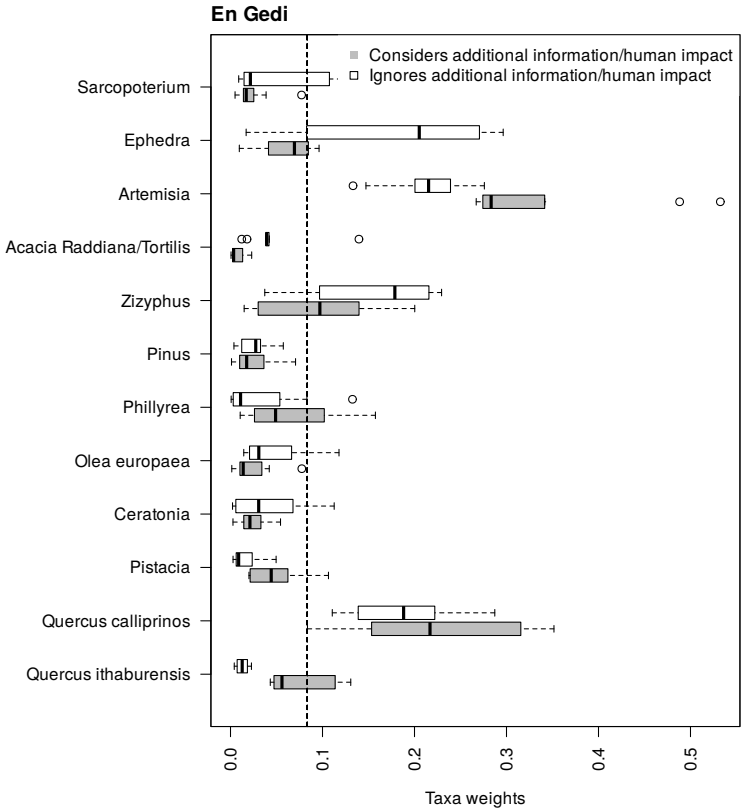


Figure 3.23: Posterior taxa weights of the reconstructions with and without additional information. The dashed line indicates the prior weights for each taxon.

samples with the whole AP/NAP curve, the corresponding R^2 is lower (from 0.15 to 0.4) than for the reconstructions with the whole AP/NAP information (about 0.55). The thinner boxplots reflect the explained variances of AP/NAP where human influence is lower. In contrast to the corresponding calculations for the entire region, R^2 is higher (about 0.65) when human influence is taken into account compared to the thin white boxplot (from 0.05 to 0.45).

This shows that the algorithm is able to properly use the information from the caves and the human influences to find a Pareto optimum between all these variables. In Fig. 3.22 (c) and (d), these differences are most pronounced for ages older than about 6 cal ka BP and thus fall within the range of lower human influence. There, the power is higher when additional information/human influences are considered, with a local maximum around 7 cal ka BP and a period of about 1000 years. Comparing this with the annual precipitation in Fig. 3.20 (b), it reflects the local maximum around this time, which is less pronounced in the blue line. As mentioned earlier, this could be due to the relatively large variation within the isotopic information. However, two taxa indicate a local maximum during these ages. Quercus ithaburensis and Pistacia show a slight increase in the pollen spectrum, which is detected by the algorithm and leads to a relatively higher weighting (see Fig. 3.23). Litt et al., 2012 associate these plants with Mediterranean vegetation, resulting in a wetter and cooler climate during this period. In addition, the reconstruction that does not take into account the information from Soreq Cave and human influences assigns a higher weight to the taxon Ephedra, resulting in the largest change in the weighting of the taxa. Moreover, the explained variance of the Ephedra pollen curve compared to AP/NAP is the highest of all plant information used, with a correlation of about -0.75. This tuning results in an overall lower R^2 of the final reconstruction with respect to the pollen reference curve. Since Ephedra belongs to the semi-arid steppe vegetation corresponding to a relatively dry and hot climate (Litt et al., 2012), the blue curve shows, on average, a reconstruction with less precipitation and higher temperatures. The maximum differences for P_{ANN} are about 100 mm and for T_{ANN} ca. 0.5°C.

Summary and conclusion The quantitative climate reconstruction of the Dead Sea region during the Holocene includes not only paleobotanical data, but also isotopic information from speleothems from nearby Soreq Cave. In addition, this section outlines a way to deal with the non-negligible impacts of humans on vegetation. Although the final reconstruction is based only on spectral information from speleothems for ages younger than 6.5 cal ka BP, the results of P_{ANN} are comparable to the Dead Sea lake level (cf. Fig. 3.21). However, it should be mentioned here that the fluctuations of the lake level are due to the freshwater balance evaporation - precipitation in the catchment area of the lake. Therefore, it may be beneficial in future work to additionally link plant proxies to evaporation. Furthermore, our reconstruction results are in agreement with those of Litt et al., 2012. These describe the MH (6.5 – 3.5 cal ka BP) as a relatively wet period with maximum values more consistent with our final reconstruction than with the lower precipitation that results when information from speleothems

and human influences are ignored. Moreover, the algorithm is able to find climate changes that can be associated with Bond events and known archaeological and cultural changes in the Levant (Schiebel, 2013; Bar-Matthews and Ayalon, 2011; Orland et al., 2009).

For these reasons, we conclude that the use of the above information is beneficial for such reconstructions. They can be easily extended by other proxy information such as lake levels or climate constraints from other quantitative reconstructions. In these cases, the first information could be used as another reference curve and the second as the data basis for additional climate anchor points. The latter option is described in Sect. 3.3.5 and a reconstruction without any climatic anchors is discussed in the following section.

Late Penultimate to Last Glacial

Introduction In this section, we aim to reconstruct the climate of the Dead Sea region over a period of about 14.5 – 147 cal ka BP. Therefore, we cannot use information from recent climate measurements for the youngest time slice of this reconstruction period. Thus, this reconstruction tests whether the algorithm is capable of generating reasonable climate values without such information.

The reconstruction is based on pollen data presented in Chen and Litt, 2018 and Miebach et al., 2019. They investigate the deepest borehole, 5017-1-A, with a total length of 455.34 m, drilled in 2010/2011 in the northern Dead Sea basin. Unlike the previous AP/NAP references, the corresponding curve of this core is difficult to correlate with any of the reconstructed climate variables. Miebach et al., 2019 describe the amount of arboreal pollen as an indicator of available moisture, which depends mainly on precipitation and evapotranspiration. In particular, when studying a period that includes glacials and interglacials in this region, information on evapotranspiration is needed to use the AP/NAP curve as a reference. However, since we only consider T_{ANN} and P_{ANN}, we will not use this reference curve for the reconstruction. Instead, we include isotopic information not only from Soreq Cave but also from the Mediterranean Sea, which is shown in Fig. 3.24. The latter are grouped together in the so-called Medstack, which comprises several locations (Colleoni et al., 2012) and can be linked to the sea surface temperature of the Mediterranean Sea (Wagner et al., 2019). Therefore, we use this proxy as a guideline for annual temperature in the Dead Sea region and the WPS of the Soreq Cave for the reconstructed annual precipitation.

The data for the Bayesian-based age-depth model can also be found in the studies of Chen and Litt, 2018 and Miebach et al., 2019. For most dates younger than about 44 cal ka BP, the uncalibrated radiocarbon ages are used and fitted to the Bacon model. The corresponding errors are relatively small, resulting in a much sharper age-depth transformation than for ages older than 40 cal ka BP. This section mainly uses the calendar ages of the uranium-thorium data from Dead Sea Core 5017-1-A. For ages older than about 120 cal ka BP, Chen and Litt, 2018 use the anchor ages of LR04 to correctly determine the transition from MIS 6 to MIS 5.

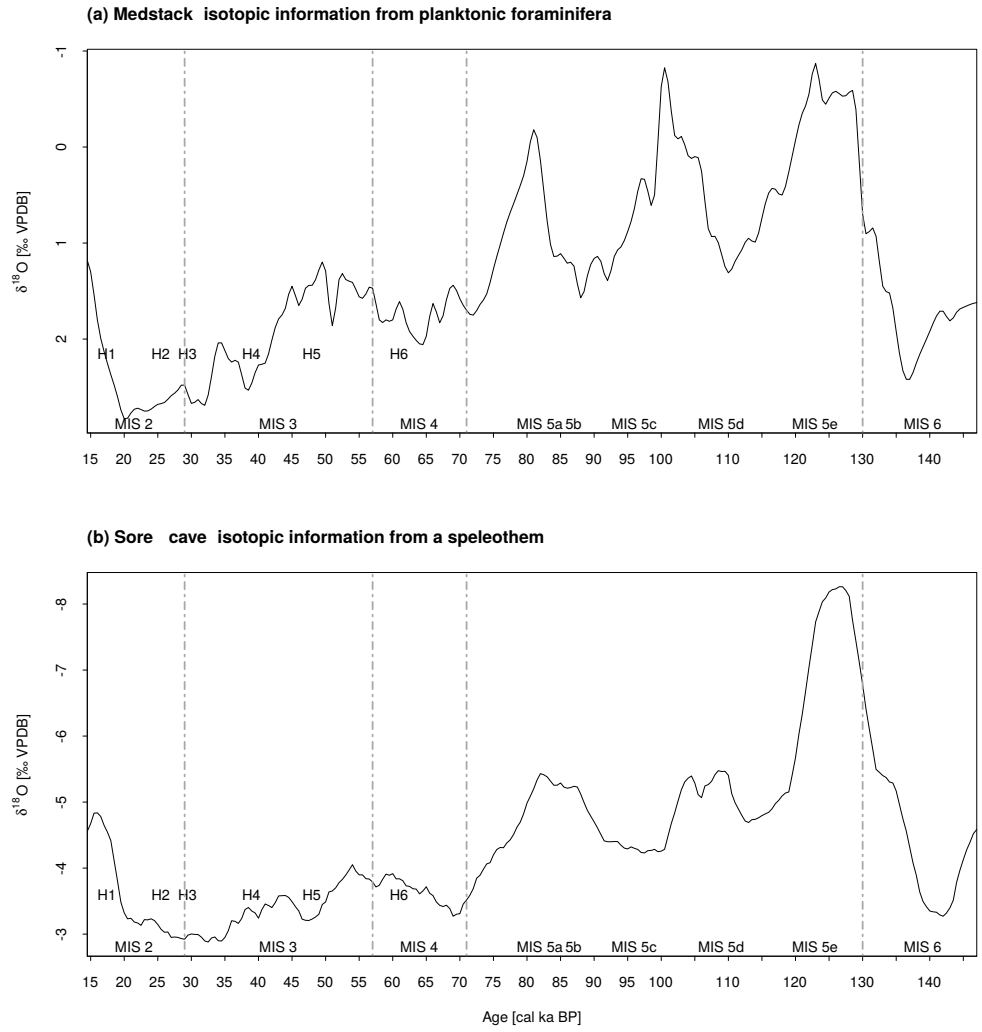


Figure 3.24: Reference curves based on (a) isotopic information from planktonic foraminifera from stacked marine cores of the Mediterranean Sea (Colleoni et al., 2012) and (b) on isotopic information from a speleothem from Soreq Cave (Grant et al., 2012). Both proxies are expressed in δ^{18}O isotopic ratios in parts per thousand relative to the VPDB scale. H $1-6$ describe the relative timing of Heinrich stadials according to Hemming, 2004 and MIS $2-6$ the marine isotopic stages following Lisiecki and Raymo, 2005. The vertical gray lines indicate the transitions of the MISs.

Reconstruction In the following, we describe the final reconstruction shown in Fig. 3.25. Panel (a) displays the reconstructed annual temperature T_{ANN} in °C and (b) the annual precipitation P_{ANN} in mm/a. At the beginning of the explanations, we focus on the solid black lines. Later, we will work out the difference between these reconstructions and others whose medians are shown as blue lines. Chen and Litt, 2018 and Miebach et al., 2019 have divided their explanations into different age ranges. These are also used for this study.

For this reason, we start with the calibrated years between **147.3 – 130.9 cal ka BP**. These include

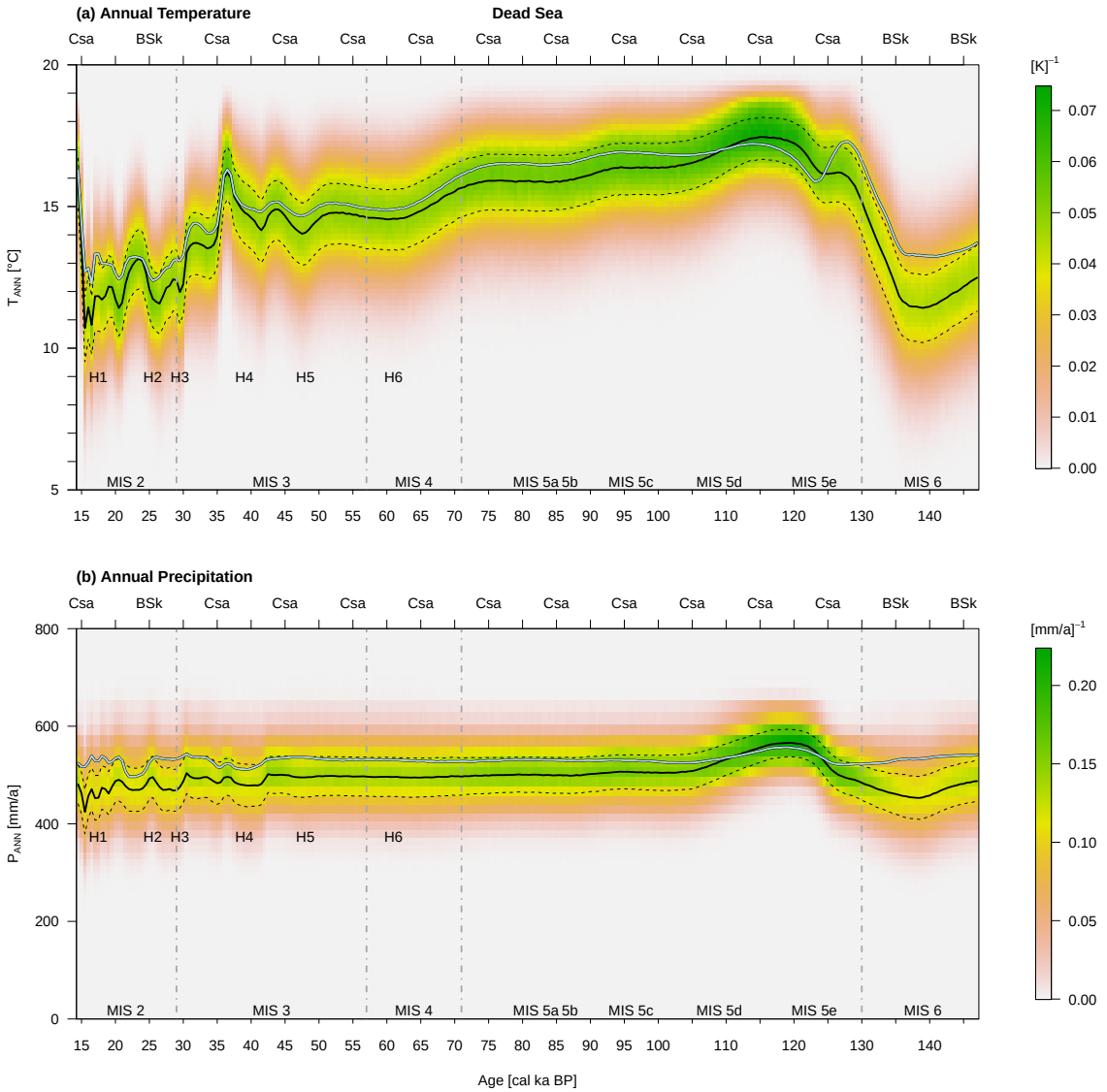


Figure 3.25: As Fig. 3.20, but for the Dead Sea. H 1 – 6 describe the relative timing of Heinrich stadials according to Hemming, 2004 and MIS 2 – 6 the Marine Isotope Stages following Lisiecki and Raymo, 2005. The vertical gray dashed dotted lines indicate the transitions of the MISs.

the late Penultimate Glacial and late MIS 6. Chen and Litt, 2018 describe that the vegetation is dominated by wormwood trees and that the proportion of trees and shrubs is relatively moderate. The reconstructed T_{ANN} averages 12.25 °C and P_{ANN} 470 mm. During 134 – 140 cal ka BP, we find the respective minimums at about 11.5 °C and 450 mm. Inserting these values into the transfer function $\mathbb{P}(\boldsymbol{P_s}|\boldsymbol{C},\boldsymbol{\psi})$ of Quercus ithaburensis and Quercus calliprinos, we obtain probabilities of approximately 45 % and 50 %, respectively. This is consistent with research by Chen and Litt, 2018 that reduced

temperature still supports survival of temperate trees. They also conclude that this reduction leads to low evaporation rates, resulting in a positive freshwater balance in the lake's catchment. Progressive aridification during the Penultimate Glacial in the Dead Sea region results in the Koeppen-Geiger approximation BSk. For comparison, this cold steppe climate is found today in some parts of Central Anatolia (cf. Fig. 3.17 (c) and (d)), where steppe vegetation predominates (Kurt et al., 2006). Looking at the isotopic data from Medstack and the Soreq cave, a very similar pattern emerges during MIS 6 with local maxima. Nevertheless, a displacement of these by about 5 ka is present. This shift is also found in the climate reconstruction, where annual temperature increases at about 135 cal ka BP and annual precipitation ca. 5 ka later.

At the beginning of the **130.9 – 115.5 cal ka BP** period, which includes the Last Interglacial optimum and corresponds to MIS 5e, frost-sensitive pistachio pollen increases (Chen and Litt, 2018). This should be accompanied by an increase in temperature, which confirms our reconstruction model. T_{ANN} is about 16 °C and P_{ANN} is about 500 mm. After ca. 125 cal ka BP, there is a further increase in both climate variables. Here Chen and Litt, 2018 identifies the onset of the LIG optimum, which duration is about 124.2 – 115.5 cal ka BP. This is coupled with an increase in grasses, trees and shrubs in the mountains. The reconstructed climate values have their maximum in this period with values around 18 °C and 570 mm. This rise in precipitation contrasts with the course of the reconstructed lake level (Torfstein et al., 2015), which shows a decline during MIS 5e. Chen and Litt, 2018 explain this with the increased vegetation mentioned above, leading to enhanced evapotranspiration in the area of the main headwaters of the Dead Sea. This could lead to a decrease in freshwater supply and thus to a drop in the lake level, even though precipitation has increased. Thus, not only the correlation with AP/NAP proves difficult, but also the comparison with the reconstructed lake level. After about 120 cal ka BP, there is a variation in the pollen spectrum with relatively constant pistachio pollen but a decline in Mediterranean sclerophylls. This change does not affect our climate reconstruction because their mean values remain comparatively stable. With respect to MIS 6, the Koeppen-Geiger classification of the Dead Sea region changes to a hot Mediterranean climate Csa throughout the MIS 5 period. This classification occurs most frequently in the climate reconstruction of En Gedi during the Holocene. Although the reconstructions do not include information on interannual climate, the Koeppen-Geiger approximation suggests mild/wet winters and hot/dry summers. This is consistent with the statements of Chen and Litt, 2018, which indicate the same climate. Furthermore, the reconstructed temperatures show similar values to those of the Holocene. In contrast, precipitation appears to be somewhat higher. This is in agreement with the results of Bar-Matthews et al., 2017, which indicate wetter conditions during the LIG in the southern Levant than today. Both isotope-based reference curves show their extreme values during the same period with an increase after 120 cal ka BP. In doing so, Medstack does not rise as steeply as the Soreq cave, resulting in a shift in extrema during the remainder of MIS 5.

The early Last Glacial within **115.5 – 89.1 cal ka BP** includes a decline in pistachio and an almost complete absence of sclerophyllous pollen (Chen and Litt, 2018). From this, they infer a cooling

trend that could be associated with the MIS 5d stadial. During the period 105-99.6 cal ka BP, there is a slight rebound of deciduous oaks. Chen and Litt, 2018 interpret this period with the interstadial MIS 5c. Both climate variations are seen for both climate variables in our reconstruction, with the corresponding amplitudes in P_{ANN} being relatively small. Since 99.6 ka, the steady decline in pistachio populations has been accompanied by a decline in both climate variables. This could be indicative of stadial MIS 5b. As mentioned earlier, the alternating MIS 5 stadials and interstadials are shifted in both isotope-based references. Nevertheless, the shift of MIS 5d with respect to the climate reconstructions is much smaller and in the opposite direction than indicated in the guidelines. In contrast, the interstadial MIS 5c fits these marine isotope data quite well, although they are not known to the algorithm in this case. This behavior emphasizes the restriction of the explained variances to not too high values, which allows the reconstruction method to find a compromise between all included proxy information.

Miebach et al., 2019 describe the age **88 – 62.6 cal ka BP** with rapid paleoenvironmental changes, with a dominance of herbs and dwarf shrubs and some moderate Mediterranean woodland elements. In addition, the pollen spectrum could contain a Dansgaard-Oeschger signature. In our reconstruction, both climate variables show stable behavior during the end of MIS 5b and during MIS 5a. It should be noted again that the age-depth transformation leads to a relatively blurry picture due to the high uncertainties in these periods. Therefore, the information about such rapid climate changes may not be resolved. After the transition to MIS 4 at about 72 cal ka BP, there is a drop from 16 °C to 14.5 °C in T_{ANN}. The local minimum at about 63 cal ka BP coincides with the Heinrich stadial 6 (Hemming, 2004). The corresponding amplitudes in P_{ANN} continue to decrease. The reference curves show a relatively correlated pattern, although the timing of H 6 in Fig. 3.24 (a) is more shifted than in Fig. 3.24 (b).

The years **62.6 – 34.7 cal ka BP** contain the late MIS 4 and the early/mid MIS 3. Miebach et al., 2019 characterize this period with the strongest variation in vegetation between MIS 5 and MIS 1 with an increase in AP and the highest mean percentages of Quercus ithaburensis. Relatively high fluctuations can also be observed for T_{ANN}. The variance of P_{ANN} increases again after about 42 cal ka BP, indicated by lower values of the probability density within the uncertainty ranges. Prior to this, T_{ANN} drops by about 1 °C at about 47 cal ka BP, which is consistent with H 5. This deviation is not as clearly visible in the corresponding reference curve (see. Fig. 3.24 (a)) as in the Soreq cave. This is followed by a temperature increase with a subsequent analogue decrease at about 42 cal ka BP. In contrast to the previous decline, this can be seen in the Medstack curve, but not so clearly in the patterns of the speleothem. Furthermore, P_{ANN} also shows a decrease. These climate fluctuations could be related to H 4, which occurs around 40 cal ka BP. The corresponding WPS of P_{ANN} in Fig. 3.26 (c) shows a relatively high non-significant power after about 45 cal ka BP with periods between 1 and 3 ka. These variations could include some Dansgaard-Oeschger (9 – 11) events (cf. Fig. 2.3). From now on, the age-depth model has a much lower variance, which results in a sharper age-depth

transformation. In this way, the rapid climate fluctuations can be resolved. After these fluctuations, both variables increase again until about 35 cal ka BP.

During late MIS 3 (**34.7 – 30.6 cal ka BP**), Miebach et al., 2019 describe a complete absence of thermophilous Olea europaea and thus a reduced temperature. There are also large proportions of steppe vegetation accompanied by Mediterranean forest components. To test these observations, they model a biome shift with a decrease of $-3\,^{\circ}$C and $-15\,\%\,P_{ANN}$. These values are taken from Stockhecke et al., 2016 for MIS 3. The result is a well-mixed equilibrium between the three biomes included, the Mediterranean woodland, Irano-Turanian steppe and Saharo-Arabian desert, with a higher probability for the second biome compared to today for the Dead Sea region. Although their temperature variations are based on the mean summer and winter temperatures, we compare them to the relative differences we have reconstructed. Our reconstruction shows a mean annual temperature of about 14.5 °C and P_{ANN} of 490 mm for MIS 3. The difference between these values and those of the reconstructed LIG optimum agrees quite well with those of Miebach et al., 2019. Both reference curves also show an increase for this period, accompanied by a decrease to 12 °C and 470 mm. This decline marks the transition to MIS 2, along with H 3, which is consistent with the climate variations we simulated.

The period **30.6 – 15.4 cal ka BP** includes a high concentration of Artemisia and a complete absence of Olea and frost-sensitive Pistacia pollen during MIS 2 (Miebach et al., 2019). They conclude that this is the coolest period after the Penultimate Glacial with a sparse vegetation cover. Again, they model the distribution of biomes with a decrease of $-6\,^{\circ}$C and $-30\,\%\,P_{ANN}/+100\,\%\,P_{ANN}$. The results are an under-/overestimation of the Mediterranean biome. In our reconstruction, we have a mean T_{ANN} of about 12 °C and P_{ANN} of 470 mm for MIS 2. This leads to differences from the LIG optimum of $-6\,^{\circ}$C and -20 %mm respectively. The reconstructed climate is comparable to that of the Penultimate Glacial. This is also indicated by the change from the hot Mediterranean climate classification Csa to the cold steppe BSk. The minimum values for both variables are about 11 °C and 425 mm and are located shortly after the H 1 event. The reconstructed lake level of Lake Lisan shows the highest values during MIS 2 compared to MIS 3 and 4 (Torfstein et al., 2013). This is in contrast to the reconstructed precipitation, which shows the lowest values during MIS 2. One explanation for this behaviour could be the changed evaporation and evapotranspiration rate for this period in this region. During the LGM, Schiebel, 2013 describe a woodland clearance in the northern mountains of Israel that could reduce plant transpiration and promote a positive freshwater balance in Lake Lisan (Miebach et al., 2019). In Miebach et al., 2017, vegetation during the late Last Glacial is described for the Sea of Galilee. In their palynological study they find a similar vegetation pattern with a dominance of steppe. Due to the lower temperatures and higher relative humidity, Miebach et al., 2019 conclude that evaporation decreases, resulting in higher effective humidity. Our reconstruction, based on several proxies, underlines these considerations. Corresponding features can be seen in the isotope curves. However, the decline of the two curves starts later than in the reconstruction. Furthermore, the WPS

of P_{ANN} shows significant power for the late MIS 2 with a local maximum of a period around 1.5 ka, although the corresponding reference WPS of the Soreq cave does not contain these features. This could indicate the last Dansgaard-Oeschger event, which can also be recognized in the pollen spectrum. At the beginning of MIS 2, H 2 is accompanied by a local minimum in T_{ANN} and a local maximum in P_{ANN}. In the corresponding WPSs we can also observe a relatively high power (not significant) with a period between 1 and 2 ka. This suggests that the pollen data also contain the information from DO 3

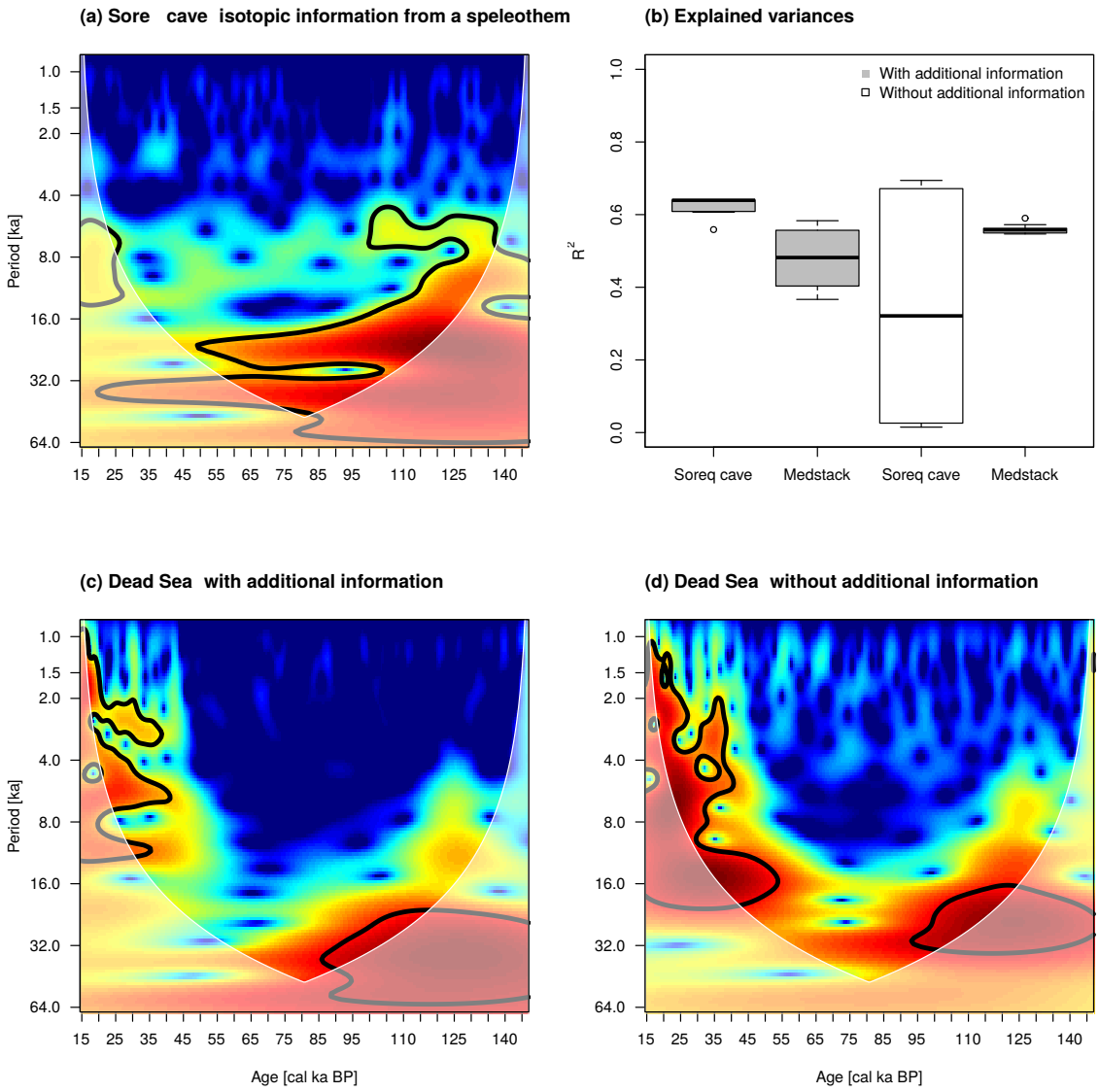


Figure 3.26: As Fig. 3.22, but for the Dead Sea. In contrast, panel (b) does not contain information on AP/NAP. Instead, the explained variances are presented in relation to Medstack. In this context, it should be added that the boxplot on the far right shows the results without additional information.

and/or DO 4 (cf. Fig. 2.3).

The Late Glacial (**15.4 – 14.2 cal ka BP**) is characterized by a spread of pioneer plants and an increase in Pistacia and Olea europaea (Miebach et al., 2019). This leads to a return to higher temperatures, which is also evident in our reconstruction. The Koeppen-Geiger classification returns to the hot Mediterranean climate Csa with temperatures above 15 °C and P_{ANN} of 480 mm. This variation is also marked in the isotopic references and is called BA interstadial or DO 1 (Rasmussen et al., 2014). All in all, the similarities between the WPS patterns in Fig. 2.3 and the WPS of the reconstructed annual precipitation for this period in Fig. 3.26 (c) are quite close. On the one hand, DO 1 is accompanied by significant power, but DO 2 is somewhat more difficult to detect, and DO 3 – 4 is again more noticeable, although significance is lacking in our reconstruction.

Sensitive study The final reconstruction contains the pollen information from the core, without the AP/NAP curve. Isotopes used are not only those of a speleothem, but also those of planktonic foraminifera from the Mediterranean Sea. In contrast, the reconstruction represented by the blue line in Fig. 3.25 does not take into account the additional information from the speleothem. This results in explained variances that vary from almost 0 to 0.65 with a mean of about 0.3 (see Fig. 3.26 (b)).

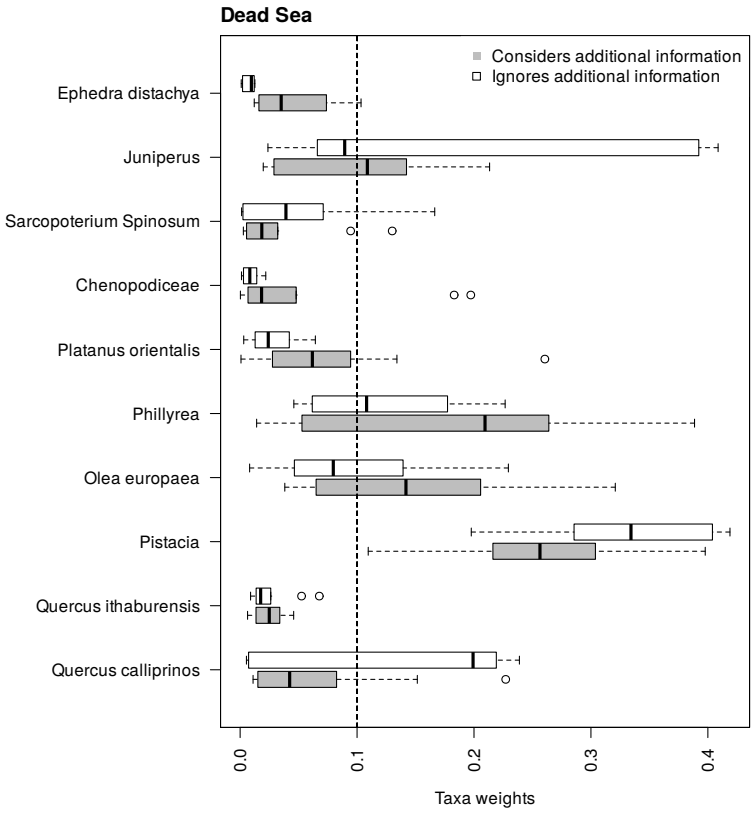


Figure 3.27: As Fig. 3.23, but for the Dead Sea.

Since the Pareto calculations in this case correspond to the highest values of a single variable, the explained variances with respect to Medstack range from 0.55 to 0.65.

By including the additional information, the Pareto optimum involves a compromise between these two proxies. Thus, R^2 from Medstack ranges from 0.4 to 0.6 and is slightly lower than those without additional information. The explained variance related to the Soreq cave is much higher, ranging from $0.6 - 0.65$ in summary. These higher values lead to the main differences in Fig. 3.26 (c) and (d). Around 50 cal ka BP, the WPS without additional information shows high significant power for periods around 16 ka. These higher values lead to the main difference from Fig. 3.26 (c) to (d). For ages around 50 cal ka BP and periods around 16 ka the WPS without additional information reveals high significant power. Instead, the WPS with additional information follows the spectral behaviour of the Soreq cave.

The main difference when looking at the reconstructions is the amplitude of the two variables over the entire period. The maximum values of the blue lines are relatively similar, while the minima differ significantly in some areas. Here they are higher for both variables, leading to the largest differences during the coldest periods, namely MIS 6 and MIS 2. In addition, the fluctuations during the Heinrich events are not as strong as in the case of the inclusion of isotopes from the Soreq cave.

All in all, the reconstruction with both information agrees better with the qualitative climate reconstruction of Chen and Litt, 2018 and Miebach et al., 2019 than when only the Medstack reference is used. This is also confirmed by the fact that the maximum of T_{ANN} in the blue line is not during the LIG optimum, but just after the onset of MIS 5e. Looking at the posterior taxa weights for both models in Fig. 3.27, it is noticeable that, with respect to oaks, Quercus calliprinos is much more weighted than for the case without additional information. This taxon has the highest pollen percentages in cooler periods, resulting in the differences described above. Phillyrea, Olea, and Pistacia, however, occur mainly during early MIS 5. These taxa have well mixed weights for the case with additional information and a maximum mean weight for pistachio. Although this is also the case for the reconstruction without additional information, the other important taxa for early MIS 5 are much less weighted. This results in the shifted LIG optimum when isotopic information from the Soreq cave are neglected.

Summary and conclusion In this section we describe a climate reconstruction of annual temperature and precipitation in the Dead Sea region over a period of about $14.5 - 147$ cal ka BP. Isotopic information from Soreq Cave and the Mediterranean Sea are used as a reference. Although no a priori climate anchor points are used for the reconstruction period, the algorithm provides reasonable climate values for both variables. The reconstructed T_{ANN} for the LIG optimum match those of En Gedi during the Middle Holocene in Sect. 3.3.1. In comparison, P_{ANN} is slightly higher than today. The corresponding Koeppen-Geiger approximations, indicating a hot Mediterranean climate, hold for most of the reconstruction period. Exceptions are the late Penultimate Glacial (late MIS 6) and the Late

Glacial (MIS 2). During these periods, the classification transitions to a cold steppe climate caused by a decrease in T_{ANN} and P_{ANN}. Nevertheless, the lowest climate averages of these glacial conditions result in occurrence probabilities $\mathbb{P}(\boldsymbol{P}|\boldsymbol{C_{Glacial}})$ for Quercus ithaburensis and Quercus calliprinos of about 45 % and 50 %, respectively. If the isotopic information from Soreq Cave is not taken into account, the reconstruction shows higher values for both variables during the cold periods. As a result, the corresponding Koeppen-Geiger classifications remain in the Mediterranean climate. This leads to the conclusion that a compromise between the two isotopic sources (Medstack and Soreq Cave) results in a climate reconstruction that is more consistent with what Chen and Litt, 2018 and Miebach et al., 2019 describe.

3.3.2 Lake Kinneret

Lake Kinneret, also known as the Sea of Galilee or Lake Tiberias, is located about 150 km north of the Dead Sea in the Jordan Valley. The water level of this freshwater lake is about 212 m bmsl and is mainly fed by the Jordan River (Schiebel and Litt, 2018). The climate diagram in Fig. 3.28 of the grid point closest to the Sea of Galilee shows a hot Mediterranean climate Csa. The mean annual temperature in this region is about 18.5 °C, and the average annual precipitation is around 470 mm. Summers are dry and hot, while winters have a mild and humid climate. Most rainfall in this region occurs between October and April. The unique orography of the Jordan Valley results in a strong precipitation and temperature gradient. Lake Kinneret thus lies on the southern edge of the

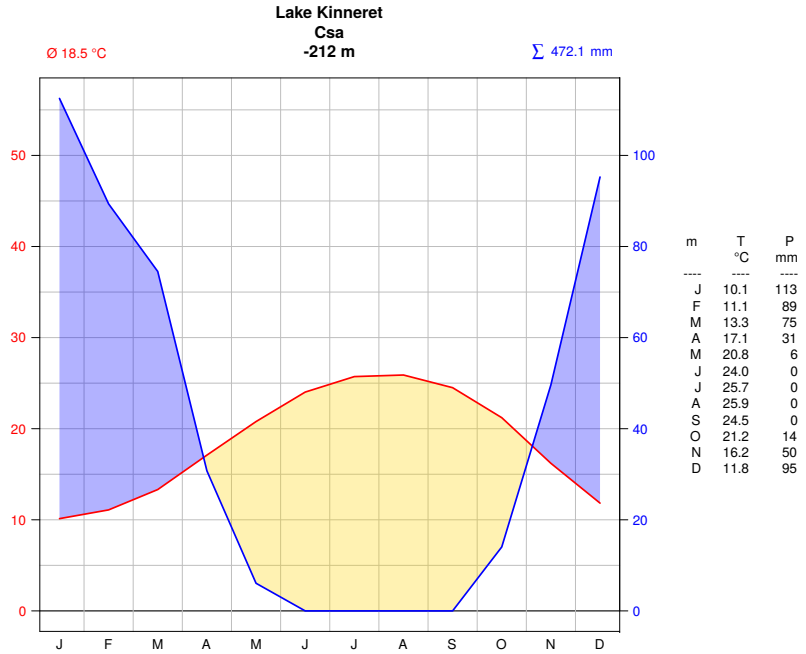


Figure 3.28: As Fig. 3.18, but for Lake Kinneret.

Mediterranean climate and borders the Irano-Turanian steppe (Miebach et al., 2019). This vegetation gradient indicates a relatively high climatic variability also during the Holocene. In the following sections we present such a quantitative climate reconstruction.

Introduction

In 2010 AD, a sediment composite core of 17, 8 m was recovered during a drilling campaign in the Sea of Galilee. The corresponding vegetational data are described in Schiebel and Litt, 2018 and covers approximately the last 9,000 years. Later, this palynological dataset was supplemented by additional samples to increase the temporal resolution (Thomas Litt, and Andrea Miebach, personal communucation). During the reconstruction process, the Koeppen-Geiger classification shown in Fig. 3.28, is used as an anchor for the most recent time slice. The problem of increasing human influence after about 6.5 cal ka BP is treated in the same way as for the Dead Sea in Sect. 3.3.1.

Several references were used for the reconstructions of the Dead Sea region. On the one hand, we include AP/NAP from the core and isotope data from Soreq Cave for a reconstruction during the Holocene. On the other hand, the AP/NAP curve was replaced by isotopic information from the Mediterranean Sea (Medstack) for the period from the Penultimate to the Last Glacial. Now all three proxy information are included, which are shown in Fig. 3.29.

The Bayesian age-depth model of this sediment core is presented in more detail in Miebach et al., 2022. It mainly includes ^{14}C measurements from terrestrial plants, sedimentary charcoal remains, and reservoir-corrected data from bulk sediments.

Reconstruction

In the following we describe the final reconstruction shown in Fig. 3.30. Panel (a) shows the reconstructed annual temperature T_{ANN} in $^{\circ}C$ and (b) the annual precipitation P_{ANN} in mm/a. At the beginning of the explanations we focus on the solid black lines. Later, we will work out the difference between these reconstructions and others whose medians are shown as blue lines. Schiebel and Litt, 2018 divide vegetation history into different ages, which is also used in this work.

The period **9 – 7 cal ka BP** can be associated mainly with the Pottery Neolithic (Schiebel, 2013). The vegetation is described in Schiebel and Litt, 2018 with a strong influence of steppe vegetation in the catchment area of Lake Kinneret. They conclude that this is due to increasing drought, which is confirmed by our reconstruction in Sect. 3.3.1, the lake levels of Lake Kinneret (Hazan et al., 2005) and the Dead Sea (Migowski et al., 2006). Furthermore, Miebach et al., 2022 infers a weak cooling trend and precipitation decrease during 7.8 – 7 cal ka BP from carbon isotope signals of the Sea of Galilee. These descriptions are consistent with our new climate reconstruction in Fig. 3.30. Panel (a) shows a slightly decreasing trend by 0.5 $^{\circ}C$ after 8 cal ka BP. This underlines the assumption that Medstack can be correlated with T_{ANN}. Although the temporal resolution of this proxy is quite low, a

corresponding increase can also be seen here, which is detected by the reconstruction algorithm. The average temperature in these years is about 18.5 °C. In (b) we see the lowest precipitation values of

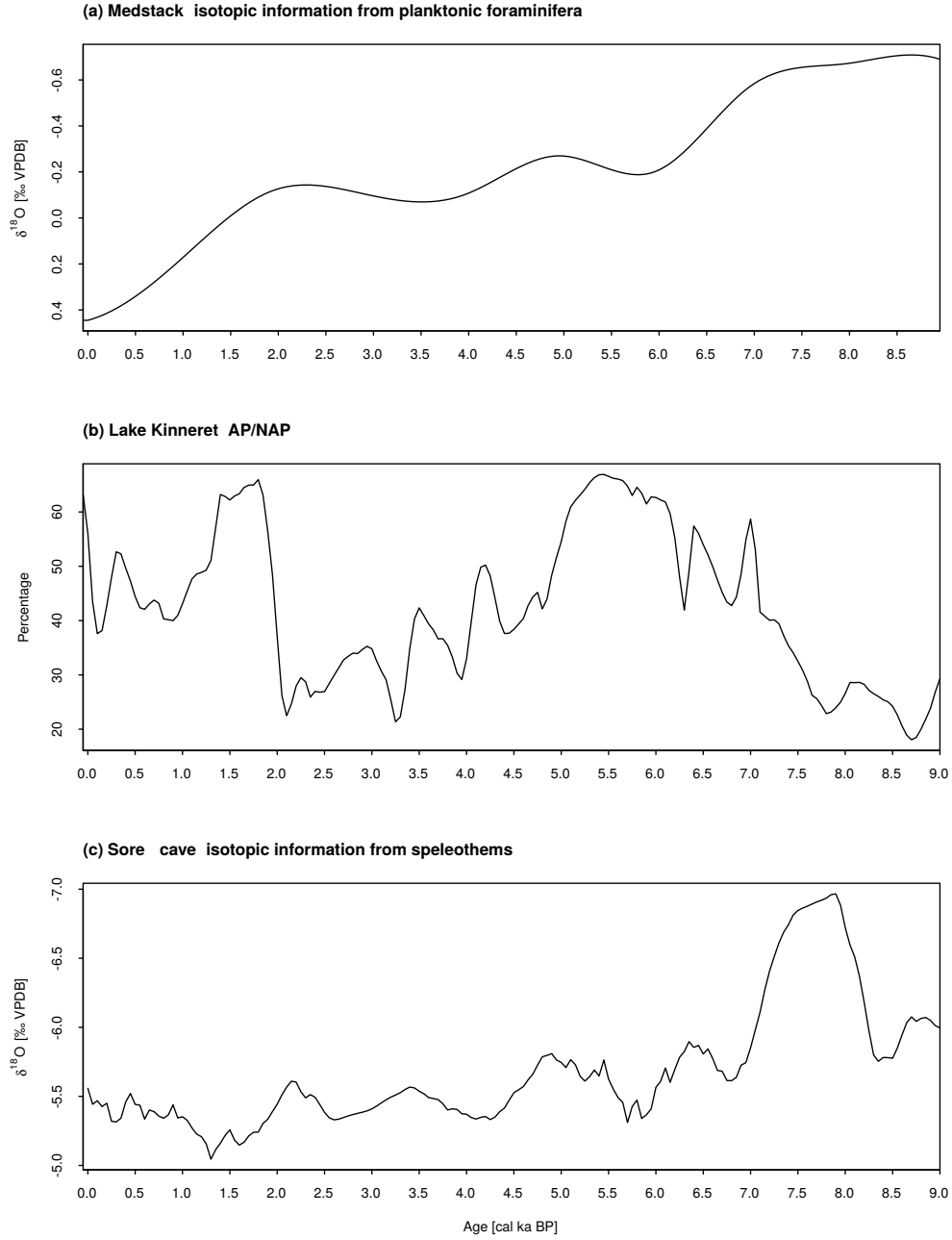


Figure 3.29: Reference curves based on (a) isotopic information from planktonic foraminifera from stacked marine cores of the Mediterranean Sea (Colleoni et al., 2012), (b) on total terrestrial pollen sums: arboreal and non-arboreal pollen taxa from Lake Kinneret, and (c) on isotopic data from speleothems from Soreq Cave (Bar-Matthews et al., 2003).

the entire reconstruction period with mean values around 400 mm and a local maximum around 8 cal ka BP. The approximate Koeppen-Geiger classification remains in the hot Mediterranean climate Csa during the last 9,000 years. Around 7 cal ka BP, an increase in P_{ANN} to values of about 550 mm is observed, with the annual temperature remaining relatively stable.

The beginning of the period **7 – 5 cal ka BP** is accompanied by an increase in Olea europaea. Schiebel and Litt, 2018 assume climate change towards more rainfall, which is also confirmed by the levels of Lake Kinneret and the Dead Sea. The isotopic data from Lake Kinneret also indicate an increase in precipitation and a stabilization of temperature during 7 – 6.6 cal ka BP. Likewise,

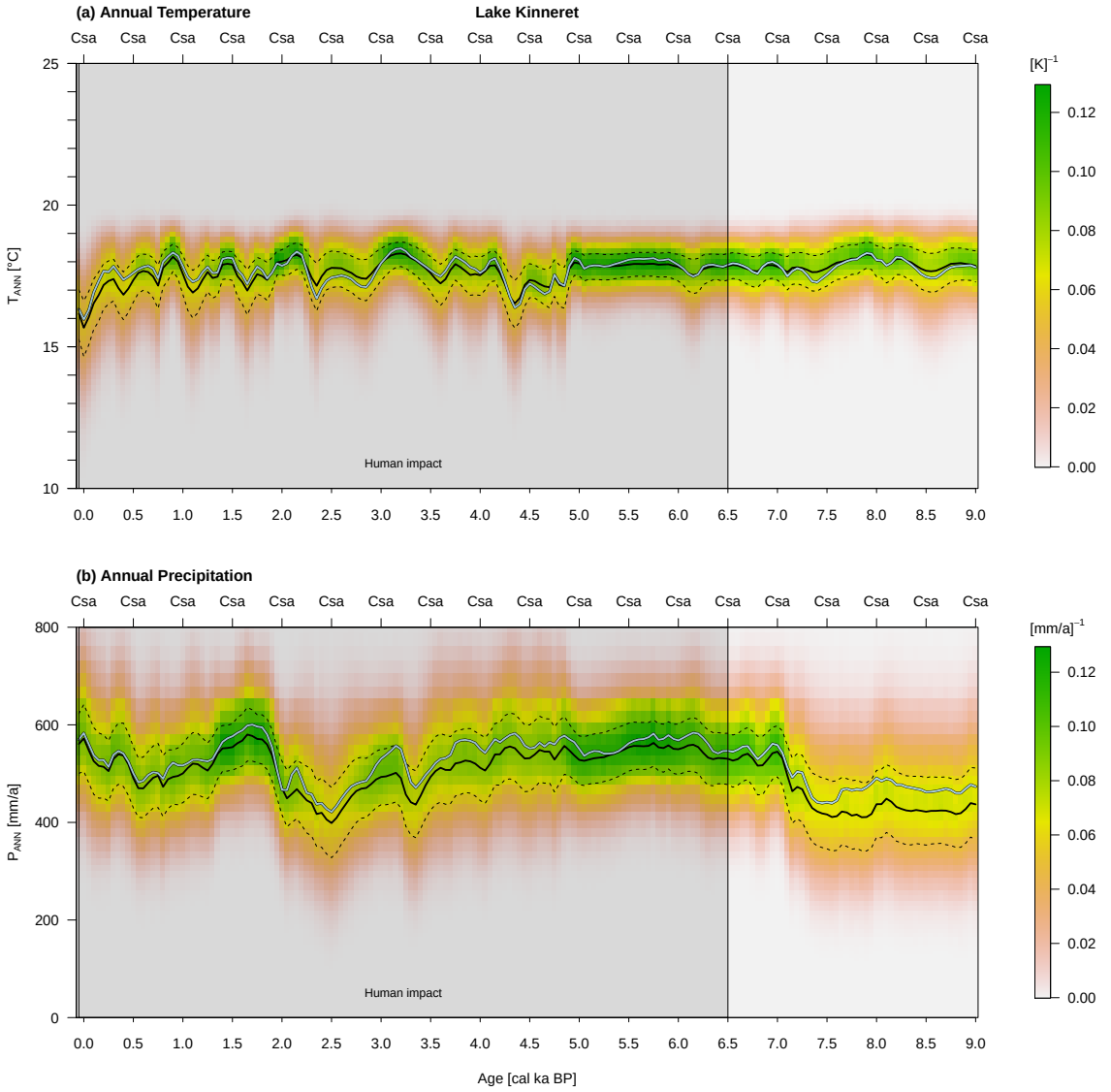


Figure 3.30: As Fig. 3.20, but for Lake Kinneret.

in our reconstruction from the Dead Sea in Sect. 3.3.1, a slight increase in P_{ANN} can be detected during 7 cal ka BP. In relation to the current reconstruction, we can see that these changes are also depicted. For both climate variables, this period shows the most stable conditions with precipitation about 550 mm and temperatures surrounding 18 °C. After the transition from the Pottery Neolithic to the Chalcolithic, ca. 6.5 cal ka BP, human influence is dealt with as described in Sect. 3.3.1. During the Chalcolithic, precipitation shows a local maximum, which decreases after about 5.5 ka BP. The

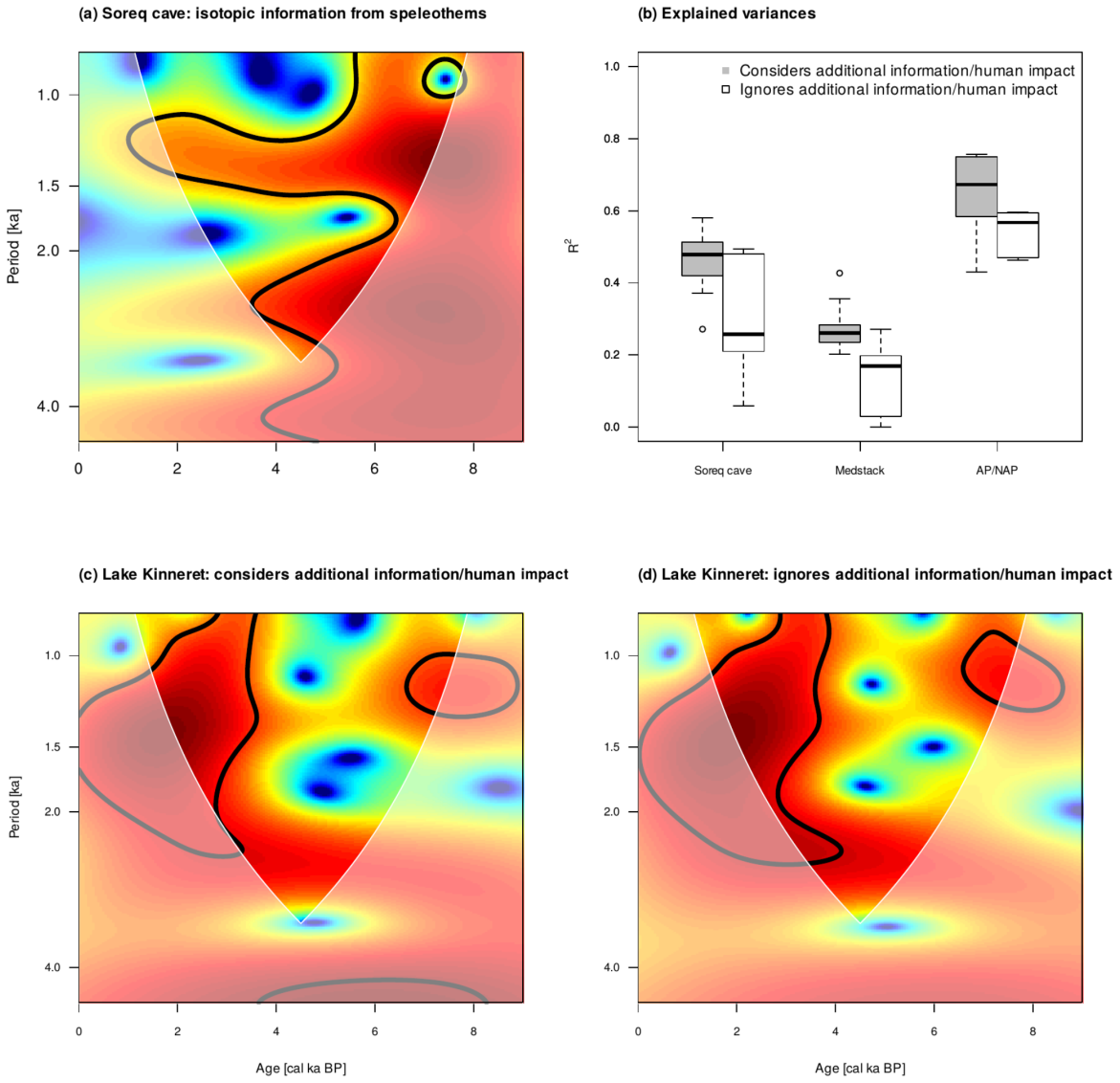


Figure 3.31: As Fig. 3.22, but for Lake Kinneret. Panel (b) is expanded by the explained variances between the medians of T_{ann} and Medstack. Moreover, R^2 in terms of AP/NAP is not divided based on human impact. Instead, only the results of the entire reconstruction period are presented.

second variation is not included in the AP/NAP intercept, but in the isotope data from the Soreq cave. Such behaviour is also described in the Dead Sea reconstruction and could be related to the transition from the Chalcolithic to the Early Bronze Age (Schiebel, 2013).

The Early Bronze Age to Iron Age within **5 – 2.3 cal ka BP** reflects not only human-induced but also climatically driven vegetation changes. On the one hand, Schiebel and Litt, 2018 describe the end of olive cultivation around 5 cal ka BP as a human influence, as this fluctuation does not match the high water levels of Lake Kinneret and the Dead Sea. On the other hand, the decrease in oak pollen by 4 and 3.3 cal ka BP could be related to the 4.2 and 3.2 drought events. Our climate reconstruction shows a slight increase in precipitation and a decrease in temperature from 18 °C to about 16.5 °C for the period 5 – 4 cal ka BP. Here, both isotopic references affect the reconstruction method, with Medstack and Soreq Cave showing an increasing trend leading to lower T_{ANN} and higher P_{ANN}. Accordingly, the climate change to lower precipitation and higher temperatures around 4 ka BP could be related to the transition from the Early to the Middle Bronze Age. The second and larger climate change during 3.3 ka BP might be related to the collapse of the Late Bronze Age (Schiebel and Litt, 2018). Due to the higher resolution of the core samples, we can detect further climate variations. The Iron Age in the Near East lasted from about 3.1 – 2.5 cal ka BP (Schiebel, 2013). This again corresponds to an increase in precipitation at the beginning and ends in a minmium with values around 400 mm. 2.5 cal ka BP marks the transition from the Iron Age to the Babylonian-Persian period, which lasted about 200 years and is accompanied by an increase in P_{ANN}.

The years from **2.3 – 0 cal ka BP** are marked by the Hellenistic and Roman-Byzantine and the Islamic periods until today. Orland et al., 2009 describe these ages, based on isotopic data from Soreq Cave, at the beginning with high precipitation values. At about 1.3 cal ka BP, they found a minmium in the deposit, which they associate with the transition from the Byzantine to the Islamic period. Our reconstruction coincides to a certain extent with these explanations. Precipitation increases up to about 1.75 cal ka BP and has a local minimum around 1.3 cal a BP. However, this is not as pronounced as in the Dead Sea reconstruction, and the reconstructed P_{ANN} of Lake Kinneret subsequently drops further to 0.5 cal ka BP. This shows how difficult it is to carry out a climate reconstruction based on vegetation data, which can be relatively easily manipulated by human activities. Finally, both reconstructed climate variables approach the most recent climate measurements for the last time interval.

Sensitivity study

Looking at the gray boxplots in Fig. 3.31 (b) we see, in contrast to the corresponding results from the Dead Sea (see Fig. 3.22 (b)), a relatively high explained variance with respect to AP/NAP. This suggests that the different proxies used for this reconstruction are more consistent with each other. Even if the additional information and human influence are ignored, R^2 has a similar median with respect to Medstack as in the other case. Nevertheless, the difference is greatest in Soreq Cave, as shown in Fig. 3.31 (c) and (d). The largest changes are found between 3 – 6 cal ka BP and periods

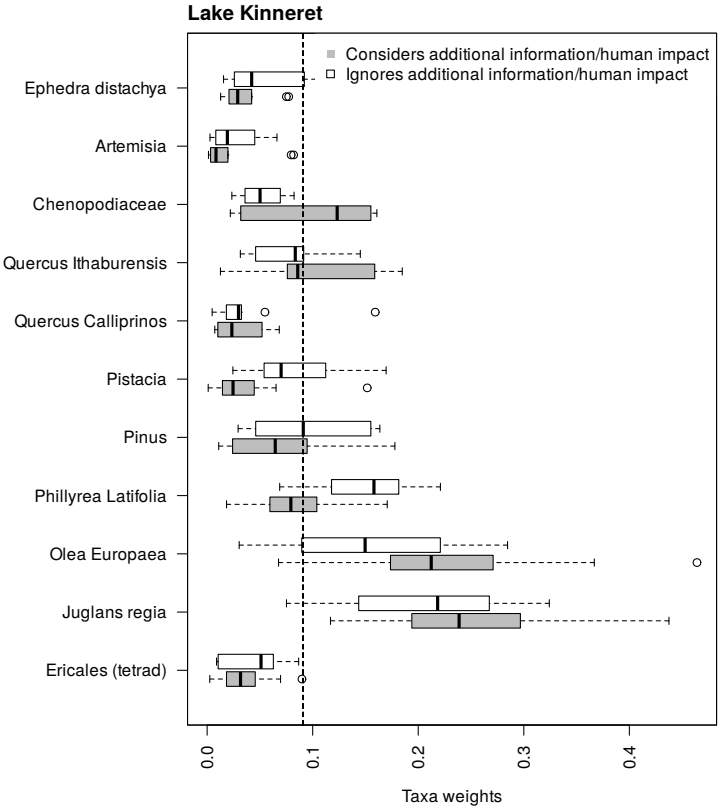


Figure 3.32: As Fig. 3.23, but for Lake Kinneret.

around 1 ka. When the reconstruction method has the information from the Soreq cave, it attempts to remove these higher frequencies and concentrate the branch around the 1.5 ka period (see Fig. 3.31 (a)).

In this context, higher weight is given to Chenopodiaceae (see Fig. 3.32). This has a WPS with relatively high power for such periods around 1.5 ka during these ages (not shown). Some of them are also characteristic members of the Irano-Turanian steppe territory (Schiebel and Litt, 2018) and occur mainly between 9 – 7 and 5 – 2.3 cal ka BP. During these periods, the blue curve in Fig. 3.30 (b) shows the largest differences with higher values compared to the black line. This is consistent with the relatively low precipitation values of these plants. No such differences can be found with respect to T_{ANN}. This is due to the relatively low variability in mean temperature values within the taxa pool used.

All in all, the differences are not as large as in the previous reconstructions. However, the algorithm again finds the best balance (in terms of R^2 and the climate anchor distributions) between the additional information from the proxies.

Summary and conclusion

The quantitative climate reconstruction of the Lake Kinneret region during the Holocene includes not only paleobotanical data, but also isotopic information from Soreq Cave and the Mediterranean Sea. In the period before 6.6 cal ka BP, the reconstructed climate variables reflect the qualitative climate reconstructions explained in Schiebel and Litt, 2018 and Miebach et al., 2022. Although the final reconstruction is compared only with isotopic information from speleothems and the Mediterranean Sea for ages younger than 6.5 cal ka BP, the results from P_{ANN} are comparable to the lake level of the Sea of Galilee (Hazan et al., 2005). Moreover, the algorithm is able to find climate changes that can be associated with Bond events and known archaeological and cultural changes in the Levant (Schiebel, 2013; Bar-Matthews and Ayalon, 2011; Orland et al., 2009).

Although these changes are not as pronounced as in the Dead Sea reconstruction and the comparatively small difference in terms of sensitivity analysis, we conclude that using the above information is beneficial. However, after 2.5 cal ka BP, human influence cannot be controlled as well as in the case of the Dead Sea. This could be addressed by expanding other proxies for this period, such as lake level or isotopic information from Lake Kinneret. In these cases, both information could be used as further reference curves.

3.3.3 Birkat Ram

Maar Lake Birkat Ram is located about 70 km north of Lake Kinneret in the northern Golan Heights. In contrast to the previously described sites in the Levant, the lake level is about 940 m amsl. Its main source of water is precipitation and local runoff (Schiebel, 2013). The climate diagram in Fig. 3.33 of the grid point closest to Birkat Ram shows a hot Mediterranean climate Csa. The mean annual temperature in this region is about 16°C, and the average annual precipitation is around 620 mm. Summers are dry and warm, while winters have a mild and humid climate. Most rainfall in this region occurs between October and April. Due to the higher elevation and mountainous orography around the lake, the annual temperature is lower and the annual precipitation is higher than at sites in the Jordan Valley. Thus, today Birkat Ram is located in the Mediterranean vegetation zone (Schiebel, 2013). How these present-day features may have changed over the past 30,000 years is discussed in the next sections.

Introduction

In 1999 AD and 2010 AD, two drilling campaigns returned 543 cm and 10, 96 m composite core, respectively. The corresponding vegetation data are described in Schwab et al., 2004 and in Schiebel, 2013 and cover approximately the last 6,500 and 30,0000 years. Schiebel, 2013 find a consistent correlation of pollen ratios as well as magnetic susceptibility signals of the two Birkat-Ram composites. For this reason, and because the shorter core has a higher sample resolution, we merge these two

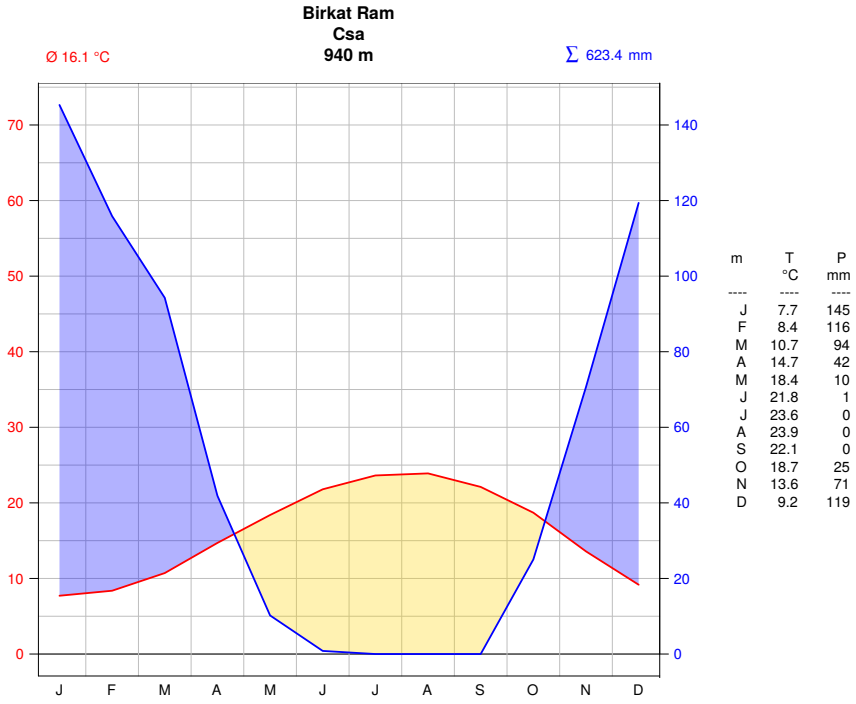


Figure 3.33: As Fig. 3.18, but for Birkat Ram.

palynological data sets. The magnetic susceptibility peaks at 516 cm (Schiebel, 2013) and at 527.7 cm (Schwab et al., 2004) are used to join both pollen spectra. That means, that for ages younger than ca. 6,250 cal a BP, the information of the shorter core are used. The reconstruction method uses the same references as for Lake Kinneret (see Sect. 3.3.2) and are mapped in Fig. 3.34.

According to the linkage of the two sediment cores, the Bayesian age-depth model uses the respective datings. This has already been done in Schiebel, 2013, where the basis of our model is described. Compared to previous age-depth models, the result shows partially discontinuous sedimentation of about 7,000 years between about 10 – 17 cal ka BP (Schiebel, 2013). This zone of low sedimentation is shown in Fig. 3.34 (b) and results in a relatively small deviation of the AP/NAP curve within this period.

Reconstruction

In the following we describe the final reconstruction shown in Fig. 3.35. Panel (a) shows the reconstructed annual temperature T_{ANN} in °C and (b) the annual precipitation P_{ANN} in mm/a. Schiebel, 2013 divide vegetation history into different ages, which is also used in this work.

During the period **30 – 18 cal ka BP** that encompasses the LGM, Schiebel, 2013 describe a steppe-like character in the vegetation composition. They argue that human influence is much less than during the Holocene and conclude that the climate is relatively cold and dry. However, as mentioned

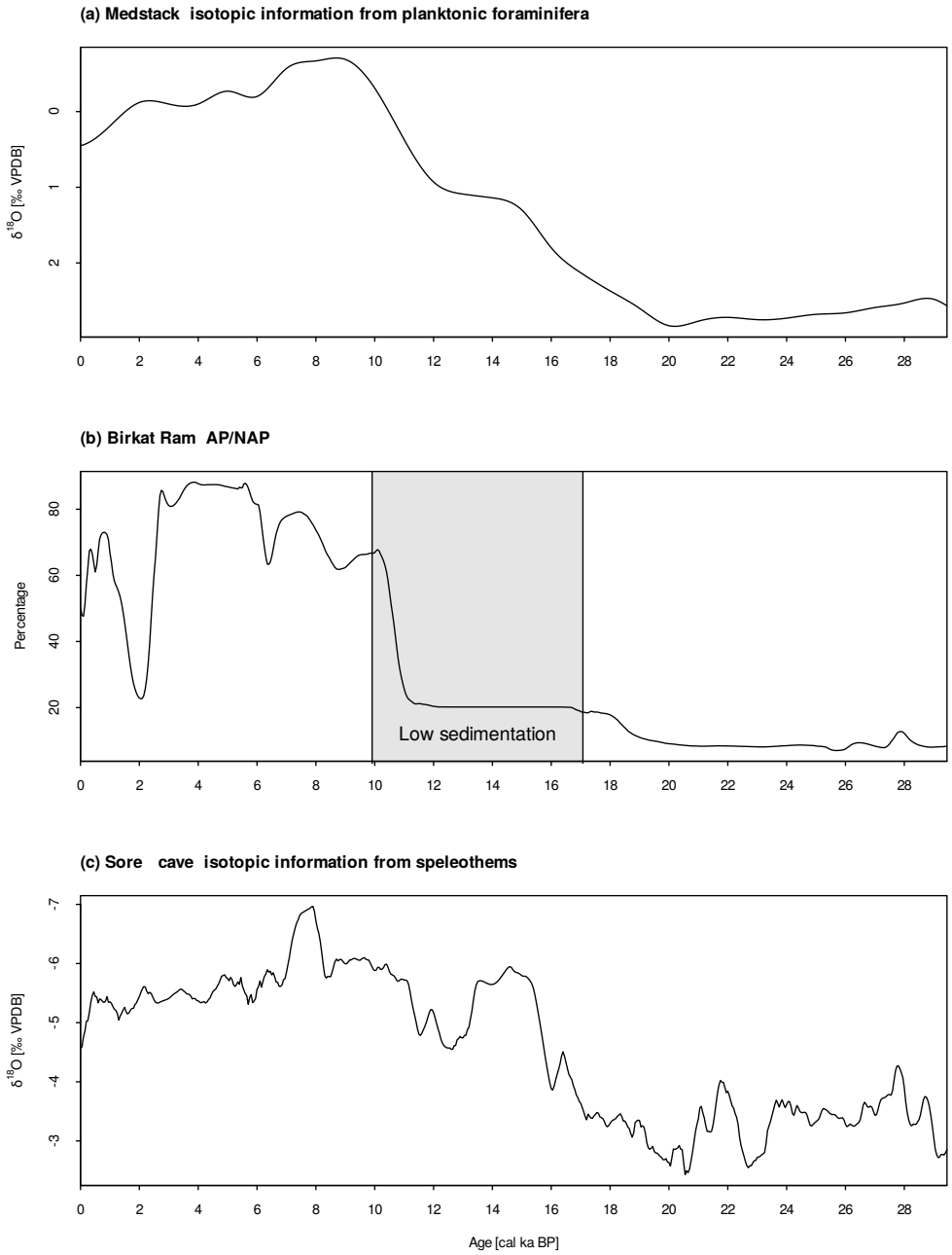


Figure 3.34: As Fig. 3.29, but for Birkat Ram. The grayed out area in (b) marks the zone with low sedimentation rates.

in Sect. 3.3.1, the low evaporation rate could lead to high water levels in Lake Lisan (Schiebel, 2013). Our reconstruction gives an average T_{ANN} of about 8 °C and P_{ANN} of 450 mm. The combination of both variables leads to an approximated Koeppen-Geiger climate of Dfa, indicating a fully humid continental climate with hot summers. For comparison, some central parts of the prairie plains in

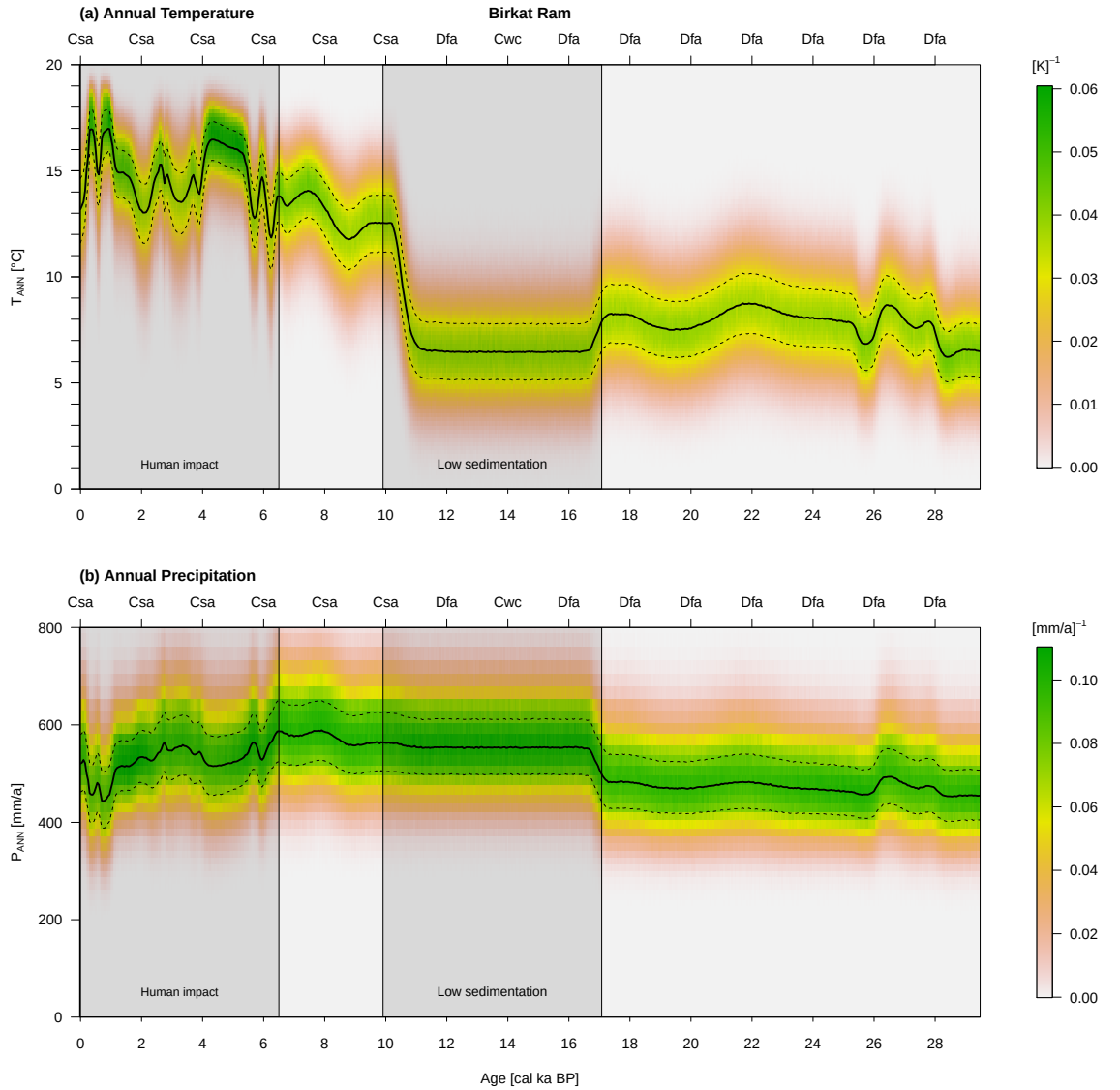


Figure 3.35: As Fig. 3.20, but for Birkat Ram. The grayed out areas mark the zone with low sedimentation rates and human influence. Since no sensitivity study is performed this time, no blue lines are shown.

the USA today have a similar climate and elevation that was dominated by the natural vegetation of the grass-prairie mixture (Weaver, 1965). The corresponding Dead Sea climate in Fig. 3.25 shows warmer temperatures but similar precipitation for this period. Between 30 – 25 cal ka BP, some fluctuations are observed in both variables. In Fig. 3.36 (b) the WPS shows a high significant power with a period of about 1.5 ka. This is consistent with Dansgaard-Oeschgar events 3 and 4 occurring between Heinrich stages 3 and 2 (cf. Fig. 2.3). In the WPS of Soreq Cave we see corresponding periods for ages older than 28 cal ka BP. Although the isotopic information from Greenland is not

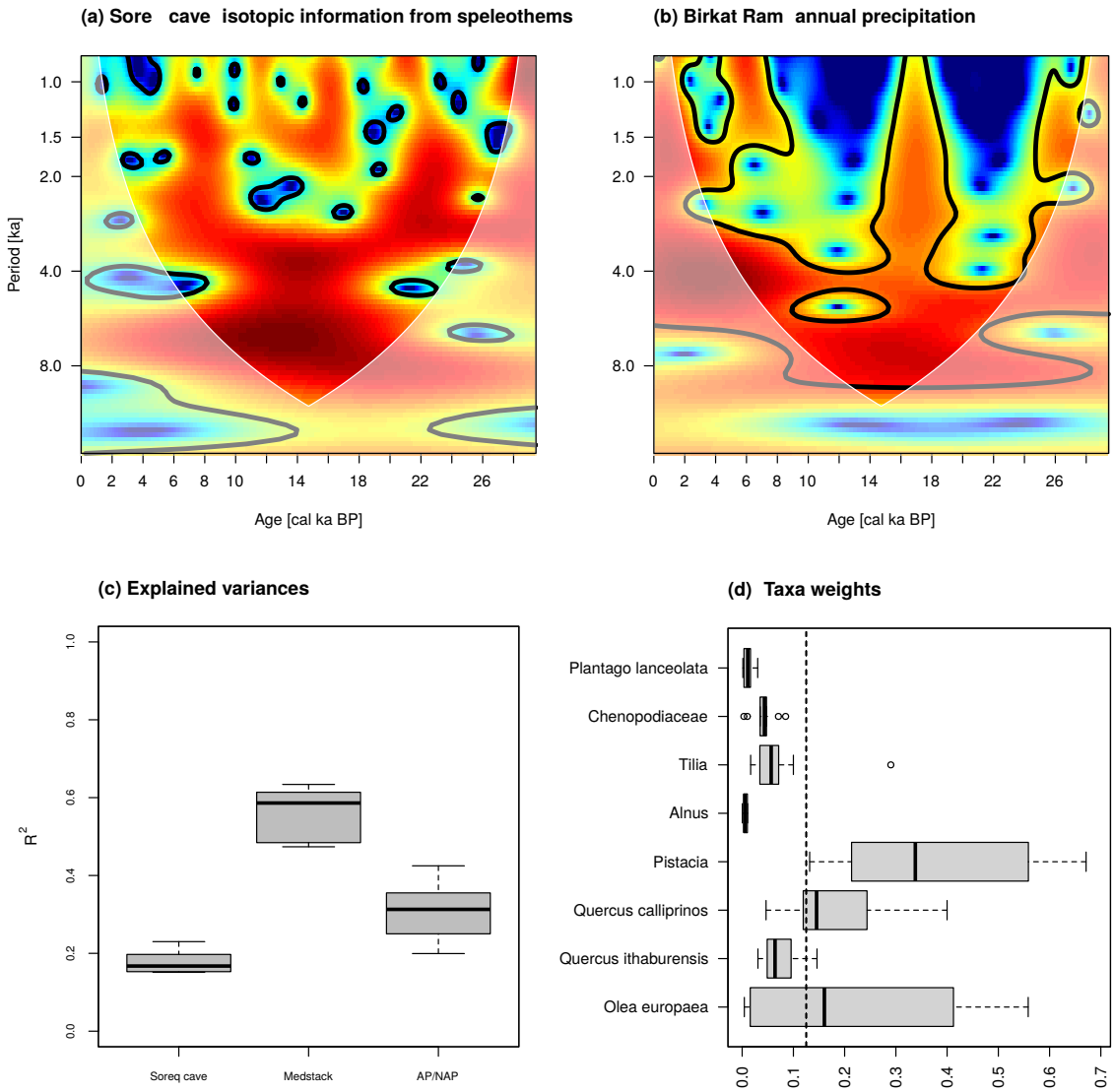


Figure 3.36: Wavelet power spectrum of (a) isotopic information from speleothems from Soreq Cave (Bar-Matthews et al., 2003) and (b) the reconstructed annual precipitation. Panel (c) shows the explained variances between the selected reconstructed medians from P_{ANN} and the patterns from Soreq Cave and the AP/NAP reference curve. The corresponding medians from T_{ANN} are compared with isotopic information from Medstack. In (d), the posterior taxa weights of the selected reconstructions are shown as boxplots and the prior weight is represented by the dashed line.

known to the algorithm, the final result is closer to it than to the data from Soreq Cave. It seems that there are corresponding features in the palynological data, as the coarse resolution of Medstack cannot capture such high-frequency variations.

Between **17 – 12.9 cal ka BP** the lake level was so low that Schiebel, 2013 suggest a possible

desiccation of Birkat Ram during the Deglaciation period. They noted an increased amount of precipitation during DO 1, possibly superimposed by increased evaporation/temperature. This could explain the discrepancy between the low lake level and the higher precipitation. Due to the vegetation gap during the Late Glacial, we cannot draw any conclusions about the reconstructed climate within the grayed-out area.

With the beginning of the Holocene in the period **11.7 – 6.5 cal ka BP**, Schiebel, 2013 describe the strongest change in pollen composition. They note a decline in Irano-Turanian plant populations and an increase in Mediterranean vegetation. From this they conclude an increase in temperature and precipitation. Around 8.7 cal ka BP, a decline in Mediterranean vegetation is noted, which they associate with the 8.2 event, accompanied by colder and drier climatic conditions. In contrast to the Dead Sea, a rapid recovery of the AP is observed after this event, as shown in Fig. 3.34 (b). Our reconstruction fits these explanations quite well. At the beginning of this period, temperatures are around 13 °C and precipitation is about 550 mm. The Koeppen-Geiger classification transitions from Dfa to the hot Mediterranean climate Csa, which remains in this category throughout the Holocene. Thus, during the 8.2 event, T_{ANN} dropped by only 1 °C, and a decrease in precipitation is hardly noticeable. After that, both variables increase to 14 °C and 600 mm. A look at the WPS in Fig. 3.36 (a) shows that Soreq Cave has significant periods of about 1.5 ka during the Holocene. With the onset of the 8.2 event, a similar pattern occurs in the WPS of P_{ANN}, which can be related with some Bond events. Compared to Lake Kinneret and the Dead Sea, it is difficult to determine what fluctuations might be associated with corresponding events.

As shown in the previous reconstructions of the Levant, the period **6.5 – 0 cal ka BP** is treated with a non-negligible human impact on the vegetation (Schiebel, 2013). It is noticeable that the reconstructed precipitation no longer follows the AP/NAP reference. The sharp decline of AP/NAP due to the deforestation of Quercus ithaburensis during the Hellenistic period (Schiebel, 2013) after ca. 2.3 cal ka BP is therefore not reconstructed in P_{ANN}. Instead, the focus is on the pattern from Soreq Cave and the curve from Medstack, which can be seen in the explained variances in Fig. 3.36 (c). The reconstruction algorithm is able to find higher values in terms of AP/NAP (not shown) than the specified range between $0.2 - 0.4$. In contrast, the explained variances of Medstack are higher and those of Soreq Cave are lower. All in all, the best combination (Pareto optimum) between these variables and the climate anchor points emerges again. The core thus contains enough variability to generate such explained variances. To achieve this, the algorithm weights the oaks, olives, and pistachios most heavily (see Fig. 3.36 (d)), with the exception of Quercus ithaburensis, which is the main component of the AP ratio (Schiebel, 2013). The decrease in annual precipitation and the simultaneous increase in temperature between $1.3 - 0.5$ cal ka BP is due to the reoccupation of vacant areas by evergreen oaks and pistachios during the Islamic period (Schiebel, 2013). Finally, the reconstructed variables approach the most recent climate anchor point (Csa) for the most recent time slice.

Comparison to older climate reconstructions

In Neumann et al., 2007 a reconstruction of the last ca. 6,500 years based on the older, higher resolution core is presented. In contrast, Thoma, 2017 presents a reconstruction of the last 30,000 years based on the newer, coarser resolution core. The techniques used in both studies are slightly different. While Neumann et al., 2007 use the indicator taxa method, Thoma, 2017 extend it with additional biome information.

The two biomes used in the latter case are the Mediterranean and the Irano-Turanian. On the one hand, the main difference between them in terms of annual precipitation is that the Mediterranean biome has higher values. The Irano-Turanian biome, on the other hand, has lower winter and higher summer temperatures. The biome ratio in Thoma, 2017 resembles the AP/NAP curve, with AP reflecting the Mediterranean biome. This leads to an overall reconstruction that also takes into account human-induced changes during the Holocene. In addition, the summer temperature during the Last Glacial is about 28 °C and the winter temperature is around 10 °C. This is in contrast to the remarks of Schiebel, 2013. In this context, the addition of another biome, such as the Irano-Turanian steppe forest (Litt et al., 2021), could help reconstruct more reasonable values during glacial periods.

In contrast, Neumann et al., 2007 present a reconstruction that shows values for the youngest time slice that are more in line with those in Fig. 3.33. In terms of P_{ANN}, it is similar to our reconstruction, with an overall slight decrease over the last 6,500 years, but with smaller fluctuations than in the reconstructed temperatures. Except for the above-mentioned variation in our reconstruction during 1.3 − 0.5 cal ka BP, the variables in Neumann et al., 2007 show greater similarity to the variations in the AP/NAP curve derived from Quercus ithaburensis. In our case, this taxon has on average a lower posterior weight compared to the prior weights.

Summary and conclusion

The quantitative climate reconstruction of the Birkat Ram region during the last ca. 30,000 years includes not only paleobotanical data, but also isotopic information from the Soreq Cave and the Mediterranean Sea. In summary, only 8 taxa information contain enough variation that a trade-off between these 3 proxies and 2 climate anchor points can be found. Furthermore, the final result resembles the qualitative reconstruction described in Schiebel, 2013. Thus, the climate classification assumes a Mediterranean climate throughout the Holocene and a continental climate during the Last Glacial.

Furthermore, in the period after 6.5 cal ka BP, most human-induced vegetation changes are not included in the reconstruction. Compared to old quantitative reconstructions, our new algorithm is better able to avoid anthropogenic fluctuations in palaeo-vegetation and to reconstruct a more appropriate climate during the Last Glacial.

3.3.4 Lake Van

Lake Van is located in the East Anatolian High Plateau in a tectonic depression (Litt et al., 2014). It is the fourth largest terminal lake on earth and is surrounded by the Bitlis massif, which reaches an altitude of more than 3500 m amsl (Pickarski et al., 2015b). The lake itself is at an elevation of about 1640 m amsl. The climate diagram in Fig. 3.37 of the grid point closest to Lake Van shows a Mediterranean-influenced continental climate (Dsb) with warm and dry summers. The average annual temperature in this region is about 8.5 °C, and the average annual precipitation is around 550 mm. The highest rainfall occurs in winter and spring and is strongly influenced by the orography. Thus, two vegetation types are found in the vicinity of the site: the Kurdo-Zagrosian oak steppe-forest and the Irano-Turanian steppe (Pickarski et al., 2015a). How these present-day features may have changed over the past 130,000 years is discussed in the next sections.

Introduction

Two drilling campaigns in 2004 and 2009 recovered a single core (VAN04-2) and a composite core from the Ahlat Ridge. The first is described in (Litt et al., 2009) and covers a continuous sequence from the LGM to the present. Part of the chronology is based on varve information. The second covers the last 600,000 years (Litt et al., 2014) and the period from about 10 – 130 cal ka BP is described in Pickarski et al., 2015b and Pickarski et al., 2015a. The corresponding chronology is based on, among

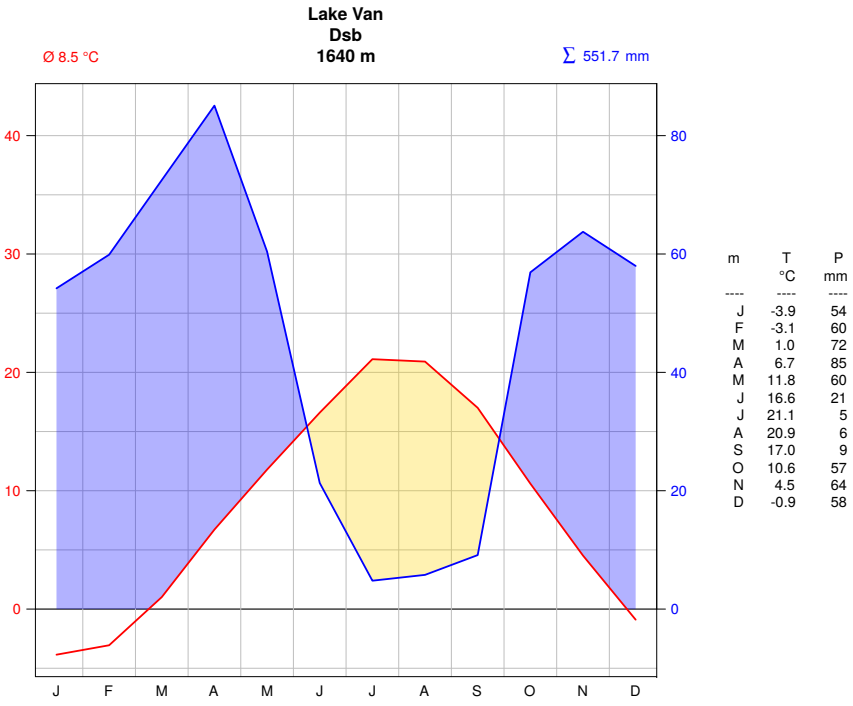


Figure 3.37: As Fig. 3.18, but for Lake Van.

other things, the same varve information as for core VAN04-2 (Stockhecke et al., 2014), so we use the same age-depth model for both cores. For this reason, and because of the proximity of the two sampling sites to each other (Litt et al., 2009), we summarize the corresponding vegetation spectra as for Birkat Ram in Sect. 3.3.3. The linkage point is at about 20 cal ka BP. The result of the combined AP/NAP reference curve in Fig. 3.38 (b) illustrates this procedure. We see a local maximum around 13 cal ka BP and a subsequent local minimum at about 12 cal ka BP. This is consistent with the results presented in Litt et al., 2009. In addition, we note a sharp decrease in percentages by 3.8 cal ka BP where Litt et al., 2009 mention an increase in human activity. For this reconstruction, the information on Pinus was removed from the AP summary. Pickarski et al., 2015b state that this taxon is likely Pinus nigra, and Pickarski et al., 2015a describe its distribution as an indicator of a colder/drier

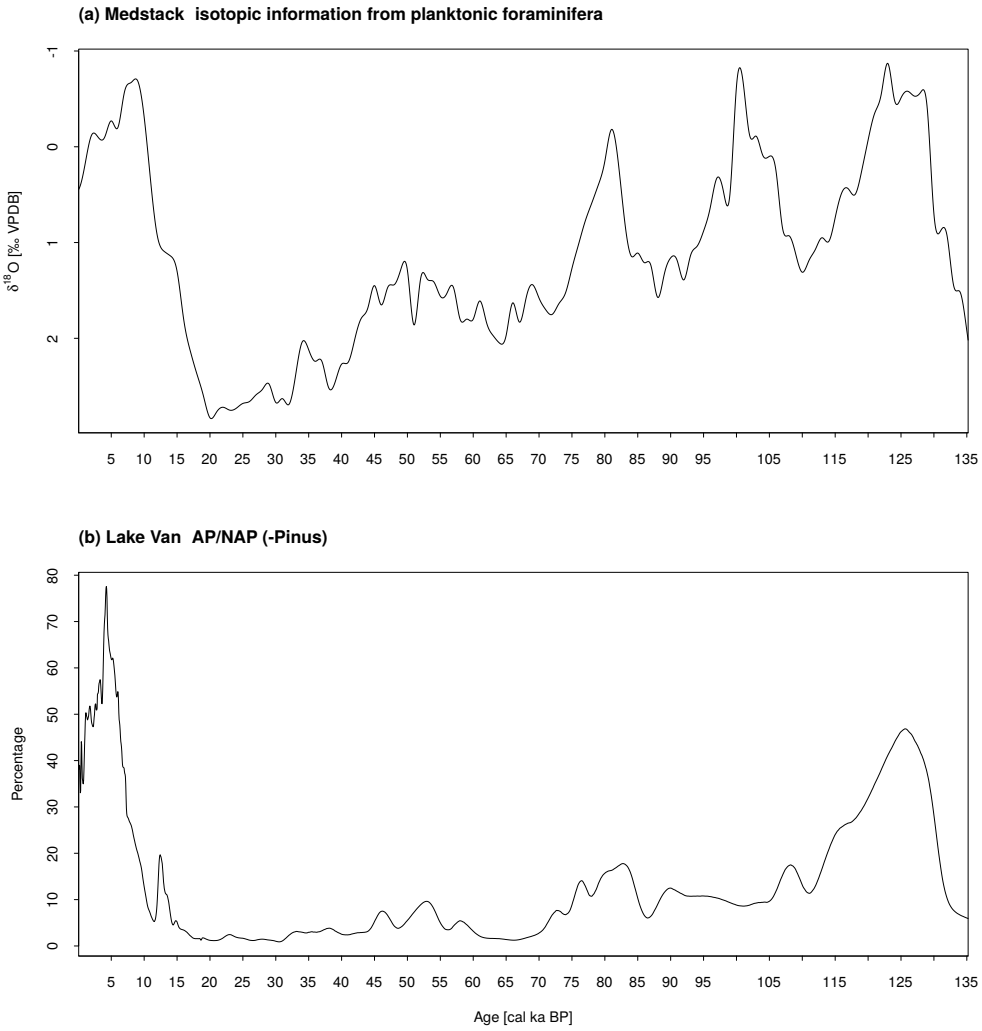


Figure 3.38: As Fig. 3.29, but for Lake Van. Pinus is removed by calculating the percentages of arboreal pollen.

environment. Therefore, the main components of the AP are the temperate taxon deciduous oak, the moisture-demanding thermophilic taxa such as Ulmus and Carpinus betulus, and the frost-sensitive Pistacia (Pickarski et al., 2015b). They also describe each interstadial as a "warm/wet" phase that could be associated with our AP. So we compare this curve not only for annual precipitation as in the Levant, but also for annual temperature.

The relationship between age and depth used in this work is described in more detail in Stockhecke et al., 2014. It is based not only on varve counts, but also on radiocarbon dates and tephra layers. Furthermore, they derive age control points from a correlation of Lake Van paleoclimate records with δ^{18}O ice core records from NGRIP and GLT_{syn} SpeleoAge. The latter is a speleothem-based synthetic Greenland ice-core record. For the last 20,000 years, the reported age errors are relatively small and increase thereafter. If no errors are specified, we use the mean error of the specified uncertainties around these ages. All data up to the first gap around 260 cal ka BP are inserted into the Bacon model. A result of this age-depth transformation can be seen in Fig. 3.38. (b), where we find a much higher age resolution for years younger than 20 cal ka BP, leading to higher frequencies in the time series.

Reconstruction

In the following, we describe the final reconstruction shown in Fig. 3.39. Panel (a) shows the reconstructed annual temperature T_{ANN} in °C and (b) the annual precipitation P_{ANN} in mm/a. Litt et al., 2009, Pickarski et al., 2015b, and Pickarski et al., 2015a divide the corresponding vegetation history into different periods, which are also used in this work.

Pickarski et al., 2015a describe the late Penultimate Glacial from **133.9 – 131.2 cal ka BP** with a dominance of NAP. They conclude cold and dry climatic conditions, which are also reflected in our reconstruction. The average T_{ANN} of about 6.5 °C and P_{ANN} of ca. 400 mm lead to a Koeppen-Geiger classification of continental climate Dsb. At the end of this period, according to the reference curves in Fig. 3.38 (a) and (b), a climate change toward warmer and wetter conditions can be observed.

The pre-temperate phase from **131.2 – 129.1 cal ka BP** shows a decrease in NAP with patches of temperate trees (Pickarski et al., 2015a). This indicates a generally warmer environment with increased precipitation, which is also reconstructed in our study. We see a temperature increase of about 2 °C and a precipitation increase of ca. 200 mm. The corresponding climate classification transitions to a Mediterranean climate with warm summers (Csb) that persisted during the Last Interglacial.

The LIG plateau, which lasted from **129.1 – 115.6 cal ka BP**, is described by Pickarski et al., 2015a as having a distribution of temperate trees and is divided into three vegetation zones. The Quercus-Ulmus zone 129.1 – 127.2 cal ka BP is characterized by a warm-temperate environment with enhanced precipitation. Here our reconstruction shows the highest annual precipitation and warmest annual temperatures during the LIG. The Carpinus zone of 127.2 – 124.1 cal ka BP is described with a slight decrease in temperature but constant humidity. This slight decrease is reconstructed for both variables. During the Pinus zone 124.1 – 115.6 cal ka BP, a spread of drought-adapted and/or

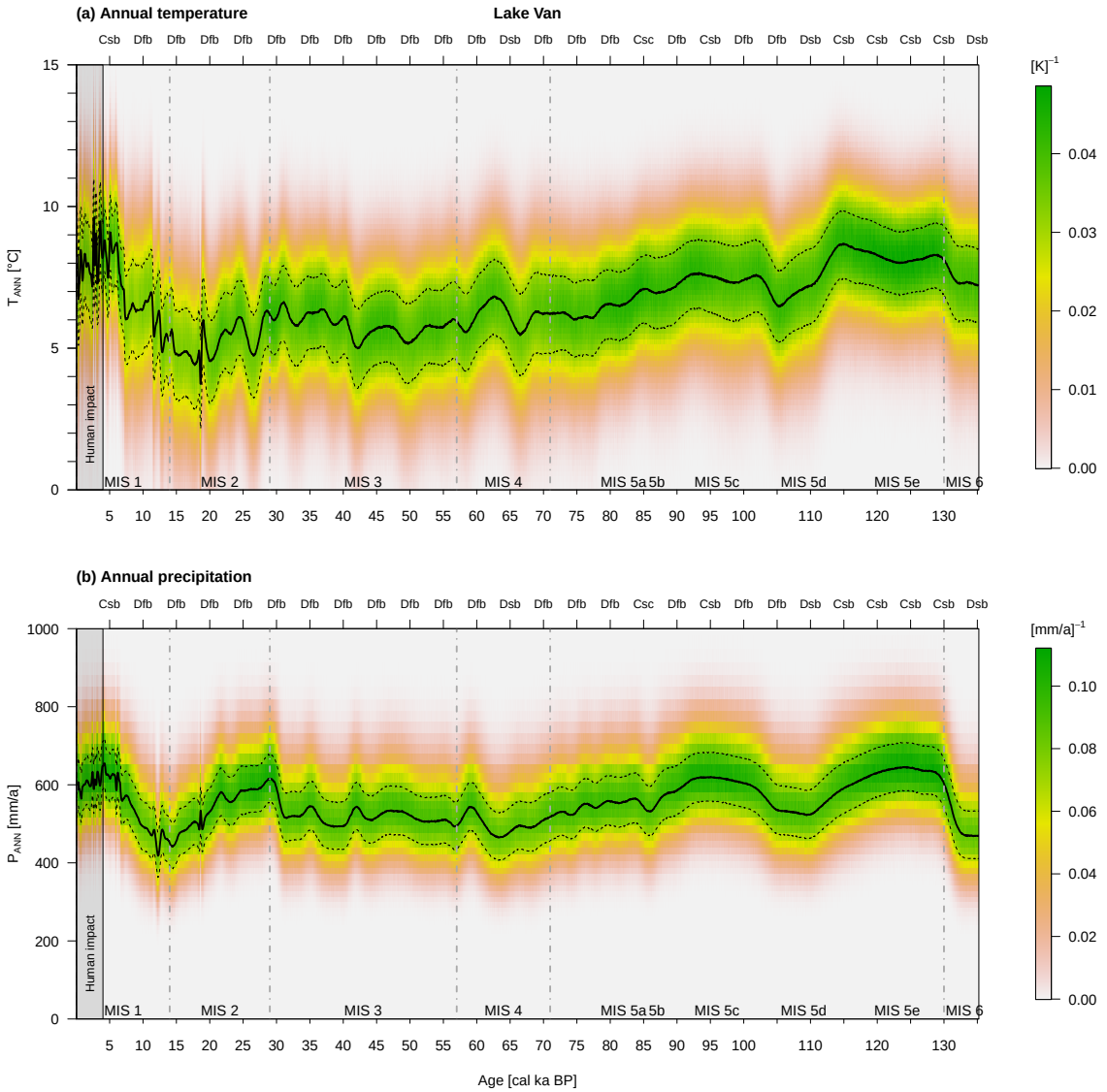


Figure 3.39: As Fig. 3.20, but for Lake Van.

cold-tolerant Pinus is observed, suggesting increasing aridity and cooler temperatures. There is a corresponding successive decline in P_{ANN}, with annual temperature increasing again, but not reaching the level of the Quercus-Ulmus zone. Although both reference curves also describe a decrease in T_{ANN}, the relatively high posterior weight for the warm-temperate Quercus deciduous shown in Fig. 3.40 (b), which still shows comparatively high abundance during this period, could lead to these temperatures.

The post-temperate phase from **115.6 – 111.5 cal ka BP** shows an expansion of NAP, which Pickarski et al., 2015a describe as a quasi-stadial state marking the end of the LIG at about 111.5 cal

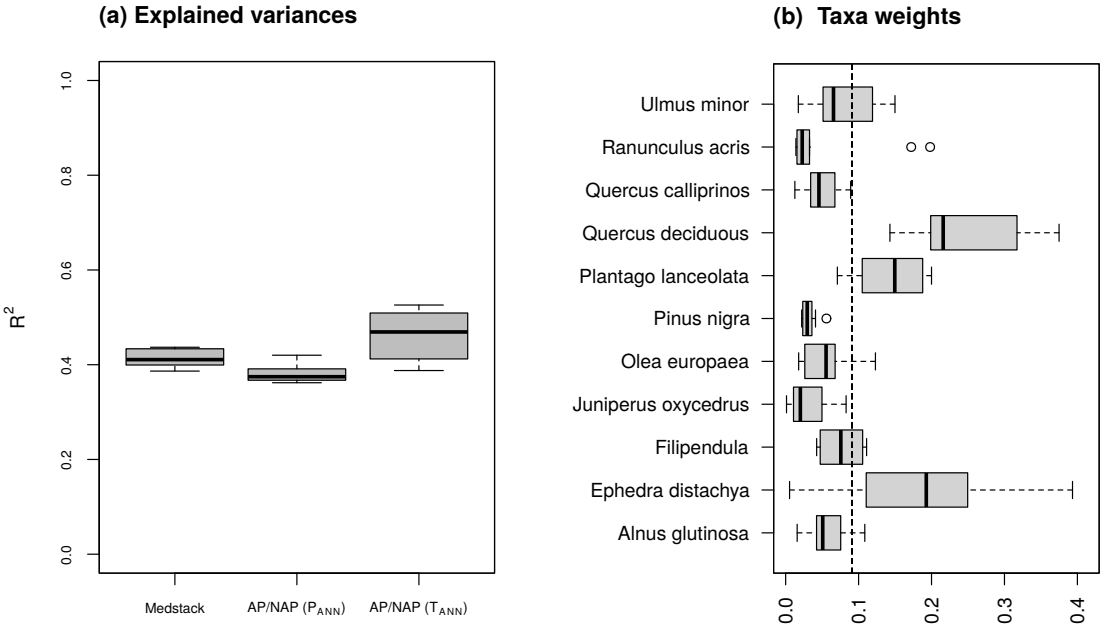


Figure 3.40: Panel (a) shows the explained variance of the posterior samples compared to the Medstack and AP/NAP reference curves. The latter is divided into T_{ANN} and P_{ANN}. Panel (b) depicts the posterior taxa weights.

ka BP. Both the reconstructed temperatures and precipitation are consistent with these explanations, as they each show a decrease. The Koeppen-Geiger classification transitions to the continental climate Dsb.

The following period of **111.5 – 107.8 cal ka BP** could be associated with MIS 5d, also known as the Herning Stadial. Pickarski et al., 2015b describe a significant spread of steppic herbaceous plants and conclude cold and/or dry climatic conditions with pronounced seasonality of precipitation during the stage. This is supported by a reconstructed summer-dry continental climate with annual temperatures around 6.5 °C and annual precipitation of about 420 mm. The relatively wide climatic range of Dsb becomes clear here when we compare the difference of the reconstructed to the current climate values from Fig. 3.37. If we look at the corresponding WPSs in Fig. 3.41, we see a comparatively high power with a period of about 23 ka, lasting at least from 30 – 100 cal ka BP. This could be related to orbital precession, which has a similar period on average and primarily controls the intensity of summer solar radiation (Huybers, 2006). Moreover, in the period of at least 50 – 80 cal ka BP, we see another branch of relatively high power with a period around 41 ka, which could indicate another Milankovic cycle, namely obliquity. Bosmans et al., 2015 mention that this variation also controls solar radiation in summer. So both cycles are connected to this forcing. In Pickarski et al., 2015b, summer solar radiation is used as an additional climate proxy, which often coincides with the vegetation variations

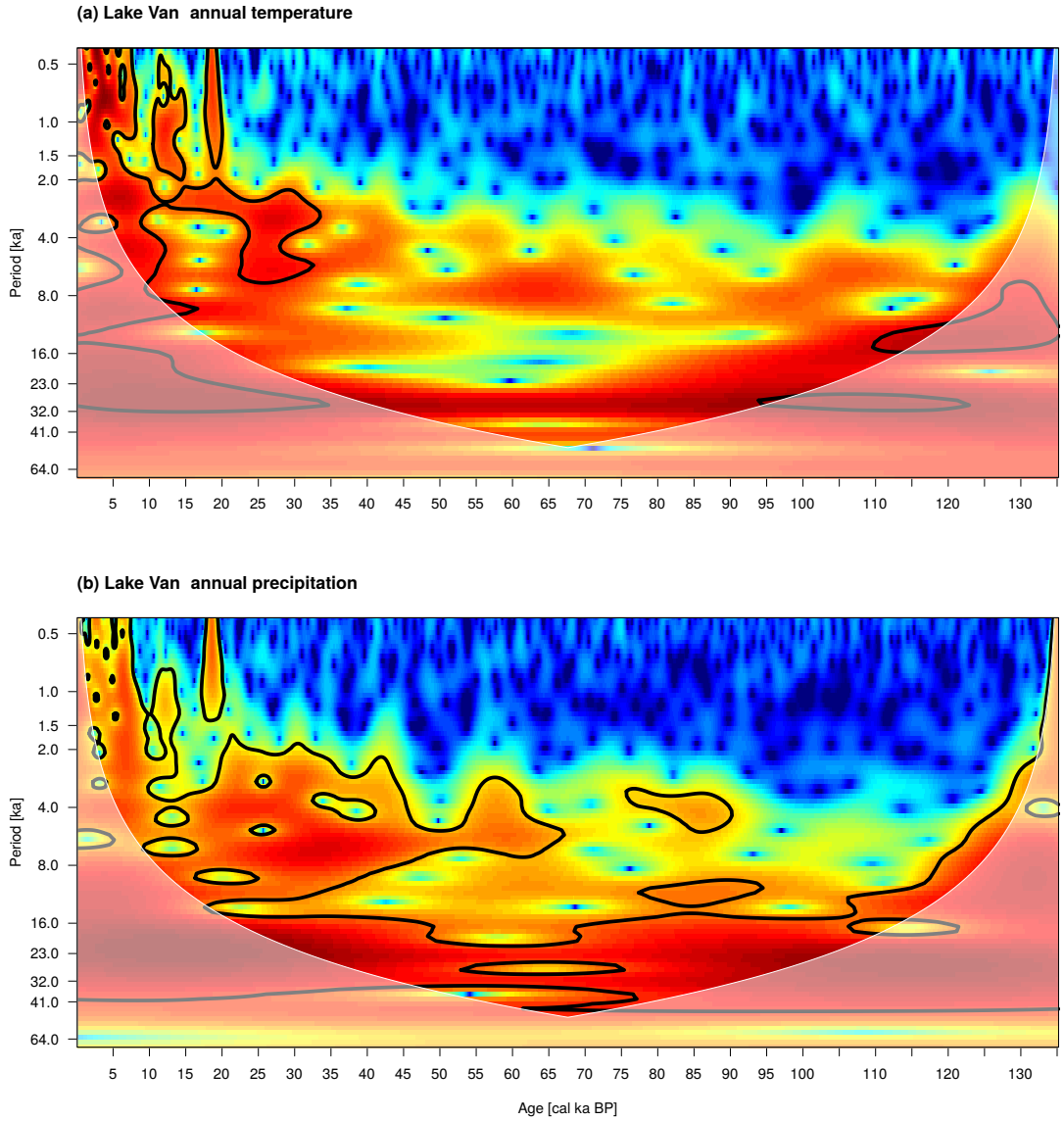


Figure 3.41: Wavelet power spectrum of (a) the median of the annual temperature and (b) the median of the annual precipitation.

of Lake Van. Although no intra-annual climate information is known to the algorithm, we see patterns in this spectral analysis that confirm these descriptions.

After the Herning stadial, the so-called Brørup interstadial (MIS 5c) lasted from ca. **107.8 – 87.3 cal ka BP**. Pickarski et al., 2015b describe it with a dominance of cold and/or summer-dry adapted Pinus. During this period, the reconstruction shows an increase in both variables. Nevertheless, the level of MIS 5e is not reached again, which could be explained by the lower proportion of Quercus deciduous during the Last Glacial. Accordingly, the Koeppen-Geiger classification alternates between

Csb and Dfb. The latter describes a continental climate without a dry season.

The Rederstall stadial of **87.3 – 84.9 cal ka BP** (MIS 5b) behaves like the Herning stadial (Pickarski et al., 2015b) and is accompanied by a reconstructed decline of T_{ANN} and P_{ANN}. The latter shows a local minimum at the end of this period, but the temperature continues to drop. A Mediterranean climate with cold summers (Csc) appears.

MIS 5a, also known as the Odderade interstadial, lasted from about **84.9 – 77.5 cal ka BP** and is accompanied by a slight increase in AP (Pickarski et al., 2015b). They infer an increased moisture supply at the beginning and a decrease in precipitation and/or an increase in evaporation during this period. This interstadial is not as clearly visible in the reconstruction as MIS 5c. Nevertheless, the reconstructed precipitation increases slightly at the beginning of this interstadial, leading to the continental climate Dfb.

Pickarski et al., 2015b describe the period **75 – 28 cal ka BP** with a dominance of Artemisia, Chenopodiaceae and Poaceae. They mention a strong aridification and cooling in Eastern Anatolia after 70 cal ka BP. Our reconstruction shows a relatively stable level in both variables, namely a moderate variation around 6.5 °C and an averaged precipitation of about 450 mm. Although the climate reveals colder temperatures and less precipitation than today, the Koeppen-Geiger estimate remains within Dfb. Pickarski et al., 2015b mention a high variability of the tree population during MIS 3. They associate these alternations of warmer/moist interstadials and cooler/drier stadials with some Dansgaard-Oeschger cycles. Due to the relatively large uncertainty taken into account when transforming depth into age within this period, these fluctuations cannot be detected. This can be verified in Fig. 3.41, where no relatively high power is shown for the typical DO period of about 1.5 ka. We tested a reconstruction with a different age-depth model based on lower age errors for this period. A WPS of these results reveals such high frequencies (not shown), highlighting the importance of accounting for uncertainty in the age-depth transformation. Although the core of Pickarski et al., 2015b and Pickarski et al., 2015a is investigated with a comparatively high resolution, the information of the DO events cannot be resolved after the conversion from depth to age.

MIS 2 from **28 – 14.5 cal ka BP** is described in Pickarski et al., 2015b with cooler and drier conditions compared to MIS 3. The temperature we reconstructed is consistent with this description. However, precipitation increases at the beginning and decreases after about 22 cal ka BP. This is due to Ranunculus acris with comparatively low T_{ANN} but high P_{ANN}. Nevertheless, the amount of precipitation is not greater than in the interstadial MIS 5c. In Litt et al., 2009, the period from circa 20 – 14 cal ka BP that includes the LGM is characterized by cold and semi-desert steppe vegetation. As a result, both climate variables decrease to a minimum of about 5.5 °C and about 400 mm.

The onset of MIS 1 from **14.5 – 11.5 cal ka BP** is indicated by higher pollen values from trees and shrubs (Litt et al., 2009). In both WPSs in Fig. 3.41 we see significant power with periods of about 0.8 – 2 ka. This could include the variations of DO 1 and YD. As far as the climate reconstruction is concerned, both variables show high-frequency fluctuations during their increase. However, the

relatively low amplitudes make it difficult to detect DO 1 and YD.

During the Holocene from **11.5 – 0 cal ka BP**, Litt et al., 2009 mention a sharp increase in moisture. From a growth delay of 3,000 years in deciduous oaks, they infer a dry spring and summer in the early Holocene. This is reconstructed in both variables and we still see a continental climate classification. During the MH (7 – 4 cal ka BP) this transitions to a warm Mediterranean climate with temperatures around 9 °C and around 620 mm. The following period until today is characterized by increased human influence (Litt et al., 2009). We treat this as with the reconstructions before (e.g. section 3.3.1). As a result, temperature and precipitation do not show a strong decrease around 3.8 ka BP as in the AP/NAP curve. The WPS show significant power at higher frequencies during the Holocene, enabled by the relatively small age errors. They also cover the typical period of Bond events around 1.5 ka, but it is difficult to determine which climate fluctuation belongs to which event. Finally, both reconstruction variables of the recent time slices approach the current climate measurements.

Summary and conclusion

The quantitative climate reconstruction of the Lake Van region during the last ca. 130,000 years includes not only paleobotanical data, but also isotopic information from the Mediterranean region. Compared to the reconstructions from the Levant, the AP/NAP values do not contain Pinus components and are correlated with both climate variables. As a result, the posterior taxa weights in Fig. 3.40 (b) show a relatively low weight for Pinus nigra. In contrast, Quercus deciduous and Ulmus minor have mean posterior weights that are higher than the prior values. They are important to determine the warmer and wetter interstadials. On the other hand, Ephedra distachya has the highest mean posterior weight and is used by the algorithm to describe the end of the late Penultimate and late Last Glacial. In summary, only 10 taxa information contain enough variation to find a compromise between these two proxy references and two climate anchor points. This can be checked in Fig. 3.40. (a), where the corresponding explained variances have values between 0.4 – 0.6.

Moreover, the final result resembles quite well the qualitative reconstructions described in Pickarski et al., 2015b; Pickarski et al., 2015a; Litt et al., 2009. Interestingly, the most recent climate classification also occurs at the end of the Penultimate Glacial. Nevertheless, the respective reconstructed climate values are about 1.5 °C and about 200 mm lower than today, which does not contradict the qualitative reconstructions mentioned above. This finally leads to a comparatively low variation of T_{ANN} and a relatively high variability of P_{ANN} throughout the reconstruction period.

3.3.5 Lake Iznik

Lake Iznik is the largest lake in the Marmara region, which lies between the Mediterranean and the Black Sea (Miebach et al., 2016). The lake is surrounded by mountain ridges with peaks between 810 and 1293 m amsl. In Fig. 3.42 we see a climate diagram of the grid point closest to Lake Iznik, describing a warm-summer Mediterranean climate Csb with warm/dry summers and mild/wet winters. The average annual temperature in this region is about 12.2 °C, and the average annual precipitation is 610 mm. Due to the orography, this value varies greatly, which affects the vegetation around Lake Iznik. Litt et al., 2021 describe it with a transition from Mediterranean woodland and mesic Euxinium forest. The latter is adapted to lower temperatures and higher precipitation than Mediterranean vegetation. How these present-day features might have changed over the last 31,000 years is discussed in the next sections.

Introduction

In 2005 AD and 2009 AD, several cores were taken during two drilling campaigns. Miebach et al., 2016 use a composite of some of these cores and investigate the corresponding vegetation history of the last ca. 31,000 years. The proxy references used are shown in Fig. 3.43, which are applied in the same way as in Sect. 3.3.4. For tree pollen, we only consider temperate trees and shrubs (Miebach et al., 2016). As with the reconstruction of the Dead Sea during the Holocene (see Sect. 3.3.1), we use

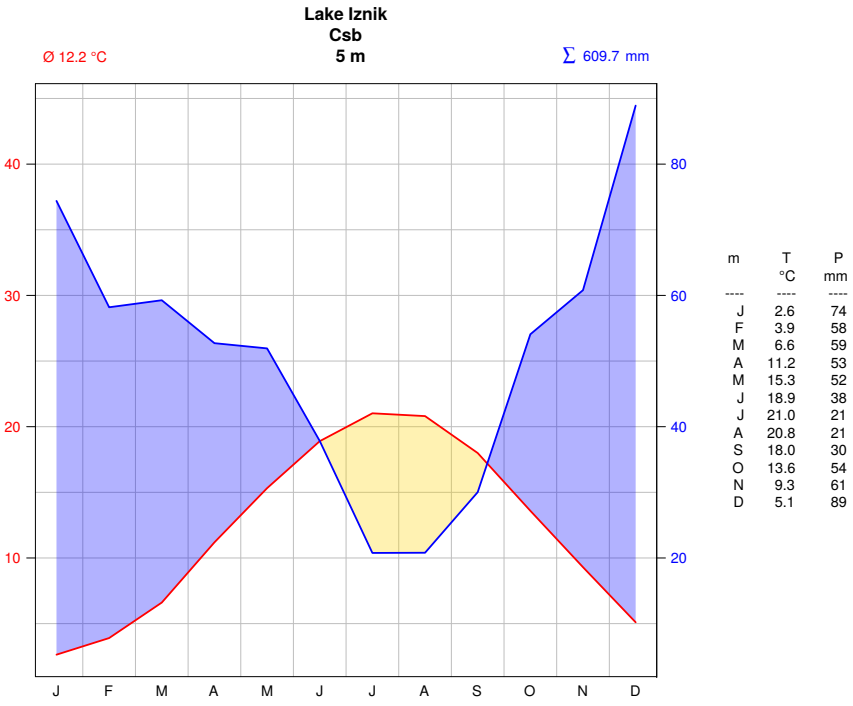


Figure 3.42: As Fig. 3.18, but for Lake Iznik.

the recent Koeppen-Geiger classification Csb as the climate anchor point, although the most recent period is 1400 AD. This assumption is confirmed by ca. 80 % of our local reconstructions over the area depicted in Fig. 3.3, where no change in classifications is apparent in the last 600 years. The marginal probability density distributions of the recent climate anchors are shown here as light green areas in Fig. 3.44.

In contrast to the previous reconstructions, we add an additional climate anchor point for a past period. This can be seen in Fig. 3.44 (b) at 21 cal ka BP, which corresponds to the LGM. Miebach et al., 2016 mention that precipitation during this period was relatively low and can be compared to that of the steppe in central Anatolia. In Fig. 3.17 (c) we see some grid points in central Anatolia indicating a cold steppe climate (BSk). Therefore, these are used to create the PDF of the annual

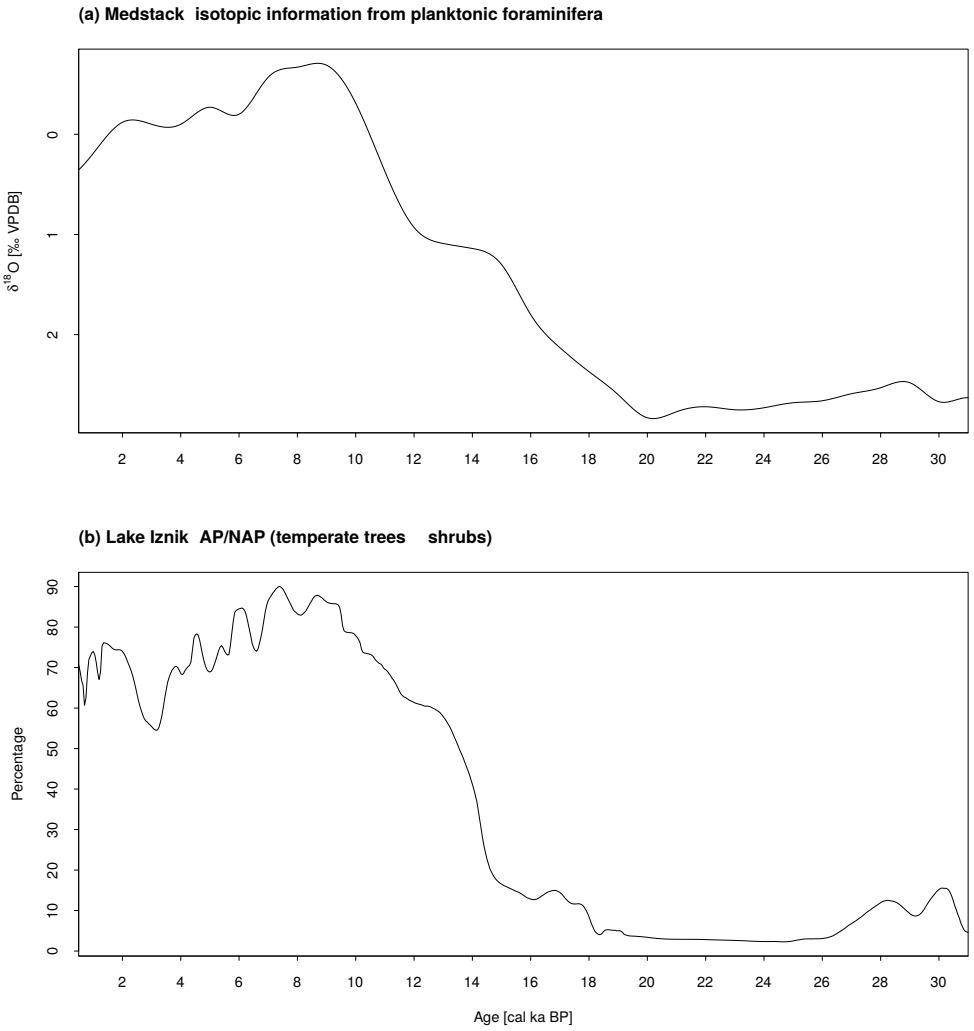


Figure 3.43: As Fig. 3.38, but for Lake Iznik. Only temperate trees and shrubs are considered.

precipitation shown in Fig. 3.44 (b).

The Bayesian age-depth model is based on the data presented in Roeser et al., 2016. They use a mixture of two tephra and radiocarbon measurements with and without reservoir corrections.

Reconstruction

In the following we describe the final reconstruction shown in Fig. 3.44. Panel (a) shows the reconstructed annual temperature T_{ANN} in °C and (b) the annual precipitation P_{ANN} in mm/a. At the beginning of the explanations we focus on the solid black lines. Later we will work out the difference between these reconstructions and others whose medians are shown as blue lines. Miebach et al., 2016 divide the corresponding vegetation history into different periods, which are also used in this work.

The period of **31.1 – 28.4 cal ka BP** is described by Miebach et al., 2016 as steppe vegetation and they conclude dry and cold climatic conditions. Our reconstruction in Fig. 3.44 shows a mean T_{ANN} of about 6.5 °C and P_{ANN} of about 600 mm. This results in an approximate Koeppen-Geiger climate classification Dfa, which describes a fully humid continental climate with hot summers. For comparison, such climate variables are similar to the current climate in central Minnesota. This area shows a transition zone from savannah to mixed forest in the time before settlement (Goring et al., 2016). Such sensitive conditions fit the explanations of Miebach et al., 2016 regarding two different rapid vegetation changes, including an increase in pine and deciduous trees. Therefore, even a relatively small increase in the two climate variables, as indicated in Fig. 3.44 with DO 3 and DO 4, can be sufficient for such a vegetation change. The WPSs in Fig. 3.45 confirm these two Dansgaard-Oeschgar events with a relatively high significant power of periods about 1.5 ka. In the corresponding reference curves in Fig. 3.43 we see that at least in (b) these features are also present. The resolution of Medstack is too coarse to resolve these high-frequency cycles.

The period that includes the LGM (**28.4 – 18.4 cal ka BP**) shows a cool and dry steppe (Miebach et al., 2016). During these ages, the reconstructed variables show a mean T_{ANN} of about 5 °C and P_{ANN} of about 550 mm. These values lead to a fully humid continental climate with warm summers (Dfb). A similar climate can be found today in eastern North Dakota, around the city of Fargo. In the pre-settlement period, this place was located on the prairie (Goring et al., 2016). In Fig. 3.44 (b) we see the additional climate anchor at 21 cal ka BP based on the explanations of Miebach et al., 2016. Without this anchor, the annual precipitation would be about 100 mm higher (blue line). This shows that the reconstruction algorithm is able to find reconstruction patterns that approximate this prior PDF.

During the post-LGM from **18.4 – 15 cal ka BP**, a change in steppe components is observed by Miebach et al., 2016. The dominance of Artemisia in combination with the presence of Ephedra indicates higher temperatures and lack of precipitation. This transition can also be seen in our reconstruction. The temperature rises to about 6.5 °C, but precipitation decreases to approximately 500 mm. This in turn leads to a fully humid continental climate with hot summers. Climate change is

also evident in both reference curves, as they each show an increase. Nevertheless, the algorithm is able to recognize the vegetation information described above.

The Late Glacial from **15 – 12.1 cal ka BP** is described by Miebach et al., 2016 with an increase of deciduous trees and a peak of Pinus. They indicate a warmer and wetter climate than DO 3 and DO 4. Although the Koeppen-Geiger classification is the same, both reconstructed climate variables have higher values. T_{ANN} is about 7 °C and P_{ANN} is around 650 mm. With the onset of 15 cal ka BP, we

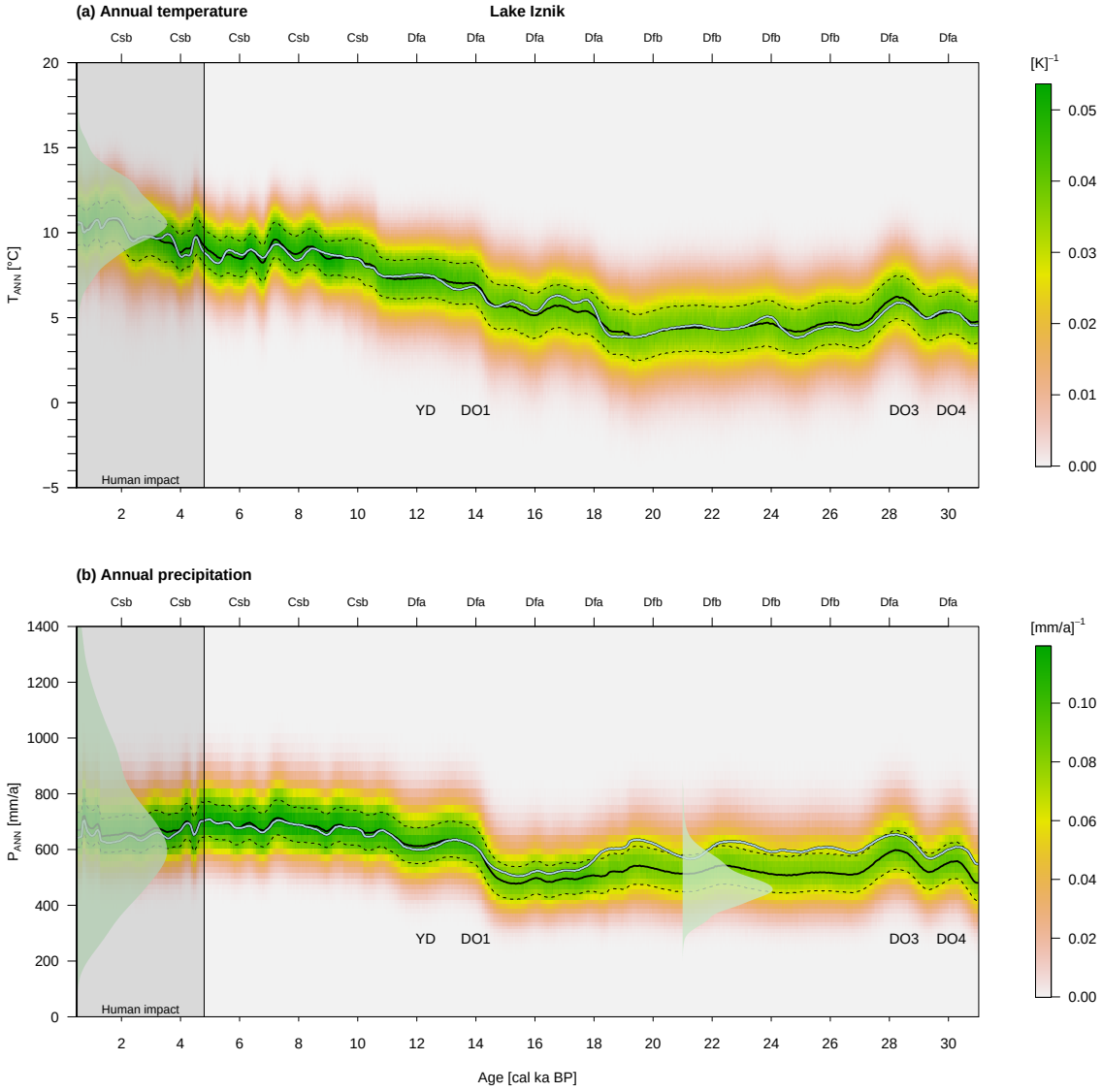


Figure 3.44: As Fig. 3.25, but for Lake Iznik. The gray shaded areas indicate the period when human influence is not negligible. Furthermore, the light green regions show the PDFs of recent climate anchor points. In terms of precipitation, another climate anchor point for the LGM is shown.

can observe relatively high power with periods around 1.5 ka BP in both WPSs. Here, Miebach et al., 2016 recognize DO 1 (Bølling-Allerød) at about 14.6 cal ka BP, which is associated with an expansion of temperate tree species. They also see a drop in AP throughout 13.3 – 12.3 cal ka BP, suggesting a drier and/or cooler climate during the YD. The first event is clearly visible in both reconstructed variables. However, the YD is only slightly seen in P_{ANN}, where a decrease of about 50 mm occurs.

During the early Holocene from **12.1 – 9 cal ka BP** a transition from open woodland to dense forests took place (Miebach et al., 2016). Accordingly, both the climate variables and the reference curves increase. The approximate climate classification transitions to a Mediterranean climate Csb, which is the current climate in this region.

Miebach et al., 2016 describe the Holocene from **9 – 0.6 cal ka BP** with a temperate deciduous forest and warm climatic conditions. They note many high-frequency fluctuations in the vegetation spectrum, which could be climatic or human-induced. After about 4.8 cal ka BP, more anthropogenic indicator taxa (olives, cereals, walnuts) emerge, and we use the corresponding method described in Sect. 3.3.1, to take this into account. This leads to ignoring the decrease in AP by 3 ka BP, which is not reconstructed in either variable. Although this coincides with the collapse of the Late Bronze Age, according to Miebach et al., 2016 this vegetation development was superimposed by human influences. Both WPSs show the typical pattern of Bond events in the Holocene. This means an increase of periods from 1 ka in the early Holocene to 1.5 ka during the Middle Holocene (cf. Fig. 2.7 (d)). With respect to P_{ANN}, the fluctuations show a much higher frequency when human influence

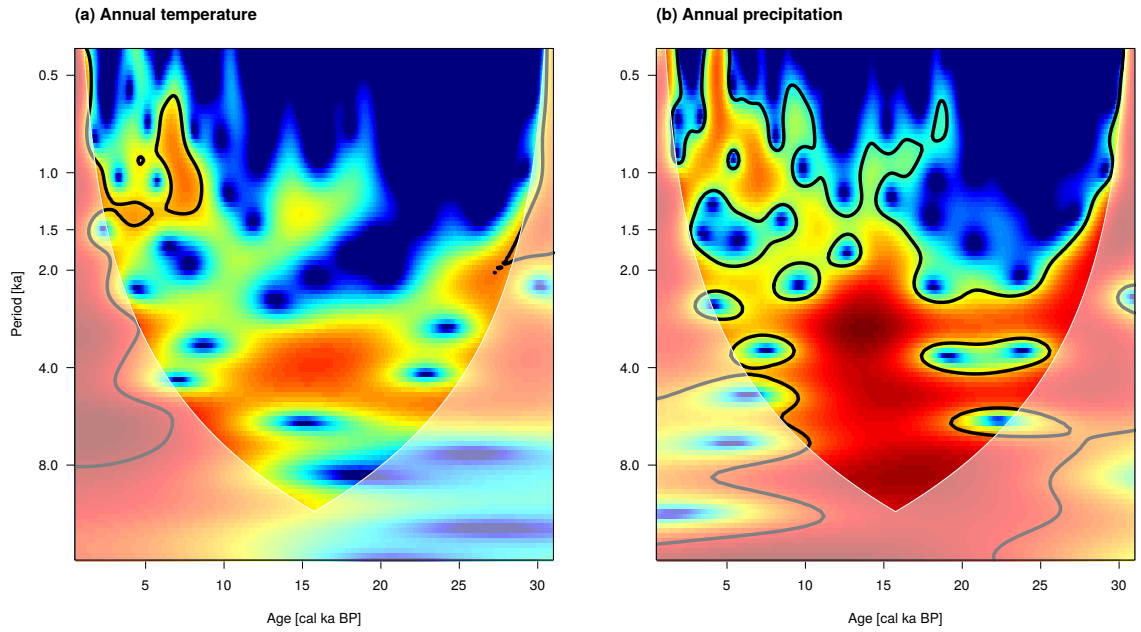


Figure 3.45: As Fig. 3.41, but for Lake Iznik.

increases (4.8 cal ka BP). This leads to another idea of how to deal with such an influence. In the zone of human impact we could use only a low-frequency part of the reference WPS, e.g. the periods higher 1000 years. The focus would increasingly be on the above-mentioned climatically driven periods, and the faster fluctuating human-induced patterns could thus be more and more neglected. Finally, both reconstructed climate variables approach the most recent climate measurements for the last time interval.

Summary and conclusion

The quantitative climate reconstruction of the region around Lake Iznik during the last 31,000 years includes not only paleobotanical data but also isotopic information from the Mediterranean Sea. During the Late Glacial, the reconstruction resembles the explanations from Miebach et al., 2016. This is shown by describing modern analogue transitions from prairie to mixed forests in the USA. Furthermore, an a priori climate anchor distribution for annual precipitation inspired by Miebach et al., 2016 is inserted into the algorithm, resulting in lower P_{ANN} during the LGM. The deglacialization period also follows the qualitative reconstruction from Miebach et al., 2016. The human influence from about 4.8 cal ka BP onwards is treated as before, resulting in much smaller fluctuations than recorded in the AP reference curve. In summary, the reconstruction reveals many well-known climate changes such as deglacialization, the Dansgaard-Oeschger and Bond events.

3.3.6 Lake Ohrid

Lake Ohrid is located on the border between Albania and Northern Macedonia at an altitude of 695 m amsl (Panagiotopoulos et al., 2020). It is surrounded by mountain ranges that reach a height of more than 2,000 m amsl. The climate diagram in Fig. 3.46 of the grid point closest to Lake Ohrid shows a warm Mediterranean climate Csb with warm/dry summers and mild/wet winters. The average annual temperature in this region is 9.2 °C, and the average annual precipitation is 790 mm. Due to the orography, these values are highly variable and influence the vegetation surrounding the lake, which acted as a vegetation refuge during the Quaternary climate cycles (Panagiotopoulos et al., 2020). Panagiotopoulos et al., 2013 describe the current vegetation as a mixture of Central European, Mediterranean and Balkan endemic plants. How these present-day features might have changed during the Early Pleistocene is discussed in the next sections.

Introduction

In 2013 AD, a composite core was drilled in the middle of the lake area as part of a drilling campaign (Wagner et al., 2019). Its length is about half a kilometre and covers the last 1.36 million years. In this study we want to focus on the period from $1,165 - 1,365$ cal ka BP within the Early Pleistocene. The corresponding vegetation survey is described in Panagiotopoulos et al., 2020. As with the

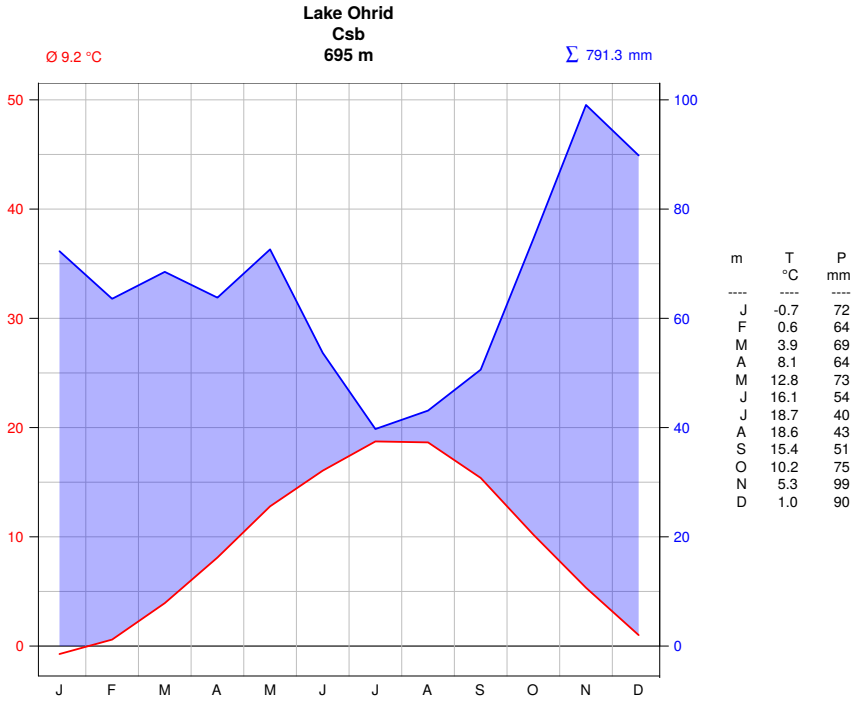


Figure 3.46: As Fig. 3.18, but for Lake Ohrid

reconstruction of the Dead Sea in Sect. 3.3.1, we cannot include climate measurements as anchor points. Medstack and the percentages of temperate trees are used as proxy references, which are shown in Fig. 3.47. We exclude Pinus from the AP to correlate it with the reconstructed annual temperature. This is based on the explanations in Wagner et al., 2019, where Medstack is linked to warm/cool SST and compared to this plant information from the lake.

An important aspect of the pollen spectrum is the occurrence of relict taxa in the Early Pleistocene Ohrid catchment (Panagiotopoulos et al., 2020). During the transition from the Early to the Middle Pleistocene, a vegetation shift led to the extinction of these taxa. In our reconstruction we include some modern distributions of Cedrus libani, Liquidambar orientalist, Pterocarya fraxinifolia, Zelkova carpinifolia, Carya, and Tsuga to establish analogies with the most important relict species (Konstantinos Panagiotopoulos, personal communication). Therefore, we compare one climate reconstruction that includes this plant information with another that excludes these taxa. At this point, it should be mentioned again that our method uses modern taxa distributions in conjunction with current climate data. That these conditions have not changed in the last 1,300,000 years is a strong assumption that must always be kept in mind in the following discussion.

We use tephra layers, orbital tuning points and two palaeomagnetic age reversals as input to our Bayesian age-depth model. These data are presented and explained in more detail in Wagner et al., 2019.

Reconstruction

In the following we describe the final reconstruction shown in Fig. 3.48. Panel (a) shows the reconstructed annual temperature T_{ANN} in °C and (b) the annual precipitation P_{ANN} in mm/a. At the beginning of the explanations we focus on the solid black lines. Later, we will work out the difference between these reconstructions and others whose medians are shown as blue lines. Panagiotopoulos et al., 2020 divide vegetation history into different periods, which is also used in this work. They are more or less consistent with the MISs presented. For the sake of simplicity, we do not go into the differences between these periods below.

We start with the period from **1,365 – 1,346 cal ka BP** (ca. MIS 43), which is described in Panagiotopoulos et al., 2020 with closed evergreen and mixed forests. Using only the two references

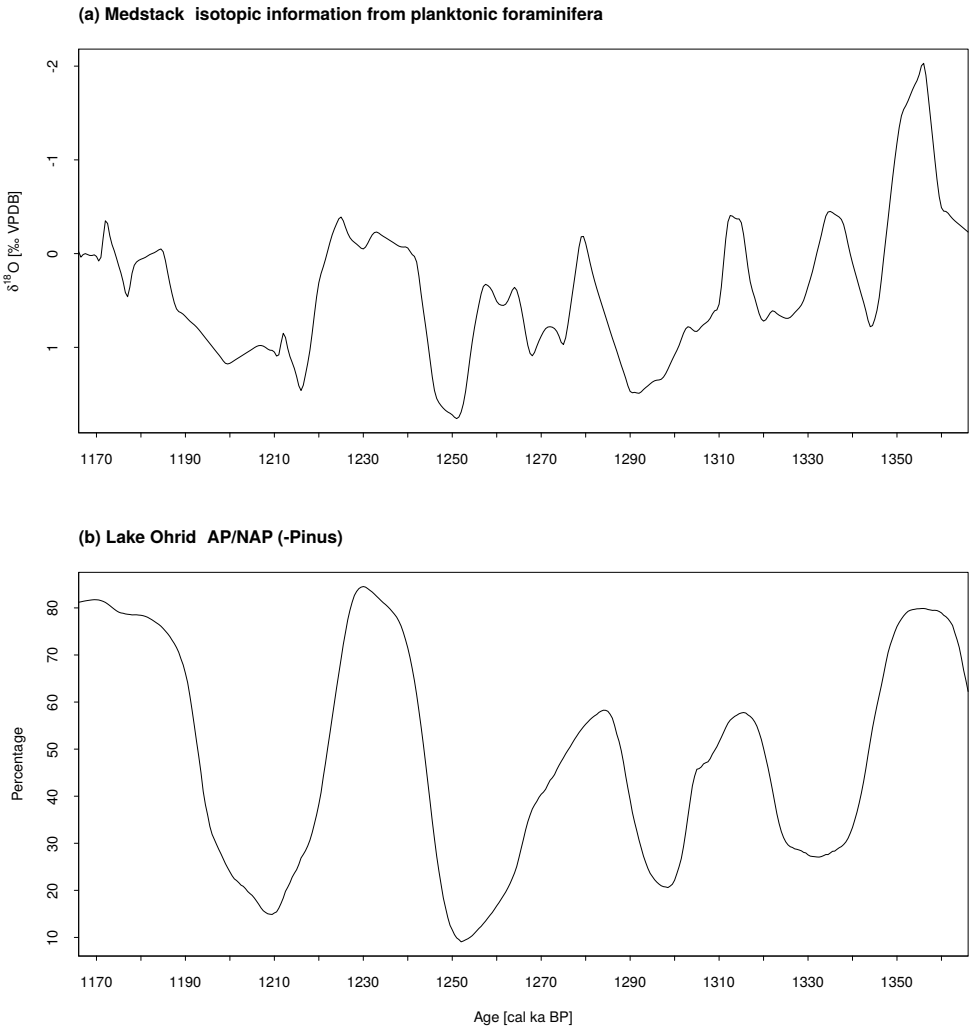


Figure 3.47: As Fig. 3.38, but for Lake Ohrid.

shown in Fig. 3.47, our reconstruction for T_{ANN} shows a mean annual temperature of 9 °C and a precipitation of about 700 mm. These are close to the present-day climate of this region (cf. Fig. 3.46) and lead to the same approximate Koeppen-Geiger climate classification, which indicates a warm summer Mediterranean climate. Both reconstructed variables reveal a phase-shifted pattern, with T_{ANN} showing a local maximum around 1,355 cal ka BP. To achieve this in accordance with the references, the algorithm uses Olea, which occurs at this time. The relatively high temperatures and low precipitation associated with this taxon result in this contrasting pattern. At the end of this

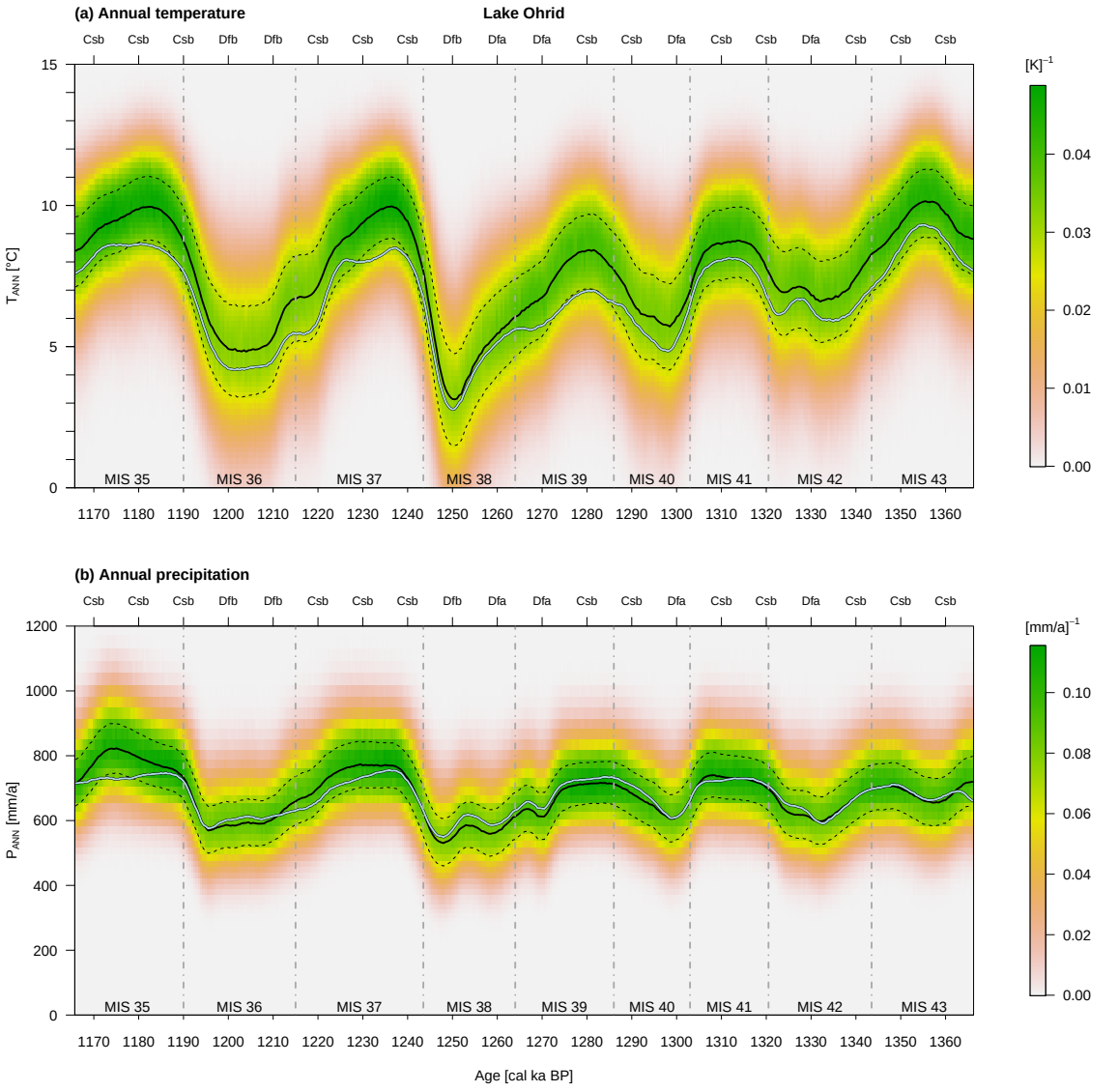


Figure 3.48: As Fig. 3.20, but for Lake Ohrid. The dashed vertical lines show the transitions of the respective Marine Isotope Stages.

interglacial, both variables show a decreasing behaviour.

The glacial period of **1,345 – 1,322 cal ka BP** (ca. MIS 42) reveals a dominance of pine forests and grassland associated with open steppe vegetation (Panagiotopoulos et al., 2020). Based on the decrease in reconstructed temperature and precipitation during 1,335 cal ka BP, the Koeppen-Geiger classification changes to a fully humid continental climate with hot summers (Dfa). The corresponding T_{ANN} and P_{ANN} resemble recent areas about 1000 km north of the Caucasus, where the natural vegetation is described in Bohn et al., 2007 as a transition from forest steppes to herb-rich and pure feather grass steppes. This also applies to the following glacial periods during MIS 36 and 40.

In the years from **1,321 – 1,305 cal ka BP** (ca. MIS 41) the vegetation is similar to that in MIS 43. Accordingly, the reconstruction shows similar values, although the annual temperature is cooler by about 1.5 °C. This is due to the reference curves and is achieved by using the relatively cold genus Abies, which is predominant during this period.

Panagiotopoulos et al., 2020 describe the periods **1,304 – 1,292 cal ka BP** (ca. MIS 40) and **1,291 – 1,265 cal ka BP** (ca. MIS 39) analogous to MIS 42 and MIS 41, respectively. According to this, our reconstruction results are quite similar.

The glacial of **1,264 – 1,243 cal ka BP** (ca. MIS 38) is described by Panagiotopoulos et al., 2020 with a dominance of pine forests and, in contrast to earlier time slices, with extensive open steppe vegetation. The reconstructed climate variables reveal their respective lows during the study period. T_{ANN} has values around 3 °C and P_{ANN} about 500 mm. Therefore, the Koeppen-Geiger classification transitions to a warm-summer continental climate Dfb without a dry season. These climate values are similar to the recent climate in northwestern Minnesota, where the natural vegetation is described in Goring et al., 2016 as a mixture of savannah and prairie.

The interglacial from **1,242 – 1,218 cal ka BP** (ca. MIS 37) shows a similar vegetation pattern as MIS 39. Nevertheless, the reconstruction shows higher maximum temperatures around 10 °C and higher precipitation amounts at about 800 mm. This is due to the rising values of Carya. As a proxy for this relict taxa we use Carya glabra/tomentosa, which occur today in the eastern USA (Konstantinos Panagiotopoulos, personal communication). They are therefore associated with relatively high precipitation and temperatures.

Panagiotopoulos et al., 2020 describe the period from **1,217 – 1,191 cal ka BP** (ca. MIS 36) analogous to the preceding glacial (MIS 38), but without extensive open steppe vegetation. This is noted by the reconstruction method, which leads to somewhat higher climatic variables.

The last interglacial within this study period from **1,190 – 1,165 cal ka BP** (ca. MIS 35) is similar to MIS 38. Again, the algorithm uses Carya to determine similar climate values.

Sensitivity study

The blue line in Fig. 3.48 shows the median of the reconstructions that do not contain the relict taxa mentioned above. Nevertheless, the AP reference curve still contains all relict information. Therefore,

we can compare both reconstructions with regard to these additional species.

At first glance, it can be seen that the blue line of the annual temperature is lower during the entire study period. Calculating the mean T_{ANN} and P_{ANN} values of the transfer functions used and comparing them with the subset where no relict species are considered results in a difference of about +4 °C and +150 mm. Furthermore, the percentage of all relict taxa used correlates with the reference curve AP in Fig. 3.47 of about 0.79. This causes the main difference between the black line and the blue line during the interglacials. In summary, the inclusion of relict taxa can increase the final reconstruction by up to about 2 °C and 100 mm.

Summary and conclusion

The quantitative climate reconstruction of the Lake Ohrid region for the period $1,165 - 1,365$ cal ka BP includes not only paleobotanical data but also isotopic information from the Mediterranean Sea. Both references are correlated with the reconstructed annual temperature. Several relict species were added to the taxa pool, which eventually led to higher temperature and precipitation values. The climate variables during the 5 interglacials are thus similar to those of the most recent measurements in this region.

We can see these glacial-interglacial cycles in the WPSs in Fig. 3.49 by looking at the branches with the highest power. These lie in a period of about 41 ka and describe the typical value of obliquity. During the Early Pleistocene, these climate cycles were driven by this Milankovitch cycle

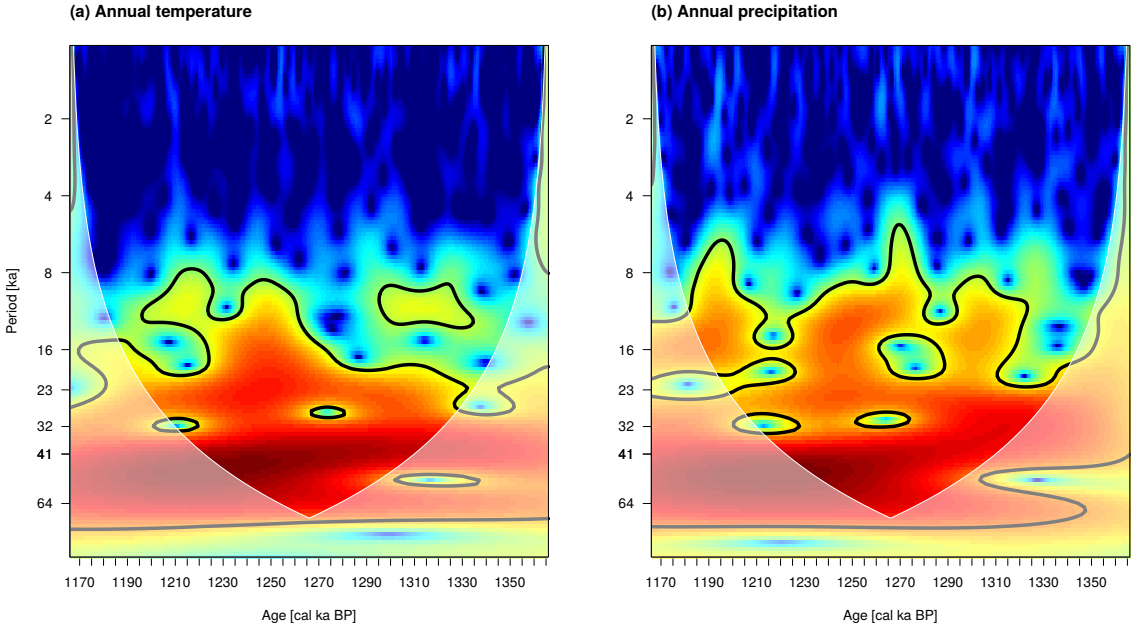


Figure 3.49: As Fig. 3.41, but for Lake Ohrid.

(Panagiotopoulos et al., 2020), which can also be seen in this study. In addition, both plots within the COI consistently show relatively high power with a period of about 23 ka. This reflects the precession that influences the intensity of summer solar radiation (Huybers, 2006) and is also noted by Panagiotopoulos et al., 2020. They mention that this could control the response of mesophilic species during interglacial periods.

Although the assumptions we make in order to carry out such a climate reconstruction are strong, all in all the result fits the explanations in the vegetation analysis in Panagiotopoulos et al., 2020. We therefore assume that our reconstruction technique provides realistic climatic values for the Early Pleistocene of Lake Ohrid.

3.4 Spatial climate reconstructions

In the following sections, we summarize 186 finally selected local reconstructions from the total 600 sites considered in this thesis. These are selected according to the criteria described in Sect. 3.2.2 and mapped onto the spatial grid shown in Fig. 2.20. This forms the background for the spatial reconstruction method described in Sect. 2.6, which is based on the 7 CMIP/PMIP experiments listed in Table 2.3. From this multi-model ensemble, we estimate a prior distribution that is updated using our local reconstruction information. These assimilated posterior climate distributions include not only the proxy details from the local reconstructions such as plant data from lakes and mires, isotopic information from speleothems, marine sediments, and ice cores, but also the appropriately adjusted boundary conditions of the CMIP/PMIP simulations such as greenhouse gas concentrations, land-sea distributions, ice sheets, topography, and orbital configurations (Abe-Ouchi et al., 2015).

We want to apply this approach for 3 different time periods. First, we consider the historical CMIP ensemble within the reference period $1961 - 1990$ AD. The corresponding assimilated spatial climate structures are verified using CRU data. Subsequently, the Middle Holocene and Last Glacial Maximum periods are reconstructed and analyzed. For consistency, all of the following anomalies are calculated with respect to the CMIP ensemble mean. Note that the results of each period are summarized at the end of the corresponding discussion. For this reason, no additional summary is provided at the end of this Sect. 3.4.

3.4.1 Historical

In this section, we will take a closer look at a spatial climate reconstruction based on the historical CMIP multi-model ensemble and 20 local reconstructions from this thesis. Therefore, we will test whether the spatial reconstruction method provides a useful result based on our reconstructed local climate information. To this end, we first consider local results that fall roughly within the HIST reference period. Specifically, we assume that due to the relatively coarse temporal resolution of our local reconstruction method of 50 years and the climate information smoothed by the age-depth transformation, time slices younger than 1950 AD can be compared to the reference period. This data is then used to assimilate the HIST CMIP ensemble. The corresponding parameter set of the shrinkage matrix is determined as described in Weitzel et al., 2019 (see Sect. 2.6).

Fig. 3.50 contains statistical information about the local reconstructions and shows where they are located. If there are multiple locations within a grid point, the mean values of the corresponding distributions are displayed. For this reason, only 13 and not 20 grid points are shown in Fig. 3.50. Significant anomalies are calculated grid box-wise. The significance level is 5 %, i.e. it is tested whether the CMIP reference is below the 2.5th percentile and above the 97.5th percentile of the respective posterior distribution. In these cases, the null hypothesis, which states that the anomaly is 0, is rejected. With respect to the annual temperature in (a), the grid points farther north show an

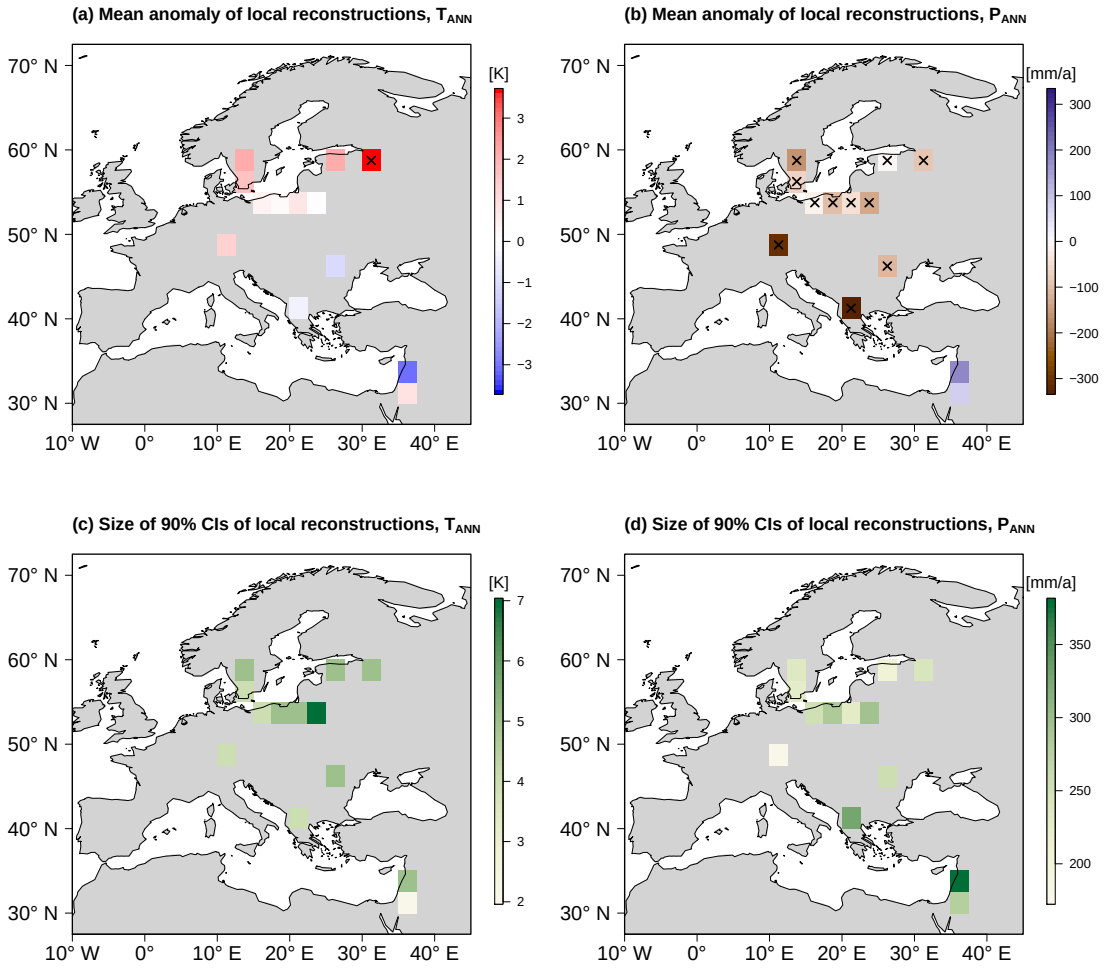


Figure 3.50: Panels (a) and (b) show the mean T_{ANN} and P_{ANN} anomaly with respect to the CMIP ensemble for the grid points for which local reconstruction information are available. The grid points with crosses indicate significant anomalies. In (c) and (d), the corresponding 90 % credible intervals are depicted.

overall positive temperature anomaly. In contrast, the corresponding locations in (b) have a negative precipitation anomaly. The three southernmost grid points include the local climate reconstructions of En Gedi (Sect. 3.3.1), Lake Kinneret (Sect. 3.3.2), Birkat Ram (Sect. 3.3.3), and Lake Prespa (Sect. 3.2.2). For further information, the reader is referred to the respective sections.

When these local data are incorporated into the spatial reconstruction scheme, the results shown in Fig. 3.51 are obtained. The panels (a) to (d) are analogous to those in Fig. 3.50. It is noticeable that the corresponding values of these two figures are similar. This shows that the uncertainty in the prior distribution is large enough to lead to a posterior reconstruction that is mostly determined by the available local information. Overall, the spatial mean of the mean posterior temperature in Fig. 3.51 (a) is 11.08 °C with a mean 90 % credible interval (CI) from 10.14 °C to 11.97 °C. The

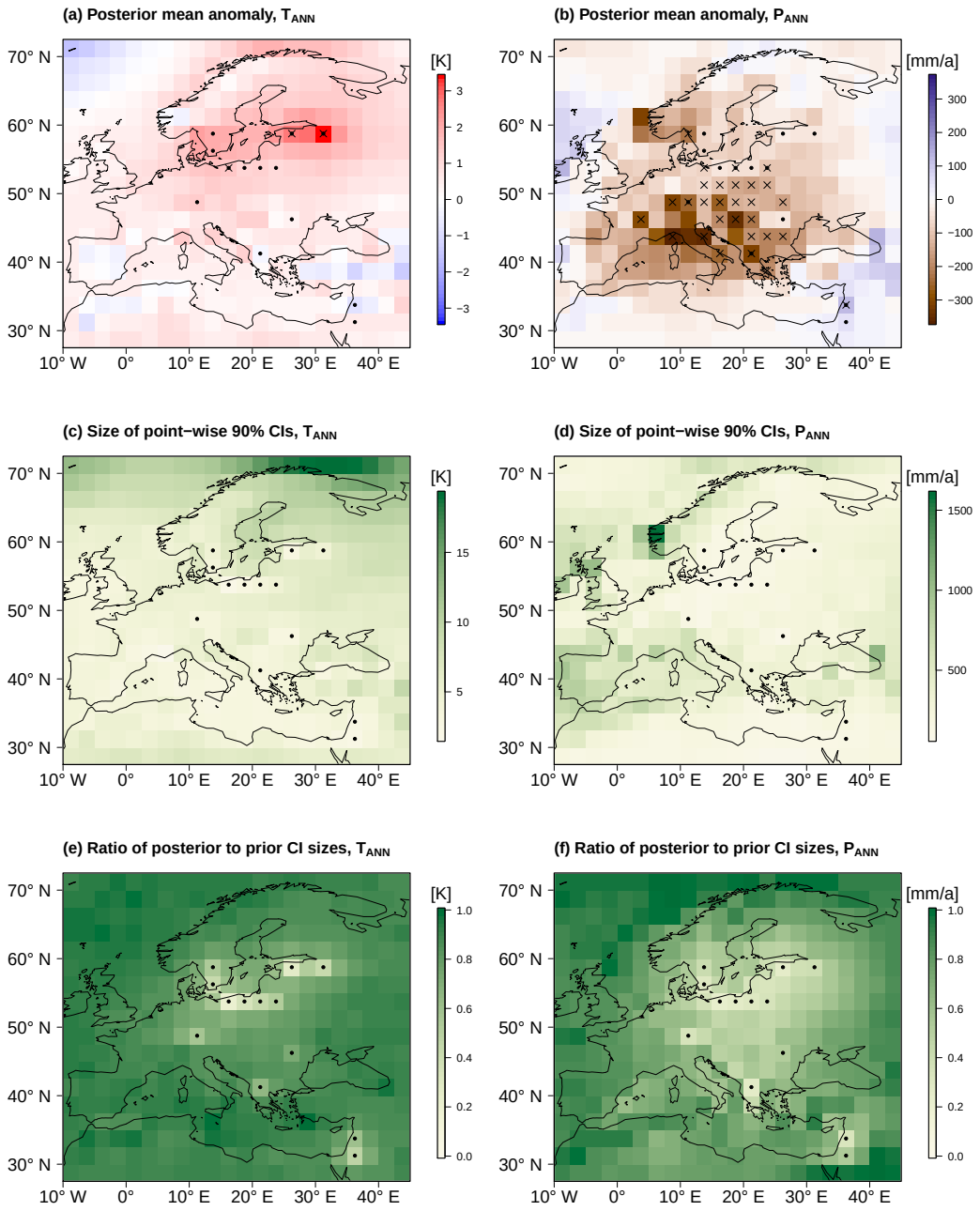


Figure 3.51: Panels (a) and (b) show the mean posterior anomaly of T_{ANN} and P_{ANN} with respect to the CMIP ensemble. The grid points with black dots contain local reconstruction information, and those with crosses indicate a significant anomaly. In (c) and (d), the corresponding 90 % credible intervals are depicted. The ratios between the 90 % credible interval of the posterior and prior simulations are given in panels (e) and (f).

corresponding values for annual precipitation are 563.85 mm with a mean 90 % CI from 506.9 mm to 624.22 mm. On average, the values are 0.62 °C warmer and precipitation is 54.16 mm less than the

CMIP reference climatology. For further analysis, we consider the posterior means over land. These are 10.48 °C and 553.26 mm, representing a decrease of 0.6 K and 10.59 mm, respectively, from the values of the full map section. The decrease in temperature is related to the fact that the grid points within the area under consideration have relatively high temperatures above water. This does not apply to precipitation, where comparatively low amounts in the south and east are balanced by high values in the northwest. In addition to the predominantly positive temperature anomalies in Fig. 3.51 (a), there are also some negative anomalies in the following areas: Iberian Peninsula, Atlas Mountains, Northern Levant, Anatolia and northwestern part of the reconstruction domain. The situation is similar with precipitation anomalies, although these have more grid points with significance. This is due to the relatively low variance of the respective local reconstructions, which thus constrain the posterior distribution more. In panel (b), the predominantly negative anomalies in the middle of the reconstruction area are surrounded by positive anomalies. This pattern emerges from local information approximating the CRU data of the reference period for the most recent time slice (see Prior climate module in Fig. 2.17). In summary, Fig. 3 (a) and (b) reveal the systematic errors between the ensemble members and observations, with temperature on average too low and precipitation on average too high.

We also see in panel (c) that the temperature uncertainties are greatest in the eastern Norwegian Sea and the western Barents Sea. This is the area where the CMIP ensemble has the greatest spread and the local data are relatively far away. The same applies to the precipitation in panel (d). Here, the highest values are found mainly on the western coasts of the reconstruction domain, with a maximum in southeast Norway, which is also the area with the highest annual precipitation. The mean values of the uncertainties in (c) are 5.71 °C and in (d) 342.69 mm. These are lowest where local information most constrain the CMIP ensemble.

In panels (e) and (f) we see the relationship between posterior and prior uncertainties. Again, it is noticeable that the grid points with local reconstruction information exhibit a reduction in uncertainties, which is 63.01 % on average. This compares to 17.77 % for the other grid points. The reduction of the total uncertainty is higher for P_{ANN} than for T_{ANN}, which is due to a stronger constraint from relatively lower variances of the respective local reconstructions.

To find a compromise between the local reconstructions and the CMIP ensemble, the spatial reconstruction method determines weights for each ensemble member. These are shown in Fig. 3.52 in the form of boxplots, showing which simulation combination best fits the local climate distributions. On average, only the MRI has higher posterior weights than the prior weights. The reason for this is that this most closely reflects the patterns of the local reconstructions in Fig. 3.50 (a) and (b). In particular, the positive/negative anomalies in temperature/precipitation in Central Europe are striking in this respect.

As mentioned earlier, the climate values of the considered local reconstructions of more recent periods are designed to approach the corresponding CRU data from 1961 – 1990 AD. This would mean that the assimilated posterior CMIP structures in Fig. 3.51 (a) and (b) should be more similar to the

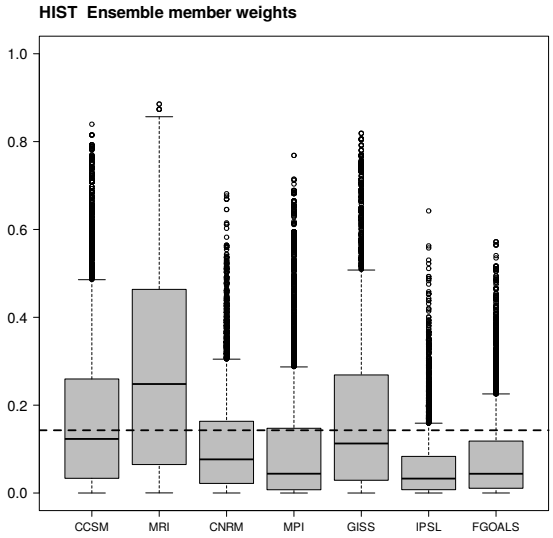


Figure 3.52: The boxplots represent the selected MCMC samples of the posterior weights of each ensemble member. The corresponding prior weight is depicted by the horizontal dashed line. The order of the boxplots is from left to right with descending horizontal resolution of the underlying models (cf. Table 2.3).

corresponding CRU climate data than the prior CMIP ensemble containing the systematic errors of the individual models. To test whether this is the case, we use the CRPS to compare the posterior and prior distributions with the CRU data. If we define $\mathrm{CRPS}_{\mathrm{prior}}$ in terms of the prior as the reference forecast and $\mathrm{CRPS}_{\mathrm{posterior}}$ as the forecast, we can calculate the CRPSS via Eq. 3.1. In contrast to CRPS, this value is positively oriented, and the respective grid point-wise results are shown in Fig. 3.53 (a) and (b). For both variables, the CRPSS is highest in Central Eastern Europe, where most local information are available. In panel (a) we see two grid points with negative scores. One is located in the Southern Carpathians, the other in En Gedi on the Dead Sea. For the latter, the mean difference between the prior CMIP ensemble and CRU is on the order of 10^{-6}. This makes it difficult to make improvements through our local reconstruction data. However, the posterior assimilated CMIP ensemble does not differ substantially either, so that the CRPSS is only slightly negative here. The situation is different with the more northern grid point in the Southern Carpathians. There, the mean difference between CMIP and CRU is about 1.3 K. Since the posterior CMIP ensemble extends this to about 1.5 K, we see the corresponding negative CRPSS. Annual precipitation shows a similar pattern to temperature. However, in this case, all grid points with local information have a positive CRPSS. If we now consider CRPS in terms of $\mathrm{T}_{\mathrm{ANN}}$ in (c), we see directly that the assimilated posterior CMIP ensemble not only achieves a mean improvement for the grid points with local reconstructions, but also for the others. Similar results are obtained for precipitation in (d), although the grid points without local sites show much lower mean improvement. This is due to the relatively large differences in Southeast Europe, as indicated by the negative CRPSS values in panel (b). However, a two-sample Kolmogorov-Smirnov

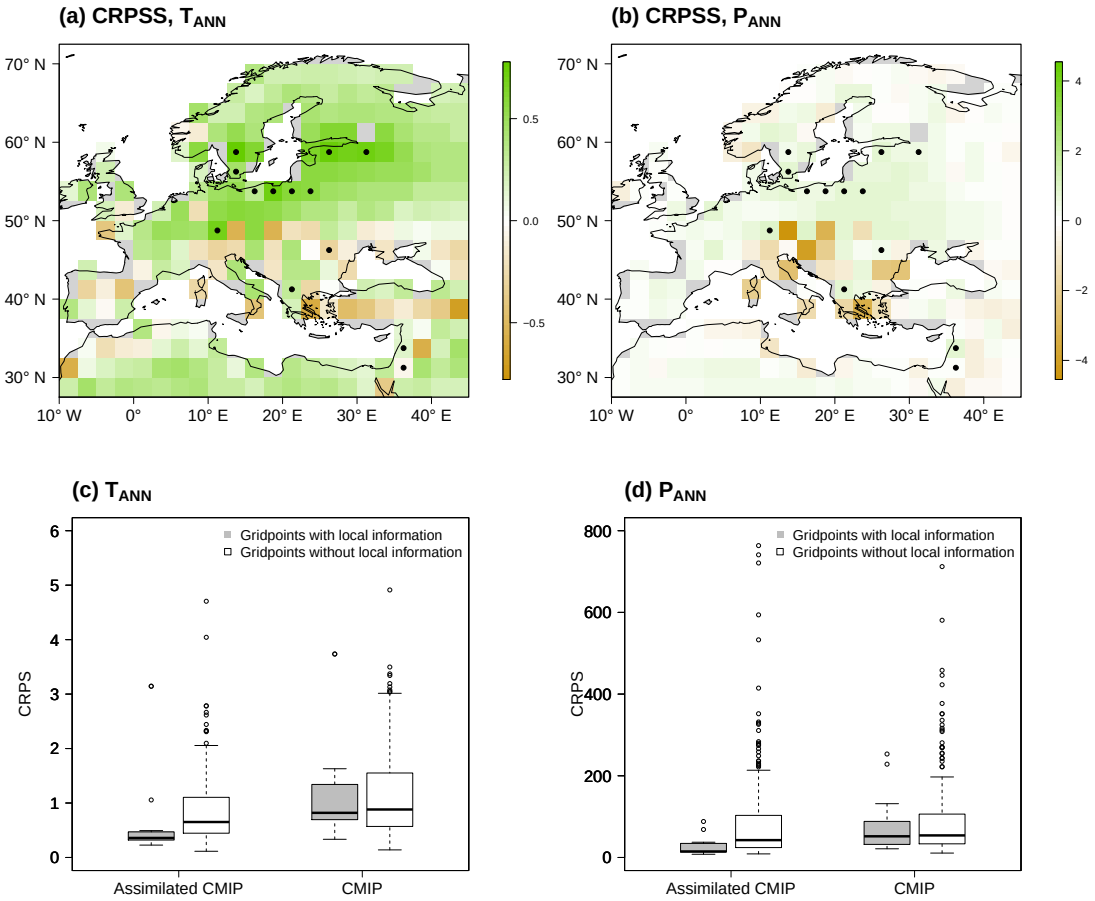


Figure 3.53: Panels (a) and (b) show the CRPSS of the posterior T_{ANN} and P_{ANN}, respectively. The grid points with the black dots contain local reconstruction information. In (c) and (d), the corresponding CRPS of the grid points with black dots (gray boxplots) and those without (white boxplots) are shown. Additionally to the output of the posterior simulations (Assimilated CMIP), the analogous results of the prior distribution (CMIP) are also presented.

test suggests that when comparing the respective samples from the boxplots, only the prior CMIP distributions of (c) and (d) were drawn from the same probability distribution. In particular, this means that the differences between the samples of the white boxplots from (d) are so large that no equal probability distribution can be assumed. This supports the statement above that not only the representative samples of grid points with local information, but also the rest of the data shows a non-negligible change to lower CRPS values after assimilation.

All in all, this demonstrates that the spatial reconstruction method can assimilate the local reconstruction information into the prior CMIP ensemble. We were not only able to determine this from the mean values between Fig. 3.50 and 3.51, but also quantify it in Fig. 3.53. It also shows that our local reconstructions are approaching the current anchor points as intended. Since the

reconstructions for the MH and LGM that now follow are based on the same CMIP models, but with modified boundary conditions, we assume that the respective uncertainties in the prior distributions are large enough to show similar flexibility to HIST.

3.4.2 Middle Holocene

In this section, we will take a closer look at a spatial reconstruction of the Middle Holocene. The procedure and presentation is analogous to that described in the previous section. Here we use 152 available local reconstructions with information on 6 cal ka BP and the shrinkage matrix parameters based on the PMIP experiments of the MH. These have orbital parameters and greenhouse gas concentrations adjusted according to the PIMP3 protocol described in Sueyoshi et al., 2013.

Fig. 3.54 shows statistical information of the local reconstructions created in this work. As for the

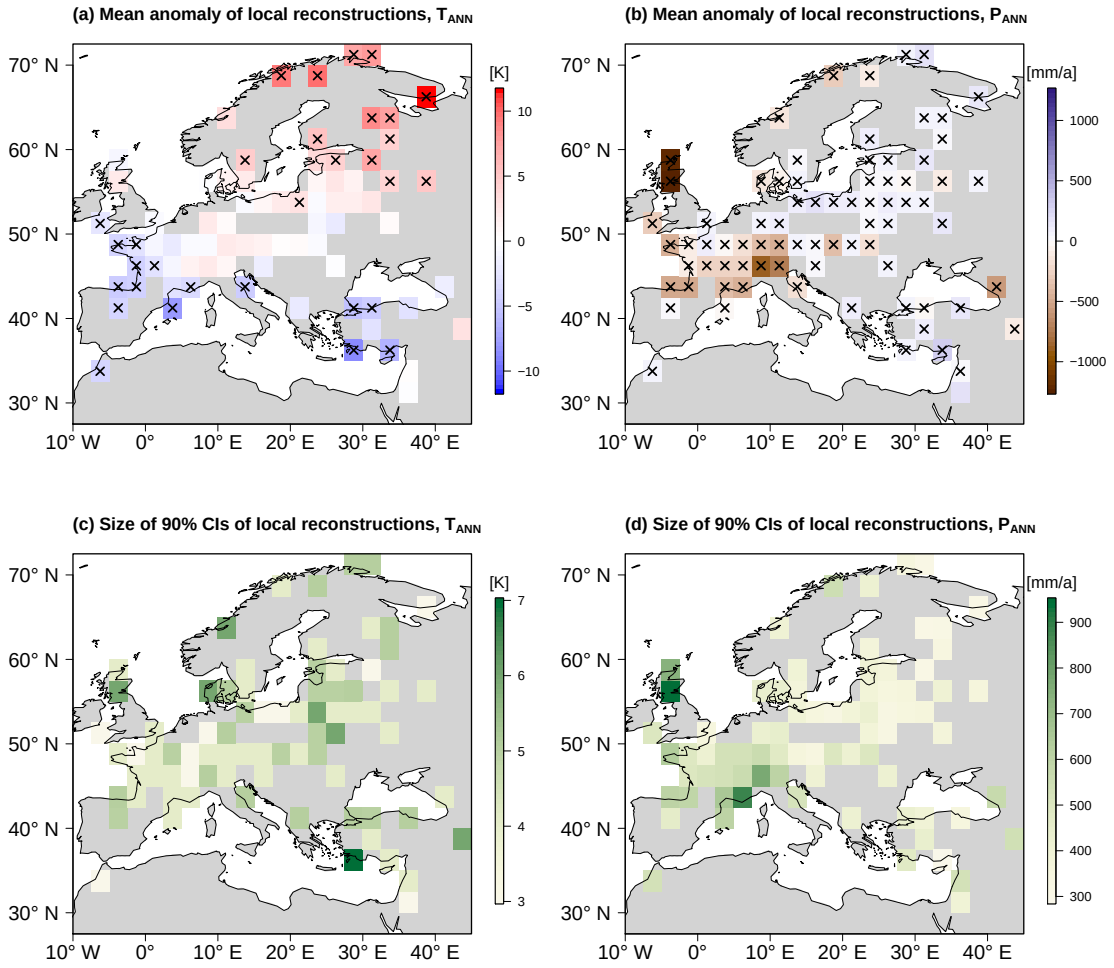


Figure 3.54: As Fig. 3.50, but for Middle Holocene. Again, the anomaly relates to the CMIP ensemble mean.

mean temperature anomalies in (a), it is noticeable that northeastern Europe is significantly warmer. In contrast, predominantly significant negative anomalies are observed in the Mediterranean region and in Western Europe. We find a similar dipole structure for the mean precipitation anomalies with the difference that the local reconstructions in Northwest Africa, Anatolia, and the Levant indicate

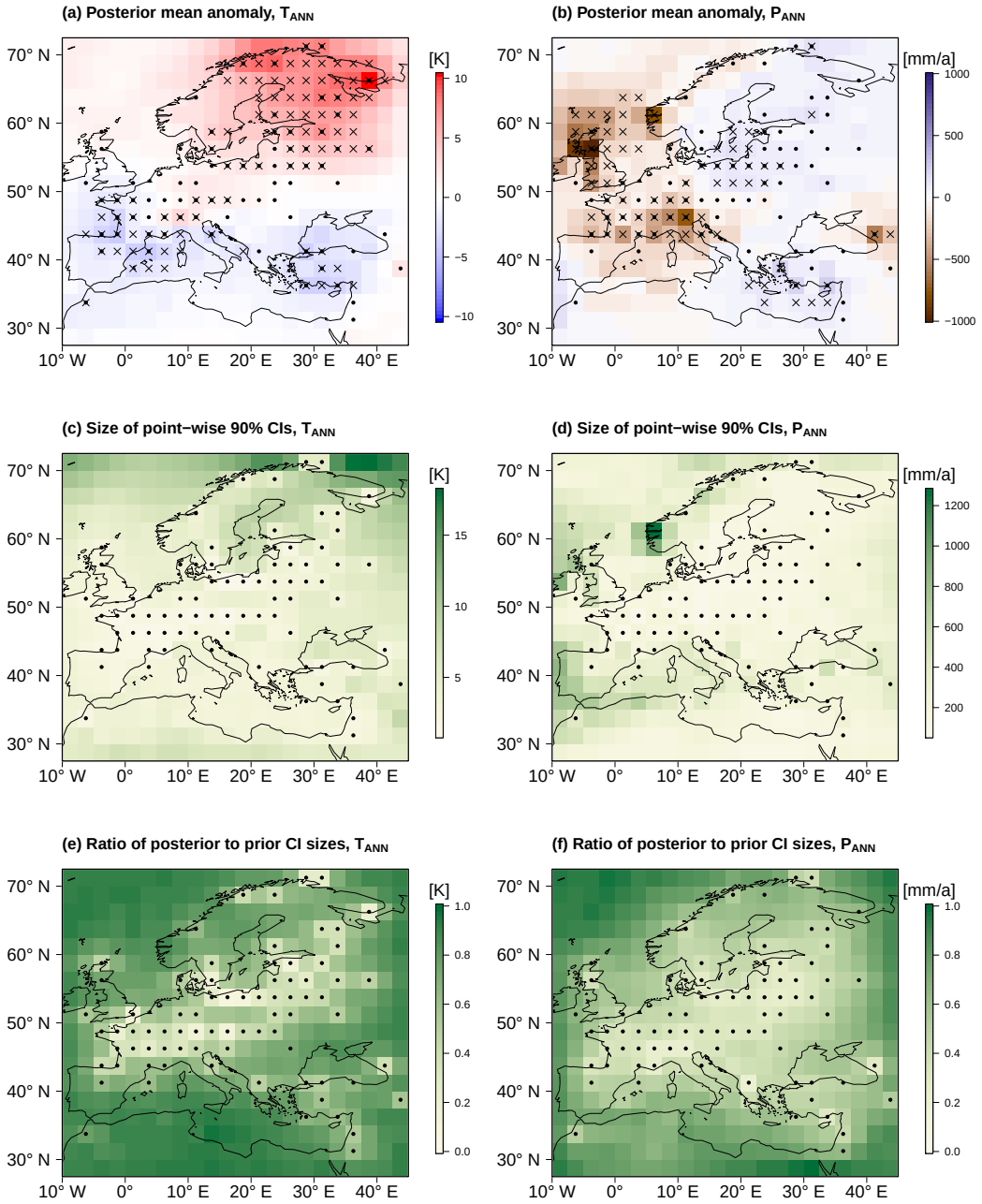


Figure 3.55: As Fig. 3.51, but for Middle Holocene. Again, the anomaly relates to the CMIP ensemble mean.

wetter conditions. The largest negative precipitation anomalies are found in the mountainous regions of the British Isles, the Alps, and the Caucasus. These are due to the large variability of P_{ANN} within the relatively coarsely resolved grid points. Dipole structures also stand out in the uncertainties of the local reconstructions in (c) and (d). These characterize in (c) the more maritime areas of Western Europe with predominantly lower uncertainties and the more continental regions of Eastern Europe with mostly higher variability. The reverse is true for the annual precipitation in (d). Such structures are due to the transfer functions of the respective taxa on which the local reconstructions are based. Thus, in more continental areas, the plants considered are those that have lower precipitation and higher temperature differences. In the more maritime areas of Western Europe, the opposite is true.

In Fig. 3.55, we see the spatial reconstructions based on the local information from Fig. 3.54. It is immediately noticeable that the prior PMIP ensemble, as in the previous section, has sufficient variability to adapt to local information. Thus, in (a) and (b) we see similar spatial structures and amplitudes as in Fig. 3.54 (a) and (b). The spatially averaged posterior temperature in Fig. 3.55 (a) is 10.96 °C with a mean 90 % CI from 10.56 °C to 11.36 °C. The corresponding values in (b) are 575.24 mm with a mean 90 % CI from 545.69 mm to 606.36 mm. Hence, overall conditions are 0.49 K warmer and precipitation is 42.77 mm less than in the CMIP reference climatology. The mean values of the grid points over land give a temperature of 10.74 °C and a precipitation amount of 574.02 mm. In contrast to the reconstruction of HIST, relatively high temperatures prevail over land, as can be seen from the positive temperature anomalies in Fig. 3.55 (a). This reduces the difference between the land and the entire map section to 0.22 K. Similar to HIST, there are no meaningful differences in precipitation. For the strongly negative anomalies in Fig. 3.55 (b), we have additionally created a reconstruction that disregards the above-mentioned local information within the mountainous regions. However, this resulted in minor differences in the posterior structures (not shown). Panels (c) - (f) show similar patterns to those of the posterior CMIP reconstruction from Fig. 3.51. These can be summarized as the multi-model CMIP/PMIP ensemble being most constrained where local information are available.

To obtain the posterior patterns of Fig. 3.55, the spatial reconstruction method assigns weights to each PMIP ensemble member depicted in Fig 3.56. In the case of the Middle Holocene, the MPI simulation receives by far the highest weights on average and is well above the prior distribution. In contrast to the results of the previous section, the second highest posterior weights from the GISS experiment are also above the corresponding prior distribution. Thus, the combinations of these two simulations best fit the local reconstruction data. Specifically, MPI shows dipole structures most similar to those in Fig. 3.55 (a) and (b).

Now, we want to compare the results of the MH with other studies. Even if in some cases not the annual temperature but the January temperature, the July temperature or the mean temperature of the coldest/warmest month (MTCO/MTWA) is reconstructed, it should be checked whether the dipole structures shown above is present. The spatial reconstructions of T_{ANN} and P_{ANN} from Mauri et al.,

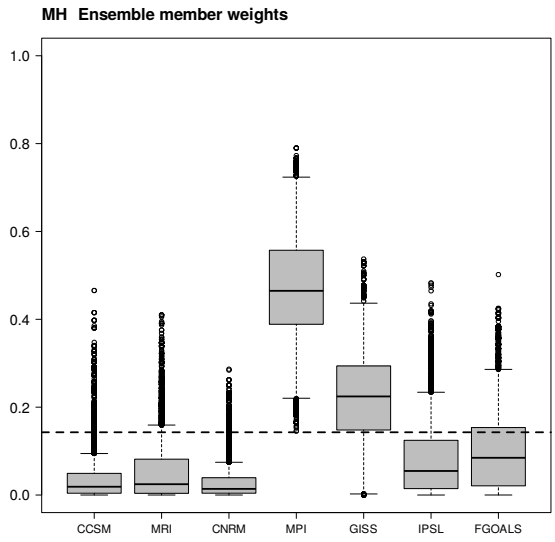


Figure 3.56: As Fig. 3.52, but for Middle Holocene.

2015 reveal a very similar overall pattern to ours. As far as annual precipitation is concerned, the only notable difference is in northwest Africa and eastern Spain. There, in contrast to our work, they reconstruct predominantly positive anomalies. The largest differences in the respective temperature anomalies are found in the southeastern Levant and on the Kola Peninsula in the northeast of the reconstruction area. For both areas Mauri et al., 2015 indicate negative temperature anomalies. In terms of their reconstruction values, both variables have lower amplitudes than ours in Fig. 3.55 (a) and (b). However, looking at the fields of the upper and lower boundaries of the 90 % CI, the results of Mauri et al., 2015 are mostly within these value ranges. As for temperature, there is a difference only over the Kola Peninsula, where we find no negative anomalies in our intervals. Regarding precipitation, the values of Mauri et al., 2015 in the Alpine region and the British Isles are outside our 90 % CI. There our reconstruction shows minima of the anomaly with relatively small interval size (cf. Fig. 3.55 (d)). The plots of the spatial fields of the upper and lower boundaries of the 90 % CI are shown in Appendix B.2. Bartlein et al., 2010 summarize a variety of pollen-based local reconstructions, whereby data from T_{ANN} and P_{ANN} can also be viewed. Their dipole structure and amplitude of T_{ANN} is similar to that of our reconstruction in Fig. 3.55 (a) and (b), with the difference that our maxima are located in northern Fennoscandia and those of Bartlein et al., 2010 in southern Sweden. In terms of precipitation, they also reconstruct negative anomalies of similar magnitude over the western part of the Scandinavian Peninsula. Our results show mainly negative anomalies in Central and Western Europe, which do not become positive within the 90 % CI (cf. Fig. 3.55 (d)). In comparison, Bartlein et al., 2010 reports slightly positive values for these regions.

In addition to our local reconstructions, we can also insert the corresponding local information from

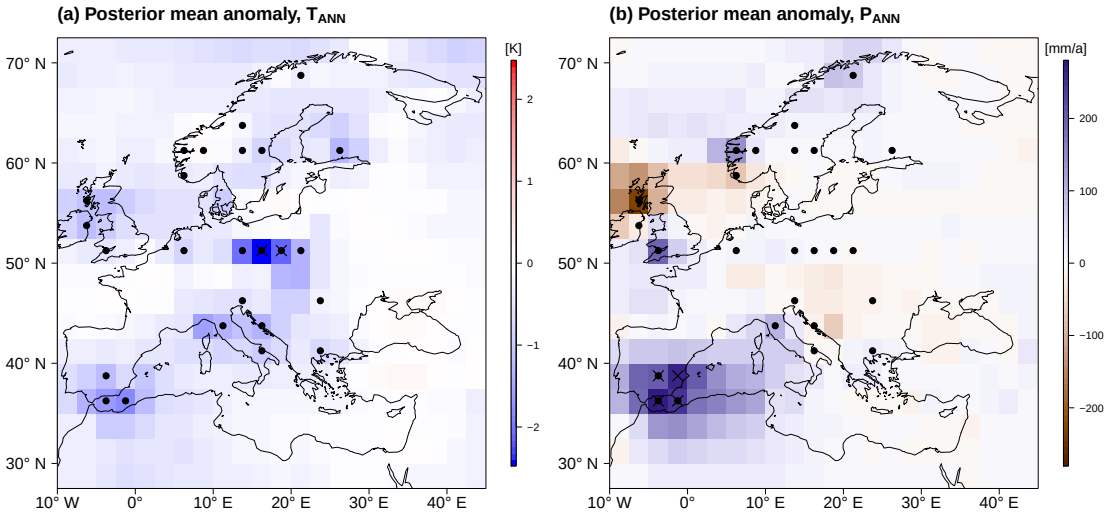


Figure 3.57: Panel (a) shows the mean temperature difference between a spatial reconstruction with additional local information from Simonis et al., 2012 and Fig. 3.55 (a). The grid points with additional information are shown with black dots and those with crosses indicate a significant anomaly. Panel (b) is like (a), but for annual precipitation.

Simonis et al., 2012 for the MH. They present a pollen and macrofossil synthesis as part of the DECVeg project (Dynamic European Climate-Vegetation impacts and interactions). In particular, information on the occurrence of taxa combining macrofossil and pollen data are provided for 50 palaeosites of the MH. We use our new transfer functions for each of these taxa to obtain the probabilities $\mathbb{P}(\boldsymbol{P_s} \mid \boldsymbol{C}, \boldsymbol{\psi})$ (see Sect. 2.5). These are integrated into the Bayesian Indicator Taxa Model (Neumann et al., 2007) to calculate $\mathbb{P}(\boldsymbol{C} \mid \boldsymbol{P_s}, \boldsymbol{\psi})$, which is done for each additional palaeosite. The resulting differences in spatial reconstruction are shown in Fig. 3.57. However, the additional data points from Simonis et al., 2012 have little impact on the posterior distribution of the two variables in northern Europe. This is because the respective local anomalies are already consistent with the patterns of the posterior reconstruction in Fig. 3.55 (a) and (b). Nevertheless, the new information leads to lower posterior temperatures in Central Eastern Europe. Here, we are at the inflection point of the dipole structure (see Fig. 3.55 (a)). Further information from the southeast of the Iberian Peninsula indicates colder, but also wetter conditions. This ultimately leads to a positive precipitation anomaly in the northern Sahara that is close to the reconstruction of Mauri et al., 2015. Overall, the inclusion of the additional data points results in a 0.36 K colder and 18.62 mm wetter reconstruction in contrast to Fig. 3.55 (a) and (b). Analogous to the plots shown here, which are based on our new local reconstructions, others are depicted in Appendix B.1 that additionally include the dataset from Simonis et al., 2012 for the MH.

We now proceed to compare our spatial reconstructions with others based solely on the dataset from Simonis et al., 2012. For example, the July temperature in Simonis et al., 2012 shows positive anomalies in Fennoscandia, the British Isles, France and south-eastern Europe. This means that the

dipole structure is not as clear as it is in our work. In addition, January temperatures in Eastern Europe exhibit positive anomalies. These differences could be due to the different methods. In contrast, the January temperature assimilated in Stolzenberger, 2017 shows a dipole structure with warmer anomalies in northeastern Europe. The corresponding July temperature, on the other hand, does not show quite such clear structures. The spatial patterns of Weitzel et al., 2019 reveal a general dipole structure for MTCO and MTWA that is similar to the annual temperature we reconstructed. By incorporating the local information from Simonis et al., 2012, these are additionally aligned, especially in Central Eastern Europe. As in our case, the PMIP3 experiment of MPI in Weitzel et al., 2019 receives the highest weighting in fitting their simulation ensemble to their local proxy data.

All in all, the spatial reconstruction of the MH reveals that the corresponding method works as intended. The temperature anomaly results show similar structures and magnitudes to many other studies, although the new local data are mostly automatically generated, meaning that our new reconstruction method requires less expert knowledge, such as setting thresholds for occurrence or the importance of specific taxa. In terms of precipitation, there is a strong similarity with Mauri et al., 2015, but also some differences with the local reconstructions in Bartlein et al., 2010. In this respect, the results of our work also show a dipole structure within Europe. Additionally, positive precipitation anomalies are reconstructed in the southwest of the Iberian Peninsula, in northwest Africa and throughout the Eastern Mediterranean, including the Levant.

3.4.3 Last Glacial Maximum

Finally, let us look at the spatial reconstruction of the Last Glacial Maximum (LGM). For this purpose, 15 available local reconstructions with information on 21 cal ka BP are combined with the PMIP3 multi-model ensemble and the corresponding shrinkage parameters are used. The PMIP3 experiments are adapted to the boundary conditions of orbital parameters, greenhouse gas concentrations, ice sheet extent, land surface elevation and coastlines (Sueyoshi et al., 2013). The latter can additionally be seen in Fig. 3.58, where the information of the local reconstructions are shown. These coastlines are based on the PMIP3 boundary condition dataset (Abe-Ouchi et al., 2015).

The temperature and precipitation anomalies of the local data indicate predominantly negative values. In Fig. 3.58 (a), the easternmost grid point is notable for a not significant temperature anomaly. However, the corresponding 90 % CI is of the order of 7 K (see panel (c)). This grid point includes Lake Van, which is located in the mountainous region of the East Anatolia High Plateau. The corresponding local reconstruction is discussed in more detail in Sect. 3.3.4. There, the low variability of annual temperature in the temporal reconstruction process is particularly striking, leading to the low anomaly shown here. In Fig. 3.58 (b) we see higher precipitation in Anatolia and the southern Levant. The latter is not significant and has the lowest 90 % CI, as seen in panel (d). This grid point contains the information about the Dead Sea, which is described in more detail in the Sect. 3.3.1. In contrast, the grid point farther north, which includes Birkat Ram (see Sect. 3.3.3), reveals significant negative

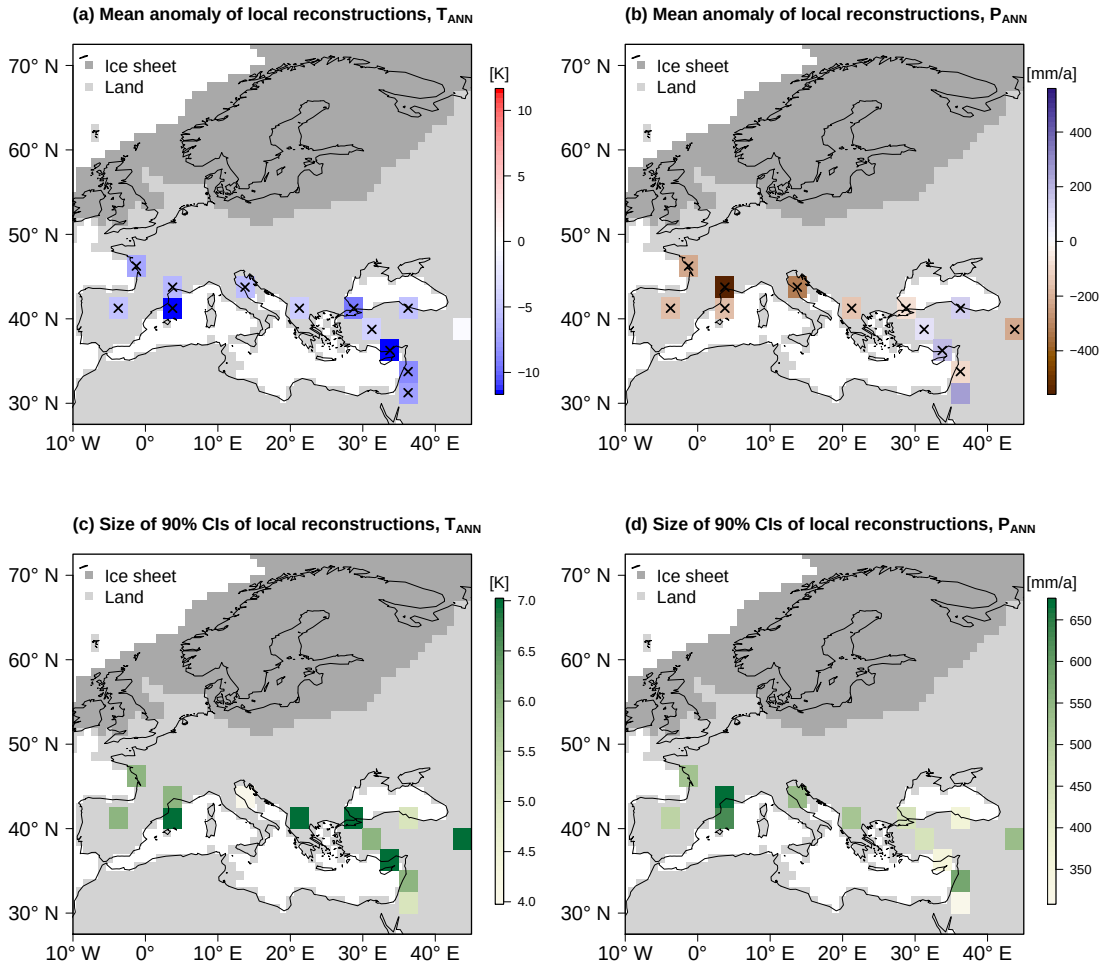


Figure 3.58: As Fig. 3.50, but for the Last Glacial Maximum. Again, the anomaly relates to the CMIP ensemble mean. In addition, the altered ice sheets and land-sea distributions are shown as gray areas.

anomalies. A comparison of our data with those of Bartlein et al., 2010 shows, with the exception of southern Anatolia, a similar pattern and amplitudes of both climate variables. They do not present any reconstruction results for the Levant. In contrast to the time slices of the local information of the MH, Fig. 3.58 (c) and (d) do not show a systematic pattern with respect to the 90 % CI.

The mean posterior anomalies for the LGM are shown in Fig. 3.59 (a) and (b). With respect to the temperature in (a), note that each grid point has a significant negative anomaly. For reasons of clarity, the respective crosses are not additionally marked. The spatially averaged posterior temperature is 0.1 °C with a mean 90 % CI from −1.02 °C to 1.32 °C. The corresponding values in (b) are 423.84 mm with a mean 90 % CI from 364.99 mm to 487.86 mm. Hence, overall conditions are 10.37 K colder and precipitation is 194.17 mm less than in the CMIP reference climatology. Furthermore, we see in

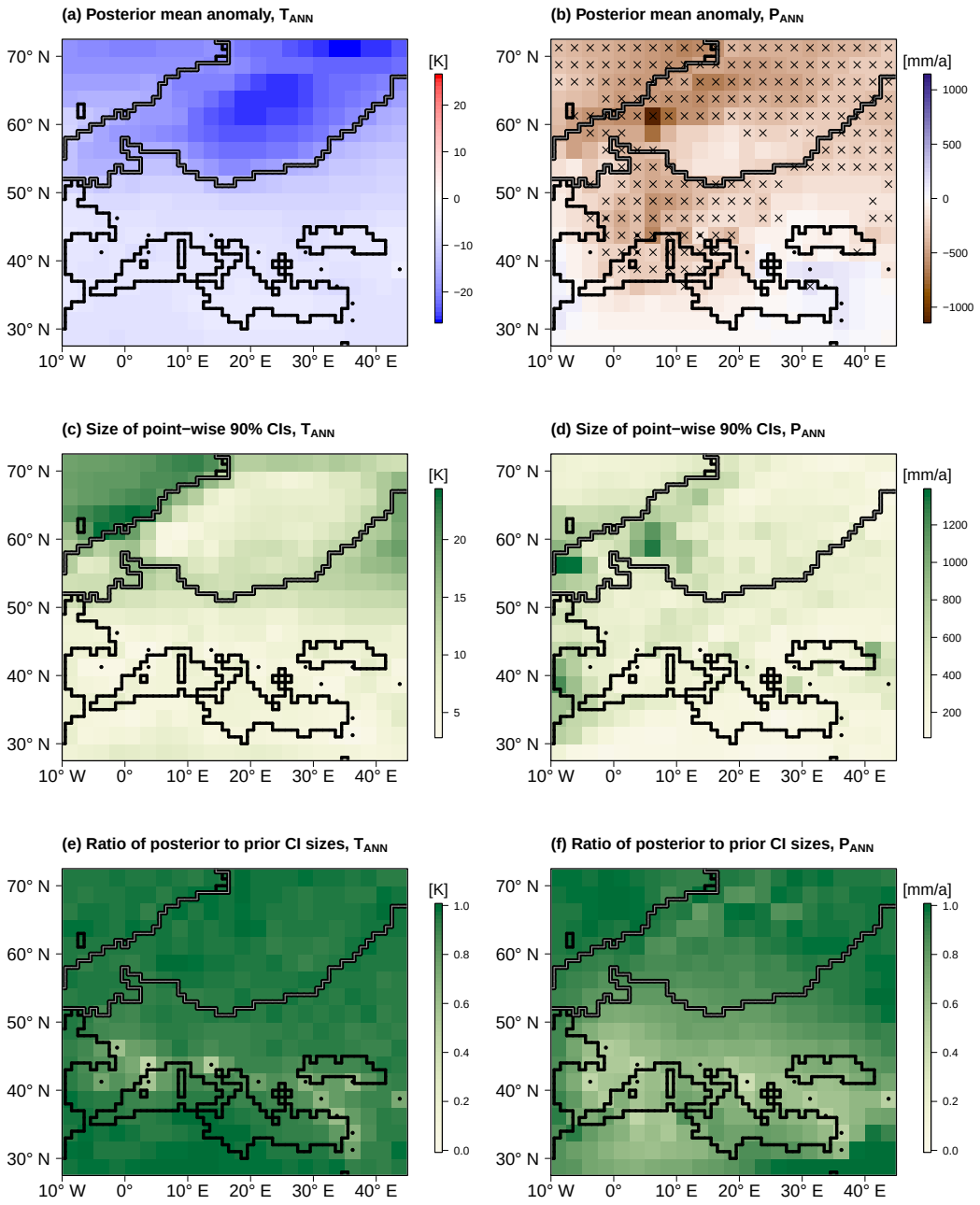


Figure 3.59: As Fig. 3.51, but for the Last Glacial Maximum. Again, the anomaly relates to the CMIP ensemble mean. In addition, contour lines are drawn to indicate modified coastlines (black) and ice sheet boundaries (gray). Furthermore, the crosses indicating significant anomalies are removed in panel (a).

(a) that according to the PMIP experiments the coldest temperatures are above the ice sheet. The grid points with local data show negative anomalies of similar magnitude as in Fig. 3.58 (a). Regarding

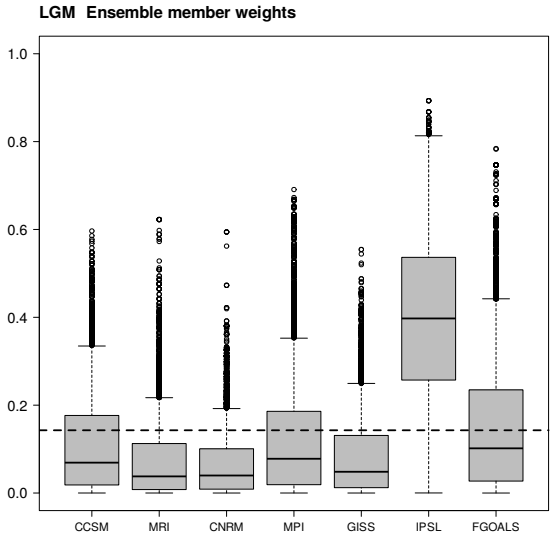


Figure 3.60: As Fig. 3.52, but for Last Glacial Maximum.

precipitation in (b), there are some areas with non-significant deviations from the reference. These are mainly located in the southern reconstruction domain and in the southern parts of the European Ice Sheet. In the latter case, there is relatively much precipitation, so that values similar to those in the recent past are simulated. An exception is the maximum of the negative anomalies. Although precipitation is also relatively high there, it does not come close to today's values in southwest Norway. In contrast, positive precipitation anomalies are simulated on the western edge of the Iberian Peninsula and in northwest Africa. The situation is similar in the Eastern Mediterranean, where the local reconstructions from Fig. 3.58 (b) record higher precipitation. Although the grid point containing Birkat Ram has a significant negative anomaly, these are not visible in Fig. 3.59 (b). This could be due to the relatively coarse spatial resolution, which makes the reconstruction method not flexible enough. However, if only the grid point with Birkat Ram in the area of Anatolia and the Levant is considered, the mean posterior precipitation anomaly in these areas shows exclusively negative deviations (not shown). To reconstruct the anomalies from Fig. 3.59 (a) and (b), we see in Fig. 3.60 that the spatial reconstruction method weights the simulations of IPSL the highest, which is also the only experiment that has higher values than the prior weights. It includes the positive precipitation anomalies in the eastern reconstruction area. The uncertainty distributions in Fig. 3.59 (c) and (d) indicate the highest values with respect to temperature northwest and east of the ice sheet. This is where the PMIP experiments of the LGM have the greatest variability. These are located in (d) in the areas with the highest precipitation. Panels (e) and (f) show where the local data most constrain the prior PMIP ensemble.

To analyse the positive precipitation anomalies from Fig. 3.59 (b), we compare them with the

experiments of Ludwig et al., 2018 and Ludwig and Hochman, 2022. In both studies, the PMIP3 experiment MPI served as the initial and boundary condition of the LGM for a dynamic downscaling with the regional weather research and forecast model (WRF). In this way, Ludwig et al., 2018 find predominantly positive precipitation anomalies in the western part of the Iberian Peninsula, which is consistent with our results. Conversely, a local reconstruction in Bartlein et al., 2010 in southern Andalusia still shows slightly negative anomalies. We therefore conclude that additional data points and higher spatial resolution could provide more insight into this area. In Ludwig and Hochman, 2022, the focus is on the Levant and it is noted that overall evaporation has decreased more than precipitation. We have already discussed this issue in detail in Sect. 3.3.1 and Sect. 3.3.3, where our local reconstructions of the Levant during the LGM show lower precipitation compared to the interglacials. This is also consistent with a pollen-based reconstruction from Lake Hula, located about 20 km southwest of Birkat Ram. In this context, Langgut et al., 2021 note 5 K colder conditions and 115 mm less precipitation compared to present-day climate. In contrast, the simulations from Ludwig and Hochman, 2022 show that precipitation increased in the southern and eastern areas of the Levant during the LGM. They assume a connection with the so-called active Red Sea Trough, which may have led to increased flood events in the Dead Sea region within this period. Moreover, Cheng et al., 2015 compare speleothems in the northern and southern Levant and find contrasting patterns as well. From this, they conclude that comparatively drier conditions prevailed in the north during the period from the LGM to Heinrich stadial 1. Such a dipole structure of precipitation in the Levant during the LGM is consistent with local information from the Dead Sea, Lake Hula, and Birkat Ram (cf. Fig. 3.58 (b)). However, these are not shown in the spatial reconstruction in Fig. 3.59 (b). Looking at the precipitation field at the lower bound of the 90 % CI, a slight dipole structure can be seen, with only the grid point containing the Dead Sea showing a positive anomaly. This pattern could be better captured with a higher resolution of the spatial reconstruction method.

In summary, this method works as intended. This time, however, it is noticeable that a higher spatial resolution would be advantageous with regard to the Levant in order to be able to analyse the dipole structure of the precipitation more precisely. The situation is similar for the areas of the western Iberian Peninsula and northwest Africa, where positive precipitation anomalies are also simulated. Since we have relatively little local data in this area, further reconstructions would be particularly useful. Apart from that, the results are largely consistent with those of other studies.

CHAPTER 4

Summary, Conclusion and Outlook

In this work, we present new techniques for generating local paleoclimate reconstructions based on multiple proxies. For this purpose, we use a newly developed BHM solved with MCMC sampling methods such as Metropolis-within-Gibbs or random walk Metropolis-Hastings. To place the various proxies in a temporal context, a new method is used to assign age information to depths in sediments and distances in speleothems. In particular, the uncertainty of age is accounted for by another BHM introduced in this thesis. Climate variables such as annual temperature and precipitation were included using transfer functions based on plant occurrence. We determined these functions with a machine learning competition. Such a systematic identification of the best model to describe the relationship between plants and climate is made here for the first time. Finally, these local scale plant information from lake and mire sediments and speleothems are matched with global scale isotopic data from marine sediments and ice cores. As a result, over 600 sites in Europe, Northwest Africa, Anatolia, and the Levant are being processed, resulting in 186 selected local climate reconstructions using these new reconstruction methods. On this basis, we additionally calculate spatial reconstruction fields of the above mentioned regions using the CMIP5/PMIP3 experiments from HIST, MH, and LGM. In this context, these ESMs are linked to our local information, resulting in assimilated spatial climate fields. The corresponding BHM is a modified version of the BHM of Weitzel et al., 2019, which is solved using a random walk Metropolis-Hastings MCMC technique.

A basis for the calculation of the local reconstructions in this work is the age-depth/distance transformation. This is considered in a Bayesian formulation by determining the uncertainty information of the depth or distance at a given age $\mathbb{P}(\boldsymbol{D} \mid \boldsymbol{A})$. In this way, we can perform a data-driven transformation of the local reconstructions that behaves like a convolution with different kernel smoothers. This approach avoids too much certainty in statements about past events. Another result of this technique is the determination of the age resolution and its projection onto a regular grid. This allows us to match proxy information from the EPD and SISAL databases and perform spectral

analysis.

Another step in computing our local reconstructions is to establish transfer functions that map climate variables to plant distributions. Various machine learning algorithms are used to solve this classification problem of absence and presence of taxa. For this purpose, 70 % of the respective imbalanced datasets are first balanced with SMOTE. To test which method works best, the respective performances are identified with a cross-validation. In addition, a number of different parameter settings of the respective algorithms are made. The winner of these tuned models is validated on the remaining 30 % unsmoted hold-out datasets. For the models and plant distributions used in this work, a simple feedforward neural network with one hidden layer wins in 70 % of the cases. Each of the machine learning algorithms considered takes into account the uncertainties in the proxy-climate relationship to avoid the risk of making overly confident inferences about past climate.

The subsequent more than 600 MCMC simulations for MG and rwMH assign weights to the taxa considered. The higher these turn out during the inference process, the more important is their inclusion in the final reconstruction, taking into account additional proxy information. These are based on isotopic information from speleothems, marine sediments, and ice cores. The respective reconstruction should approximate these reference curves to a certain degree. This is verified not only in temporal space but also in spectral space using wavelet power spectra. Such a comprehensive use of the spectral behavior of proxy information is possible due to the new age-depth/distance transformation and has therefore not been performed before. In addition, a priori information on the actual climate distribution in specific time periods are incorporated as further constraints. Specifically, anchor points of recent climate are often used to ensure that local reconstructions approach them. Overall, both MCMC techniques show similar flexible and stable results when all simulations are considered.

The local reconstructions discussed in this study show realistic results if the corresponding qualitative reconstructions of Aufgebauer et al., 2012, Panagiotopoulos et al., 2013, Litt et al., 2012, Chen and Litt, 2018, Miebach et al., 2019, Schiebel and Litt, 2018, Miebach et al., 2022, Schiebel, 2013, Neumann et al., 2007, Pickarski et al., 2015b, Pickarski et al., 2015a, Litt et al., 2009, Miebach et al., 2016, and Panagiotopoulos et al., 2020 are used as a comparison. Relatively small-scale climate changes are simulated, which can be identified as Bond, Heinrich, and Dansgaard-Oeschger events. In addition, large-scale changes such as the last deglacialization and different glacial-interglacial cycles can also be detected. During certain time periods, human influence on plant information in lake and mire sediments is clearly evident. This thus also affects the local reconstructions, where additional climate variations are simulated. These can be minimized to some extent with our new method by giving more attention to isotope-based proxies during the inference process.

The final 186 accepted local paleoclimate reconstructions are summarized using a modified version of the spatial reconstruction method of Weitzel et al., 2019. In this context, the posterior spatial climate distributions of the reference period show that the prior multi-model ensemble of CMIP5 experiments, in combination with the local data fitted to the current climate, converges toward the

CRU fields. This applies not only to temperature, but also to precipitation. Moreover, the areas with more local information show a closer approximation of the prior ensemble to the CRU reference period, indicating sufficient flexibility of this reconstruction method. This also holds true for the reconstructions of the MH and the LGM. The former shows warming in the northeast and cooling in southern Europe according to local data, which is similar to Stolzenberger, 2017, Weitzel et al., 2019, Mauri et al., 2015, and Bartlein et al., 2010. Annual precipitation indicates an increase in the northeast and a decrease in the west of Europe. In addition, we note an increase in precipitation in the Eastern Mediterranean and Levant, which is also true for Mauri et al., 2015. The spatial reconstruction of the LGM reveals an increase in precipitation in the west of the Iberian Peninsula and in Northwest Africa, Anatolia, and the Levant. The latter shows a dipole structure with higher precipitation in the southern Levant, which is consistent with the work of Ludwig and Hochman, 2022 and the interpretations of Cheng et al., 2015. T_{ANN} exhibits no particular pattern during this time interval, only a cooling of the entire reconstruction area.

Overall, it appears that our new method provides an additional way to calculate quantitative paleoclimate reconstructions. Comparisons with qualitative and other quantitative studies show realistic and to some extent similar results. We therefore conclude that more automatic, statistics-based methods are an alternative to those that require additional expert knowledge. Furthermore, our method provides complementary information such as the importance of taxa with corresponding uncertainty estimates. From this, we can gain new insights into possible biological mechanisms behind past climate changes.

Therefore, an examination of important taxa and why they lead to certain patterns at certain ages could be useful for future work. This might be done by an additional BHM based on the local reconstructions produced in this thesis. Considering these as constraints, one could focus on specific time periods and analyze the resulting taxa weights in more detail. Finally, using all available local reconstructions, this procedure would yield a spatial distribution of the most important taxa for specific time slices.

In addition, the techniques presented in this work could be performed at higher resolutions. This might be done by replacing the CRU data used for the transfer functions and climate anchor points with regional reanalysis data such as COSMO-REA6 (Bollmeyer et al., 2015). On the other hand, it could be helpful to increase the number of reference curves used. Examples of additional proxies are lake level curves, CO_2 data, and solar radiation information. Since not only reference curves but also climate distributions were used as boundary conditions, such additional information could also be helpful. Examples would include various simulation results from ESMs or available permafrost information on the sites in question. Moreover, a larger number of taxa considered would allow higher flexibility of our MCMC methods and provide a more comprehensive view of past climate variability. Performing spatial reconstructions based on the more recent CMIP6/PMIP4 experiments could also provide new understanding. It would also allow the preparation of climate reconstructions for the Last Interglacial

period. Furthermore, simulation of additional variables such as evaporation could be beneficial. This would provide more information on dry periods and areas. Machine learning techniques could be used to test the importance of these additional climate variables for each plant distribution. For comparison with reference curves, other metrics besides explained variance can provide useful additional findings. An example of this is the so-called target redundance of Glowienka-Hense et al., 2020. This is the information that, for example, two reference curves provide equally. In terms of our age-depth/distance transformation, an attempt could be made to apply this technique to existing proxies such as Medstack, LR04, and NGRIP. In particular, for stacked data, the alignment of the underlying time series may provide new insights.

Analogous to transfer functions based on plant data applied to the appropriate sites, one could attempt to do this for speleothem data currently being collected. In the course of this work, initial tests have already been conducted with the SISAL database. However, we found no relationship between the geographic coordinates, altitude, and geologic characteristics of the caves, the temporal characteristics of the growth of the individual dripstones, and the various climatic variables related to temperature and precipitation. This may change with the steady expansion of speleothem data in future studies.

It would also be interesting to try to use our method of local reconstruction to detect human activity where little is known about past civilizations. This could be done by analyzing the spectral and spatiotemporal patterns of known activities. If the differences between the plant data and the other proxies stand out in a similar way during certain time periods, this could be an indication of human activity.

In summary, the reconstruction methods presented in this work represent an attempt to solve complex BHMs with low computational cost, capable of considering multiple proxy sources. This can not only diagnose past climate changes, but also provide a better understanding of its mechanisms. Extending this method and applying it to other geographic areas could further enhance this knowledge.

Bibliography

Abe-Ouchi, Ayako et al. (2015),
Ice-sheet configuration in the CMIP5/PMIP3 Last Glacial Maximum experiments,
Geoscientific Model Development Discussions **8**, p. 4293 (cit. on pp. 46, 123, 134).

Agassiz, Louis (2012), *Études sur les glaciers*, Cambridge Library Collection - Earth Science,
Cambridge University Press (cit. on p. 7).

Agresti, A. (2003), *Categorical Data Analysis*, Wiley Series in Probability and Statistics, Wiley
(cit. on p. 42).

Akers, Pete D. et al. (2016), *An extended and higher-resolution record of climate and land use from stalagmite MC01 from Macal Chasm, Belize, revealing connections between major dry events, overall climate variability, and Maya sociopolitical changes*,
Palaeogeography, Palaeoclimatology, Palaeoecology **459**, p. 268 (cit. on p. 24).

Alvarez-Solas, Jorge et al. (2011),
Heinrich event 1: An example of dynamical ice-sheet reaction to oceanic changes,
Climate of the Past **7**, p. 1297 (cit. on p. 9).

Ammann, Caspar and Eugene Wahl (2007), *Importance of the geophysical context for statistical evaluation of climate reconstruction procedures*, Clim. Change **85**, p. 71 (cit. on p. 45).

Andersen, Katrine et al. (2004),
High-resolution record of Northern Hemisphere climate extending into the last interglacial period,
Nature **431**, p. 147 (cit. on p. 9).

Arrigo, Rosanne, Rob Wilson and Gordon Jacoby (2006),
On the long-term context for late twentieth century warming, Journal of Geophysical Research **111**
(cit. on p. 45).

Asmerom, Y., V. Polyak and S. Burns (2010),
Variable winter moisture in the southwestern United States linked to rapid glacial climate shifts,
Nature Geoscience **3**, p. 114 (cit. on p. 38).

Atsawawaranunt, K. et al. (2018), *The SISAL database: a global resource to document oxygen and carbon isotope records from speleothems*, Earth System Science Data **10**.3, p. 1687 (cit. on p. 15).

Aufgebauer, Anne et al. (2012), *Climate and environmental change in the Balkans over the last 17 ka recorded in sediments from Lake Prespa (Albania/F.Y.R. of Macedonia/Greece)*, Quaternary International **274**, p. 122, Temporal and spatial corridors of Homo sapiens sapiens population dynamics during the Late Pleistocene and Early Holocene (cit. on pp. 57, 58, 66, 67, 140).

Ayer, Turgay et al. (2010), *Comparison of Logistic Regression and Artificial Neural Network Models in Breast Cancer Risk Estimation*, RadioGraphics **30**.1, p. 13, PMID: 19901087 (cit. on p. 42).

Bar-Matthews, M., A. Ayalon, A. Vaks and A. Frumkin (2017), "Climate and Environment Reconstructions Based on Speleothems from the Levant", *Quaternary of the Levant: Environments, Climate Change, and Humans*, ed. by Yehouda Enzel and OferEditors Bar-Yosef, Cambridge University Press, p. 151 (cit. on p. 83).

Bar-Matthews, Miryam (2013), *Dansgaard-Oeschger events and their reflection in speleothems (Hans Oeschger Medal Lecture)*, p. 14186 (cit. on p. 16).

Bar-Matthews, Miryam and Avner Ayalon (2007), "Speleothems as palaeoclimate indicators, a case study from Soreq Cave located in the Eastern Mediterranean Region, Israel", p. 363 (cit. on pp. 17, 72, 76).

– (2011), *Mid-Holocene climate variations revealed by high-resolution speleothem records from Soreq Cave, Israel and their correlation with cultural changes*, Holocene **21**, p. 163 (cit. on pp. 17, 72, 80, 96).

Bar-Matthews, Miryam, Avner Ayalon, Mabs Gilmour, Alan Matthews and Chris J. Hawkesworth (2003), *Sea–land oxygen isotopic relationships from planktonic foraminifera and speleothems in the Eastern Mediterranean region and their implication for paleorainfall during interglacial intervals*, Geochimica et Cosmochimica Acta **67**.17, p. 3181, A Special Issue Dedicated to Robert Clayton (cit. on pp. 16, 38, 73, 77, 91, 100).

Bar-Matthews, Miryam, Avner Ayalon and Aaron Kaufman (1997), *Late Quaternary Paleoclimate in the Eastern Mediterranean Region from Stable Isotope Analysis of Speleothems at Soreq Cave, Israel*, Quaternary Research **47**.2, p. 155 (cit. on pp. 14, 16).

Bar-Matthews, Miryam, Avner Ayalon, Aaron Kaufman and Gerald J Wasserburg (1999),
The Eastern Mediterranean paleoclimate as a reflection of regional events: Soreq cave, Israel,
Earth and Planetary Science Letters **166**.1, p. 85 (cit. on pp. 15, 17, 75).

Bar-Matthews, Miryam, Avner Ayalon and Aharon Kaufman (1998),
"Middle to Late Holocene (6,500 Yr. Period) Paleoclimate in the Eastern Mediterranean Region from Stable Isotopic Composition of Speleothems from Soreq Cave, Israel",
Water, Environment and Society in Times of Climatic Change: Contributions from an International Workshop within the framework of International Hydrological Program (IHP) UNESCO, held at Ben-Gurion University, Sede Boker, Israel from 7–12 July 1996,
ed. by Arie S. Issar and Neville Brown, Dordrecht: Springer Netherlands, p. 203 (cit. on p. 18).

Bartlein, Patrick et al. (2010),
Pollen-based continental climate reconstructions at 6 and 21 ka: A global synthesis,
Climate Dynamics **37**, p. 775 (cit. on pp. 5, 132, 134, 135, 138, 141).

Bayes, Mr. and Mr. Price (1763),
An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.,
Philosophical Transactions (1683-1775) **53**, p. 370 (cit. on p. 22).

Benson, Larry V. et al. (1996), *Climatic and Hydrologic Oscillations in the Owens Lake Basin and Adjacent Sierra Nevada, California*, Science **274**.5288, p. 746 (cit. on p. 38).

Bernhardi, Albrecht Reinhard (1832),
Wie kamen die aus dem Norden stammenden Felsbruchstücke und Geschiebe, welche man in Norddeutschland und den benachbarten Ländern findet, an ihre gegenwärtigen Fundorte?,
Jahrbuch für Mineralogie, Geognosie, Geologie und Petrefaktenkunde **3**, p. 257 (cit. on p. 7).

Betancourt, Michael (2017), *The Convergence of Markov Chain Monte Carlo Methods: From the Metropolis Method to Hamiltonian Monte Carlo*, Annalen der Physik **531** (cit. on p. 24).

Blaauw, Maarten (2019), *clam: Classical Age-Depth Modelling of Cores from Deposits*,
R package version 2.3.2 (cit. on p. 22).

Blaauw, Maarten and J. Andrés Christen (2011),
Flexible paleoclimate age-depth models using an autoregressive gamma process,
Bayesian Anal. **6**.3, p. 457 (cit. on pp. 22, 24).

Blaauw, Maarten, J. Andres Christen and Marco A. Aquino L. (2020),
rbacon: Age-Depth Modelling using Bayesian Statistics, R package version 2.4.0 (cit. on p. 24).

Bohn, Udo, Nugzar Zazanashvili and George Nakhutsrishvili (2007),
The Map of the Natural Vegetation of Europe and its application in the Caucasus Ecoregion,
Bulletin of the Georgia Academy of Science **175**, p. 112 (cit. on pp. 66, 120).

Bollmeyer, C. et al. (2015),
Towards a high-resolution regional reanalysis for the European CORDEX domain,
Quarterly Journal of the Royal Meteorological Society **141**.686, p. 1 (cit. on p. 141).

Bond, G.C. et al. (2002), *Persistent Solar Influence on North Atlantic Climate During the Holocene*,
Science (New York, N.Y.) **294**, p. 2130 (cit. on pp. 18, 76).

Bond, Gerard et al. (1997),
A Pervasive Millennial-Scale Cycle in North Atlantic Holocene and Glacial Climates,
Science **278**.5341, p. 1257 (cit. on p. 9).

Bond, Gerard C. and Rusty Lotti (1995),
Iceberg Discharges into the North Atlantic on Millennial Time Scales During the Last Glaciation,
Science **267**.5200, p. 1005 (cit. on pp. 3, 9).

Bosmans, Joyce, Frits Hilgen, E. Tuenter and Lucas Lourens (2015),
Obliquity forcing of low-latitude climate, Climate of the Past **11**, p. 1335 (cit. on p. 107).

Bowyer, Kevin W., Nitesh V. Chawla, Lawrence O. Hall and W. Philip Kegelmeyer (2011),
SMOTE: Synthetic Minority Over-sampling Technique, CoRR **abs/1106.1813** (cit. on p. 41).

Braconnot, P. et al. (2011),
The Paleoclimate Modeling Intercomparison Project contribution to CMIP5,
CLIVAR Exchanges **56**, p. 15 (cit. on p. 45).

Bradley, Raymond (2014), *Paleoclimatology (3rd edition)* (cit. on pp. 1, 3, 7, 8, 10, 14, 21, 24, 38).

Breiman, Leo (2001), *Random Forests*, Machine Learning **45**.1, p. 5 (cit. on p. 44).

Brodersen, Kay H., Cheng Soon Ong, Klaas Stephan and Joachim Buhmann (2010),
The Balanced Accuracy and Its Posterior Distribution,
Pattern Recognition, International Conference on **0**, p. 3121 (cit. on p. 50).

Brooks, Stephen and Andrew Gelman (1998),
General Methods for Monitoring Convergence of Iterative Simulations,
J. Comput. Graphi. Stat. **7**, p. 434 (cit. on p. 53).

Bubenzer, Olaf and Ulrich Radtke (2007), “Natürliche Klimaänderungen im Laufe der Erdgeschichte”,
ed. by Wilfried Endlicher and Friedrich-Wilhelm Gerstengarbe, Humboldt-Universität zu Berlin,
Mathematisch-Naturwissenschaftliche Fakultät II, Geographisches Institut, p. 17 (cit. on p. 7).

Budsky, Alexander et al. (2019), *Western Mediterranean Climate Response to Dansgaard/Oeschger Events: New Insights From Speleothem Records*, Geophysical Research Letters **46**.15, p. 9042 (cit. on p. 16).

Butcher, Brandon and Brian J. Smith (2020), *Feature Engineering and Selection: A Practical Approach for Predictive Models*, The American Statistician **74**.3, p. 308 (cit. on p. 40).

Chen, Chunzhu and Thomas Litt (2018), *Dead Sea pollen provides new insights into the paleoenvironment of the southern Levant during MIS 6–5*, Quaternary Science Reviews **188**, p. 15 (cit. on pp. 80–84, 88, 89, 140).

Chen, Deliang and Hans Weiteng Chen (2013), *Using the Köppen classification to quantify climate variation and change: An example for 1901–2010*, Environmental Development **6**, p. 69 (cit. on p. 70).

Cheng, H. et al. (2015), *The climate variability in northern Levant over the past 20,000 years*, Geophysical Research Letters **42**.20, p. 8641 (cit. on pp. 3, 15, 16, 18, 38, 138, 141).

Chevalier, M., R. Cheddadi and B. M. Chase (2014), *CREST (Climate REconstruction SofTware): a probability density function (PDF)-based quantitative climate reconstruction method*, Climate of the Past **10**.6, p. 2081 (cit. on p. 40).

Chevalier, Manuel et al. (2020), *Pollen-based climate reconstruction techniques for late Quaternary studies*, Earth-Science Reviews **210** (cit. on pp. 19, 20, 30, 31).

Chicco, Davide, Matthijs J. Warrens and Giuseppe Jurman (2021), *The Matthews correlation coefficient (MCC) is more informative than Cohen's kappa and Brier score in binary classification assessment*, English, IEEE Access **9**, p. 78368 (cit. on p. 50).

Christen, J. and Colin Fox (2010), *A General Purpose Sampling Algorithm for Continuous Distributions (the t-walk)*, Bayesian Analysis **5** (cit. on p. 53).

Colleoni, Florence, Simona Masina, Alessandra Negri and Alice Marzocchi (2012), *Plio–Pleistocene high–low latitude climate interplay: A Mediterranean point of view*, Earth and Planetary Science Letters **319-320**, p. 35 (cit. on pp. 58, 59, 80, 81, 91).

Constantin, Silviu, Ana-Voica Bojar, Stein-Erik Lauritzen and Joyce Lundberg (2007), *Holocene and Late Pleistocene climate in the sub-Mediterranean continental environment: A speleothem record from Poleva Cave (Southern Carpathians, Romania)*, Palaeogeography, Palaeoclimatology, Palaeoecology **243**.3, p. 322 (cit. on p. 16).

Cook, Edward R., David M. Meko, David W. Stahle and Malcolm K. Cleaveland (1999), *Drought Reconstructions for the Continental United States*, Journal of Climate **12**.4, p. 1145 (cit. on p. 45).

Coplen, Tyler B. (1996), *New guidelines for reporting stable hydrogen, carbon, and oxygen isotope-ratio data*, Geochimica et Cosmochimica Acta **60**.17, p. 3359 (cit. on p. 8).

Daniel Simonis (2010), *Reconstruction of possible realisations of the Late Glacial and Holocene near surface climate in Central Europe* (cit. on p. 40).

Dansgaard, W. (1964), *Stable isotopes in precipitation*, Tellus **16**.4, p. 436 (cit. on pp. 8, 9).

Davis, B.A.S., S. Brewer, A.C. Stevenson and J. Guiot (2003), *The temperature of Europe during the Holocene reconstructed from pollen data*, Quaternary Science Reviews **22**.15, p. 1701 (cit. on p. 45).

De'ath, Glenn (2007), *Boosted Trees for Ecological Modeling and Prediction*, Ecology **88**, p. 243 (cit. on p. 44).

Debret, M. et al. (2007), *The origin of the 1500-year climate cycles in Holocene North-Atlantic records*, Climate of the Past **3**.4, p. 569 (cit. on pp. 13, 18).

Deininger, Michael et al. (2016), *Coherency of late Holocene European speleothem δ18O records linked to North Atlantic Ocean circulation*, Climate Dynamics **49**, p. 595 (cit. on p. 16).

Demény, Attila et al. (2021), *Holocene hydrological changes in Europe and the role of the North Atlantic ocean circulation from a speleothem perspective*, Quaternary International **571**, p. 1 (cit. on p. 16).

DiVenere, Dr. Vic (2017), *Stable Isotopes*, `http://www.columbia.edu/~vjd1/stable_isotopes.htm`, Accessed: 2022-08-01 (cit. on p. 10).

Dokken, Trond M., Kerim H. Nisancioglu, Camille Li, David S. Battisti and Catherine Kissel (2013), *Dansgaard-Oeschger cycles: Interactions between ocean and sea ice intrinsic to the Nordic seas*, Paleoceanography **28**.3, p. 491 (cit. on p. 9).

Drăguşin, V. et al. (2014), *Constraining Holocene hydrological changes in the Carpathian–Balkan region using speleothem δ^{18}O and pollen-based temperature reconstructions*, Climate of the Past **10**.4, p. 1363 (cit. on pp. 16, 24–27).

Edwards, Kevin, Ralph Fyfe and Stephen Jackson (2017), *The first 100 years of pollen analysis*, Nature Plants **3**, p. 17001 (cit. on p. 19).

Fahrmeir, Ludwig and Gerhard Tutz (1994), *Multivariate statistical modelling based on generalized linear models / Ludwig Fahrmeir, Gerhard Tutz*, SERBIULA (sistema Librum 2.0) (cit. on p. 42).

Fairchild, I. et al. (2006), *Modification and preservation of environmental signals in speleothems*, Earth-Science Reviews **75**, p. 105 (cit. on pp. 14, 18).

Fleitmann, Dominik et al. (2003), *Holocene Forcing of the Indian Monsoon Recorded in a Stalagmite from Southern Oman*, Science (New York, N.Y.) **300**, p. 1737 (cit. on p. 16).

Freund, Yoav and Robert E. Schapire (1996), *Experiments with a New Boosting Algorithm*, ICML'96, p. 148 (cit. on p. 44).

Friedman, Jerome, Trevor Hastie and Robert Tibshirani (2000), *Additive logistic regression: a statistical view of boosting*, Ann. Statist. **28**.2, p. 337 (cit. on p. 44).

Friedman, Jerome H. (2001), *Greedy function approximation: A gradient boosting machine.*, The Annals of Statistics **29**.5, p. 1189 (cit. on pp. 30, 44).

Fyfe, Ralph et al. (2009), *The European Pollen Database: Past efforts and current activities*, Vegetation History and Archaeobotany **18**, p. 417 (cit. on p. 21).

Gams, Helmut (1964), "Systematische und genetische Pflanzengeographie", *Bericht über das Jahr 1963*, ed. by Erwin Bünning and Heinz Ellenberg, Berlin, Heidelberg: Springer Berlin Heidelberg, p. 94 (cit. on p. 38).

Gebhardt, Christoph, Norbert Kühl, Andreas Hense and Thomas Litt (2004), "Multi-Scale Processes and the Reconstruction of Palaeoclimate", vol. 97, p. 325 (cit. on p. 40).

– (2008), *Reconstruction of Quaternary temperature fields by dynamically consistent smoothing*, Climate Dynamics **30**.4, p. 421 (cit. on pp. 4, 22, 30, 39, 45).

Gelman, A. et al. (2013), *Bayesian Data Analysis, Third Edition*, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis (cit. on pp. 23, 36, 53).

Gentle, James E. (2004), *Random Number Generation and Monte Carlo Methods*, 2nd (cit. on p. 40).

Genty D., et al (2003), *Precise Dating of Dansgaard-Oeschger Climate Oscillations in Western Europe from Stalagmite Data*, Nature **421**, p. 833 (cit. on pp. 15, 38).

Giesecke, T. et al. (2013), *Towards mapping the late Quaternary vegetation change of Europe*, Vegetation History and Archaeobotany **23**, p. 75 (cit. on pp. 21, 22).

Giesecke, Thomas, Simon Brewer, Walter Finsinger, Michelle Leydet and Richard H.W. Bradshaw (2017), *Patterns and dynamics of European vegetation change over the last 15,000 years*, Journal of Biogeography **44**.7, p. 1441 (cit. on pp. 56, 67, 68).

Glowienka-Hense, R., A. Hense, S. Brune and J. Baehr (2020), *Comparing forecast systems with multiple correlation decomposition based on partial correlation*, Advances in Statistical Climatology, Meteorology and Oceanography **6**.2, p. 103 (cit. on pp. 34, 142).

Göppert, H. R. (1836), *Über den Zustand, in welchem sich die fossilen Pflanzen befinden, und über den Versteinerungsprocess insbesondere*, Annalen der Physik **114**.8, p. 561 (cit. on p. 19).

Goring, Simon J. et al. (2016), *Novel and Lost Forests in the Upper Midwestern United States, from New Estimates of Settlement-Era Composition, Stem Density, and Biomass*, PLOS ONE **11**.12, p. 1 (cit. on pp. 113, 120).

Grant, K. M. et al. (2012), *Rapid coupling between ice volume and polar temperature over the past 150,000 years*, Nature **491**, p. 744 (cit. on p. 81).

Greenwell, Brandon, Bradley Boehmke, Jay Cunningham and GBM Developers (2019), *gbm: Generalized Boosted Regression Models*, R package version 2.1.5 (cit. on p. 49).

Grichuk, V.P. (1969), *An attempt to reconstruct certain elements of the climate of the northern hemisphere in the Atlantic period of the Holocene.*, Proceedings of the 3rd International Palynological Congress, p. 41 (cit. on p. 29).

Harris, I. and P.D. Jones (2017), *CRU TS4.01: Climatic Research Unit (CRU) Time-Series (TS) version 4.01 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2016)*, (cit. on pp. 38, 58, 70, 71).

Haslett, John and Andrew C Parnell (2008), *A simple monotone process with application to radiocarbon-dated depth chronologies*, Journal of the Royal Statistical Society: Series C (Applied Statistics) **57**.4, p. 399 (cit. on p. 25).

Hastie, T., R. Tibshirani and J.H. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer series in statistics, Springer (cit. on pp. 44, 45).

Havinga, A.J. (1967), *Palynology and pollen preservation*, Review of Palaeobotany and Palynology **2**.1, p. 81 (cit. on p. 20).

Hazan, N. et al. (2005),
The late Quaternary limnological history of Lake Kinneret (Sea of Galilee), Israel,
Quaternary Research **63**.1, p. 60 (cit. on pp. 90, 96).

Hegerl, Gabriele and Francis Zwiers (2011),
Use of models in detection and attribution of climate change,
Wiley Interdisciplinary Reviews: Climate Change **2**, p. 570 (cit. on p. 45).

Heinrich, Hartmut (1988), *Origin and Consequences of Cyclic Ice Rafting in the Northeast Atlantic Ocean During the Past 130,000 Years*, Quaternary Research **29**, p. 142 (cit. on p. 9).

Hemming, Sidney R. (2004), *Heinrich events: Massive late Pleistocene detritus layers of the North Atlantic and their global climate imprint*, Reviews of Geophysics **42**.1 (cit. on pp. 81, 82, 84).

Hendy, C.H (1971), *The isotopic geochemistry of speleothems—I. The calculation of the effects of different modes of formation on the isotopic composition of speleothems and their applicability as palaeoclimatic indicators*, Geochimica et Cosmochimica Acta **35**.8, p. 801 (cit. on pp. 13, 14).

Hense, A. (2005), *2 Processing of observational data and its implication for climate analysis: Datasheet from Landolt-Börnstein - Group V Geophysics · Volume 6: "Observed Global Climate" in SpringerMaterials (https://doi.org/10.1007/10730550_3)*, ed. by M. Hantel (cit. on p. 1).

Hersbach, Hans (2000),
Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems,
Weather and Forecasting **15**.5, p. 559 (cit. on p. 57).

Hill, Theodore (2008), *Conflations of Probability Distributions*,
Transactions of the American Mathematical Society **363** (cit. on pp. 63–65).

Hoens, T. Ryan and Nitesh V. Chawla (2013), "Imbalanced Datasets: From Sampling to Classifiers",
Imbalanced Learning, John Wiley and Sons, Ltd, chap. 3, p. 43 (cit. on p. 41).

Honerkamp, J. (2002), *Statistical Physics: An Advanced Approach with Applications*,
Graduate Texts in Physics, Springer Berlin Heidelberg (cit. on pp. 13, 65, 66).

Huntley, Brian (2012), *Reconstructing palaeoclimates from biological proxies: Some often overlooked sources of uncertainty*, Quaternary Science Reviews **31**, p. 1 (cit. on p. 20).

Huybers, Peter (2006),
Early Pleistocene Glacial Cycles and the Integrated Summer Insolation Forcing,
Science **313**.5786, p. 508 (cit. on pp. 107, 122).

Iversen, Johs (1944), *Viscum, Hedera and Ilex as Climate Indicators*,
Geologiska Föreningen i Stockholm Förhandlingar **66**.3, p. 463 (cit. on pp. 19, 29).

James, Gareth, Daniela Witten, Trevor Hastie and Robert Tibshirani (2013), *An Introduction to Statistical Learning: with Applications in R*, Springer (cit. on pp. 41, 43).

Jed Wing, Max Kuhn. Contributions from et al. (2019), *caret: Classification and Regression Training*, R package version 6.0-84 (cit. on pp. 30, 41).

Jergensen, G. Eli, Amy McGovern, Ryan Lagerquist and Travis Smith (2020), *Classifying Convective Storms Using Machine Learning*, Weather and Forecasting **35**.2, p. 537 (cit. on p. 45).

Johnson, R.A. and D.W. Wichern (2007), *Applied Multivariate Statistical Analysis*, Applied Multivariate Statistical Analysis, Pearson Prentice Hall (cit. on p. 43).

Kim, Sang-Tae and James R. O'Neil (1997), *Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates*, Geochimica et Cosmochimica Acta **61**.16, p. 3461 (cit. on pp. 14, 15).

Klotz, Stefan, Ulrich Müller, Volker Mosbrugger, Jacques-Louis de Beaulieu and Maurice Reille (2004), *Eemian to early Würmian climate dynamics: history and pattern of changes in Central Europe*, Palaeogeography, Palaeoclimatology, Palaeoecology **211**.1, p. 107 (cit. on p. 30).

Kühl, Norbert, Christoph Gebhardt, Thomas Litt and Andreas Hense (2002), *Probability Density Functions as Botanical-Climatological Transfer Functions for Climate Reconstruction*, Quaternary Research **58**.3, p. 381 (cit. on pp. 3, 30, 38–40, 44).

Kühl, Norbert and T. Litt (2003), *Quantitative time series reconstruction of Eemian temperature at three European sites using pollen data*, Vegetation History and Archaeobotany **12**, p. 205 (cit. on pp. 22, 30, 39).

Kurt, Latif, Tug Tuğ and Osman Ketenoğlu (2006), *Synoptic View of the Steppe Vegetation of Central Anatolia (Turkey)*, Asian Journal of Plant Sciences **5** (cit. on p. 83).

Lachniet, Matthew S. (2009), *Climatic and environmental controls on speleothem oxygen-isotope values*, Quaternary Science Reviews **28**.5, p. 412 (cit. on p. 15).

Langenbrunner, B. and J. D. Neelin (2017), *Multiobjective constraints for climate model parameter choices: Pragmatic Pareto fronts in CESM1*, Journal of Advances in Modeling Earth Systems **9**.5, p. 2008 (cit. on p. 53).

Langgut, Dafna, Rachid Cheddadi and Gonen Sharon (2021), *Climate and environmental reconstruction of the Epipaleolithic Mediterranean Levant (22.0–11.9 ka cal. BP)*, Quaternary Science Reviews **270**, p. 107170 (cit. on p. 138).

Lauritzen, Stein and J. Lundberg (1999), *Calibration of the speleothem delta function: An absolute temperature record for the Holocene in northern Norway*, hol **9**, p. 659 (cit. on pp. 3, 14–16, 38).

Lechleitner, Franziska A. et al. (2018), *The Potential of Speleothems from Western Europe as Recorders of Regional Climate: A Critical Assessment of the SISAL Database*, Quaternary **1**.3 (cit. on p. 15).

LeCun, Y. et al. (1989), *Backpropagation Applied to Handwritten Zip Code Recognition*, Neural Computation **1**.4, p. 541 (cit. on p. 43).

Leisch, Friedrich, Kurt Hornik and Brian D. Ripley. (2017), *mda: Mixture and Flexible Discriminant Analysis*, R package version 0.4-10 (cit. on p. 49).

Li, Camille, David Battisti and Cecilia Bitz (2010), *Can North Atlantic Sea Ice Anomalies Account for Dansgaard-Oeschger Climate Signals?*, Journal of Climate - J CLIMATE **23**, p. 5457 (cit. on p. 9).

Lisiecki, Lorraine E. and Maureen E. Raymo (2005), *A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records*, Paleoceanography **20**.1 (cit. on pp. 2, 10, 81, 82).

Litt, T., J. Richter and F. Schäbitz (2021), *The Journey of Modern Humans from Africa to Europe: Culture-Environmental Interaction and Mobility*, E. Schweizerbartsche Verlagsbuchhandlung (cit. on pp. 31, 102, 111).

Litt, Thomas, Christian Ohlwein, Frank H. Neumann, Andreas Hense and Mordechai Stein (2012), *Holocene climate variability in the Levant from the Dead Sea pollen record*, Quaternary Science Reviews **49**, p. 95 (cit. on pp. 18, 24, 31, 34, 40, 72–76, 78, 79, 140).

Litt, Thomas, Nadine Pickarski, Georg Heumann, Mona Stockhecke and Polychronis C. Tzedakis (2014), *A 600,000 year long continental pollen record from Lake Van, eastern Anatolia (Turkey)*, Quaternary Science Reviews **104**, p. 30, Special Issue: Results from the PALEOVAN Drilling Project: a 600,000 year long continental archive in the Near East (cit. on pp. 2, 4, 103).

Litt, Thomas et al. (2009), *'PALEOVAN', International Continental Scientific Drilling Program (ICDP): site survey results and perspectives*, Quaternary Science Reviews **28**.15, p. 1555 (cit. on pp. 103–105, 109, 110, 140).

Ludwig, Patrick and Assaf Hochman (2022), *Last glacial maximum hydro-climate and cyclone characteristics in the Levant: a regional modelling perspective*,
Environmental Research Letters **17**, p. 014053 (cit. on pp. 5, 138, 141).

Ludwig, Patrick, Yaping Shao, Martin Kehl and Gerd-Christian Weniger (2018),
The Last Glacial Maximum and Heinrich event I on the Iberian Peninsula: A regional climate modelling study for understanding human settlement patterns,
Global and Planetary Change **170**, p. 34 (cit. on p. 138).

Lynch, Scott M. (2007), "Evaluating Markov Chain Monte Carlo Algorithms and Model Fit",
Introduction to Applied Bayesian Statistics and Estimation for Social Scientists,
ed. by Scott M. Lynch, New York, NY: Springer New York, p. 131 (cit. on p. 54).

Mann, Michael, Raymond Bradley and Malcolm Hughes (1998),
Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries,
Nature **392**, p. 779 (cit. on p. 45).

Mann, Michael et al. (2008), *Proxy-based Reconstructions of Hemispheric and Global Surface Temperature Variations over the Past Two Millennia*,
Proceedings of the National Academy of Sciences of the United States of America **105**, p. 13252 (cit. on p. 45).

Manten, A.A. (1967), *Lennart Von Post and the foundation of modern palynology*,
Review of Palaeobotany and Palynology **1**.1, p. 11 (cit. on p. 19).

Mauri, A., Basil Davis, Pamela Collins and Jed Kaplan (2015), *The climate of Europe during the Holocene: A gridded pollen-based reconstruction and its multi-proxy evaluation*,
Quaternary Science Reviews **112** (cit. on pp. 4, 5, 38, 45, 131–134, 141).

Mcculloch, Warren and Walter Pitts (1943),
A Logical Calculus of Ideas Immanent in Nervous Activity,
Bulletin of Mathematical Biophysics **5**, p. 127 (cit. on p. 42).

McDermott, F., T.C. Atkinson, I.J. Fairchild, L.M. Baldini and D.P. Mattey (2011),
A first evaluation of the spatial gradients in δ^{18}O recorded by European Holocene speleothems,
Global and Planetary Change **79**.3, p. 275, Rapid climate change: lessons from the recent geological past (cit. on p. 15).

McDermott, Frank et al. (1999), *Holocene climate variability in Europe: Evidence from δ^{18}O, textural and extension-rate variations in three speleothems*,
Quaternary Science Reviews **18**.8, p. 1021 (cit. on p. 14).

McKay, Nicholas, Julien Emile-Geay and Deborah Khider (2020),
GeoChronR – an R package to model, analyze and visualize age-uncertain paleoscientific data,
(cit. on p. 29).

Meusel, H., E. Weinert and E. Jäger (1974),
Vergleichende Chorologie der zentraleuropäischen Flora: T. 1 Text. T. 2 Karten, Literatur, Register,
Bd. II, Vergleichende Chorologie der zentraleuropäischen Flora, G. Fischer (cit. on p. 39).

Miebach, Andrea, Chunzhu Chen, Markus J. Schwab, Mordechai Stein and Thomas Litt (2017),
Vegetation and climate during the Last Glacial high stand (ca. 28–22 ka BP) of the Sea of Galilee, northern Israel, Quaternary Science Reviews **156**, p. 47 (cit. on p. 85).

Miebach, Andrea, Phoebe Niestrath, Patricia Roeser and Thomas Litt (2016),
Impacts of climate and humans on the vegetation in northwestern Turkey: Palynological insights from Lake Iznik since the Last Glacial, Climate of the Past **12**, p. 575
(cit. on pp. 4, 34, 111–116, 140).

Miebach, Andrea, Sophie Stolzenberger, Lisa Wacker, Andreas Hense and Thomas Litt (2019),
A new Dead Sea pollen record reveals the last glacial paleoenvironment of the southern Levant,
Quaternary Science Reviews **214**, p. 98 (cit. on pp. 24, 71, 72, 80, 81, 84, 85, 87–90, 140).

Miebach, Andrea et al. (2022), *Changing fire regimes during the first olive cultivation in the Mediterranean Basin: New high-resolution evidence from the Sea of Galilee, Israel*,
Global and Planetary Change **210**, p. 103774 (cit. on pp. 90, 96, 140).

Migowski, Claudia, Mordechai Stein, Sushma Prasad, Jörg F.W. Negendank and Amotz Agnon (2006),
Holocene climate variability and cultural evolution in the Near East from the Dead Sea sedimentary record, Quaternary Research **66**.3, p. 421, Holocene Climate and Cultural Evolution in Late Prehistoric-Early Historic West Asia (cit. on pp. 75, 90).

Moore, G.W. (1952), *Speleothem—a new cave term*, National Speleological Society of America **10**
(cit. on p. 13).

Neugebauer, Ina et al. (2016), *Hydroclimatic variability in the Levant during the early last glacial (117–75 ka) derived from micro-facies analyses of deep Dead Sea sediments*,
Climate of the Past **12**, p. 75 (cit. on p. 72).

Neumann, Frank, Christian Schölzel, Thomas Litt, Andreas Hense and Mordechai Stein (2007),
Holocene vegetation and climate history of the northern Golan heights (Near East),
Vegetation History and Archaeobotany **16**.4, p. 329 (cit. on pp. 22, 24, 30, 34, 102, 133, 140).

Nieto-Lugilde, Diego, Francisca Alba-Sanchez and Grettel Vargas-Azofeifa (2019),
EPDr: Interact with the European Pollen Database (EPD), R package version 1.0 (cit. on p. 21).

Niggemann, Stefan, Augusto Mangini, Manfred Mudelsee, Detlev Richter and Georg Wurth (2003), *Sub-Milankovitch climatic cycles in Holocene stalagmites from Sauerland, Germany*, Earth and Planetary Science Letters - EARTH PLANET SCI LETT **216** (cit. on p. 16).

Orland, Ian et al. (2009), *Climate deterioration in the Eastern Mediterranean as revealed by ion microprobe analysis of a speleothem that grew from 2.2 to 0.9ka in Soreq Cave, Israel*, Quaternary Research **71**, p. 27 (cit. on pp. 18, 77, 80, 94, 96).

Overpeck, J.T., T. Webb and I.C. Prentice (1985), *Quantitative interpretation of fossil pollen spectra: Dissimilarity coefficients and the method of modern analogs*, Quaternary Research **23**.1, p. 87 (cit. on p. 30).

PAGES, Past Interglacials Working Group of (2016), *Interglacials of the last 800,000 years*, Reviews of Geophysics **54**.1, p. 162 (cit. on p. 1).

Panagiotopoulos, Konstantinos, Anne Aufgebauer, Frank Schäbitz and Bernd Wagner (2013), *Vegetation and climate history of the Lake Prespa region since the Lateglacial*, Quaternary International **293**, p. 157, Advancing Pleistocene and Holocene climate change research in the Carpathian-Balkan region (cit. on pp. 4, 34, 57–59, 66–68, 116, 140).

Panagiotopoulos, Konstantinos et al. (2020), *Insights into the evolution of the young Lake Ohrid ecosystem and vegetation succession from a southern European refugium during the Early Pleistocene*, Quaternary Science Reviews **227**, p. 106044 (cit. on pp. 4, 57, 116–118, 120, 122, 140).

Parnell, Andrew et al. (2014), *Bayesian inference for palaeoclimate with time uncertainty and stochastic volatility*, Journal of the Royal Statistical Society: Series C (Applied Statistics) **64** (cit. on pp. 22, 29).

Peel, David (2005), *W. Dansgaard 2004. Frozen annals: Greenland ice sheet research. Copenhagen, University of Copenhagen. Department of Geophysics of the Niels Bohr Institute. 122 pp. ISBN 87-990078-0-0.*, Journal of Glaciology **51**, p. 170 (cit. on p. 9).

Peyron, Odile et al. (1998), *Climatic Reconstruction in Europe for 18,000 YR B.P. from Pollen Data*, Quaternary Research **49**, p. 183 (cit. on p. 3).

Pickarski, Nadine, Ola Kwiecien, Morteza Djamali and Thomas Litt (2015a), *Vegetation and environmental changes during the last interglacial in eastern Anatolia (Turkey): A new high-resolution pollen record from Lake Van*, Palaeogeography, Palaeoclimatology, Palaeoecology **435** (cit. on pp. 3, 103–106, 109, 110, 140).

Pickarski, Nadine, Ola Kwiecien, D. Langgut and T. Litt (2015b),
Abrupt climate variability of eastern Anatolia vegetation during the last glacial,
Climate of the Past Discussions **11**, p. 3341 (cit. on pp. 103–105, 107–110, 140).

Prentice, Colin, J. Guiot, B. Huntley, D. Jolly and R. Cheddadi (1996), *Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka*, Climate Dynamics **12**.3, p. 185 (cit. on p. 31).

Prentice, I. Colin et al. (1992), *Special Paper: A Global Biome Model Based on Plant Physiology and Dominance, Soil Properties and Climate*, Journal of Biogeography **19**.2, p. 117 (cit. on p. 31).

Prest, V. K. (1990), *Laurentide Ice-Flow Patterns: A Historical Review, and Implications of the Dispersal of Belcher Islands Erratics*, Geographie Physique Et Quaternaire **44**, p. 113 (cit. on pp. 7, 9).

Psomiadis, D., E. Dotsika, K. Albanakis, B. Ghaleb and C. Hillaire-Marcel (2018),
Speleothem record of climatic changes in the northern Aegean region (Greece) from the Bronze Age to the collapse of the Roman Empire,
Palaeogeography, Palaeoclimatology, Palaeoecology **489**, p. 272 (cit. on p. 24).

R Core Team (2018), *R: A Language and Environment for Statistical Computing*,
R Foundation for Statistical Computing, Vienna, Austria (cit. on pp. 24, 49).

Rasmussen, S. et al. (2014), *A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: Refining and extending the INTIMATE event stratigraphy*, Quaternary Science Reviews **106** (cit. on pp. 4, 12, 87).

Rausch, Joseph and Ken Kelley (2009),
A comparison of linear and mixture models for discriminant analysis under nonnormality,
Behavior research methods **41**, p. 85 (cit. on p. 44).

Ravelo, Ana Christina and Claude Hillaire-Marcel (2007),
"Chapter Eighteen The Use of Oxygen and Carbon Isotopes of Foraminifera in Paleoceanography",
Proxies in Late Cenozoic Paleoceanography, ed. by Claude Hillaire–Marcel and Anne De Vernal, vol. 1, Developments in Marine Geology, Elsevier, p. 735 (cit. on pp. 8, 10).

Richards, David and Jeffrey Dorale (2003),
Uranium-series Chronology and Environmental Applications of Speleothems,
Reviews in Mineralogy and Geochemistry - REV MINERAL GEOCHEM **52**, p. 407 (cit. on p. 15).

Roesch, Carla and Kira Rehfeld (2019),
Automatising age-depth modelling for hundreds of speleothems, (cit. on p. 25).

Roeser, Patricia, Sven Franz and Thomas Litt (2016), *Aragonite and calcite preservation in sediments from Lake Iznik related to bottom lake oxygenation and water column depth*, Sedimentology **63**, p. 2253 (cit. on p. 113).

Rubner, Yossi, Carlo Tomasi and Leonidas J. Guibas (2000), *The Earth Mover's Distance as a Metric for Image Retrieval.*, International Journal of Computer Vision **40**.2, p. 99 (cit. on p. 66).

Rumelhart, D., Geoffrey E. Hinton and R. J. Williams (1986), *Learning representations by back-propagating errors*, Nature **323**, p. 533 (cit. on p. 42).

Sánchez-Fernández, David et al. (2018), *The deep subterranean environment as a potential model system in ecological, biogeographical and evolutionary research*, Subterranean Biology **25**, p. 1 (cit. on pp. 14, 38).

Schiebel, Vera (2013), *Vegetation and climate history of the southern Levant during the last 30,000 years based on palynological investigation*, PhD thesis: Rheinische Friedrich-Wilhelms-Universität Bonn (cit. on pp. 34, 72, 75, 76, 80, 85, 90, 94, 96–98, 100–102, 140).

Schiebel, Vera and Thomas Litt (2018), *Holocene vegetation history of the southern Levant based on a pollen record from Lake Kinneret (Sea of Galilee), Israel*, Vegetation History and Archaeobotany **27**.4, p. 577 (cit. on pp. 4, 20, 21, 24, 76, 89, 90, 92, 94–96, 140).

Schmidhuber, Jürgen (2015), *Deep learning in neural networks: An overview*, Neural Networks **61**, p. 85 (cit. on pp. 42, 43).

Scholz, Denis and Dirk L. Hoffmann (2011), *StalAge – An algorithm designed for construction of speleothem age models*, Quaternary Geochronology **6**.3, p. 369 (cit. on p. 24).

Schölzel, C. (2006), *Palaeoenvironmental Transfer Functions in a Bayesian Framework with Application to Holocene Climate Variability in the Near East*, Bonner Meteorologische Abhandlungen 62, p. 104 (cit. on pp. 30, 72).

Schwab, Markus, Frank Neumann, Thomas Litt, Jörg Negendank and Mordechai Stein (2004), *Holocene palaeoecology of the Golan Heights (Near East): Investigation of lacustrine sediments from Birkat Ram crater lake*, Quaternary Science Reviews **23**, p. 1723 (cit. on pp. 96, 97).

Seppä, H., D. Hammarlund and K. Antonsson (2005), *Low-frequency and high-frequency changes in temperature and effective humidity during the*

Holocene in south-central Sweden: implications for atmospheric and oceanic forcings of climate, Climate Dynamics **25**.2-3, p. 285 (cit. on p. 24).

Shakun, Jeremy D., David W. Lea, Lorraine E. Lisiecki and Maureen E. Raymo (2015), *An 800-kyr record of global surface ocean δ^{18}O and implications for ice volume-temperature coupling*, Earth and Planetary Science Letters **426**, p. 58 (cit. on p. 8).

Simard, Richard and Pierre L'Ecuyer (2011), *Computing the Two-Sided Kolmogorov-Smirnov Distribution*, Journal of Statistical Software **39**.11, p. 1 (cit. on p. 47).

Simonis, Daniel, Andreas Hense and Thomas Litt (2012), *Reconstruction of late Glacial and Early Holocene near surface temperature anomalies in Europe and their statistical interpretation*, Quaternary International **274**, p. 233, Temporal and spatial corridors of Homo sapiens sapiens population dynamics during the Late Pleistocene and Early Holocene (cit. on pp. 5, 22, 30, 45, 133, 134).

Spackman, Kent (1991), *Maximum likelihood training of connectionist models: comparison with least squares back-propagation and logistic regression*, Proceedings. Symposium on Computer Applications in Medical Care, p. 285 (cit. on p. 42).

Stockhecke, Mona et al. (2014), *Chronostratigraphy of the 600,000 year old continental record of Lake Van (Turkey)*, Quaternary Science Reviews **104**, p. 8, Special Issue: Results from the PALEOVAN Drilling Project: a 600,000 year long continental archive in the Near East (cit. on pp. 104, 105).

Stockhecke, Mona et al. (2016), *Millennial to orbital-scale variations of drought intensity in the Eastern Mediterranean*, Quaternary Science Reviews **133**, p. 77 (cit. on p. 85).

Stolzenberger, S. (2017), *On the probabilistic evaluation of decadal and paleoclimate model predictions*, Bonner Meteorologische Abhandlungen 81, p. 122 (cit. on pp. 3, 31, 40, 42, 45, 47, 134, 141).

Stolzenberger, Sophie (2011), *Untersuchungen zu botanischen Paläoklimatransferfunktionen*, diplomathesis: Meteorologisches Institut der Universitäts Bonn (cit. on pp. 30, 40).

Sueyoshi, T. et al. (2013), *Set-up of the PMIP3 paleoclimate experiments conducted using an Earth system model, MIROC-ESM*, Geoscientific Model Development **6**.3, p. 819 (cit. on pp. 129, 134).

Thoma, B.M. (2017), *Palaeoclimate Reconstruction in the Levant and on the Balkans*, Bonner Meteorologische Abhandlungen 78, p. 266 (cit. on pp. 30, 40–42, 102).

Tingley, Martin and Peter Huybers (2010),
A Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time. Part I: Development and Applications to Paleoclimate Reconstruction Problems,
Journal of Climate - J CLIMATE **23**, p. 2759 (cit. on p. 47).

Torfstein, Adi, Steven L. Goldstein, Mordechai Stein and Yehouda Enzel (2013),
Impacts of abrupt climate changes in the Levant from Last Glacial Dead Sea levels,
Quaternary Science Reviews **69**, p. 1 (cit. on pp. 72, 85).

Torfstein, Adi et al. (2015),
Dead Sea drawdown and monsoonal impacts in the Levant during the last interglacial,
Earth and Planetary Science Letters **412**, p. 235 (cit. on pp. 24, 83).

Torrence, Christopher and Gilbert P. Compo (1998), *A Practical Guide to Wavelet Analysis*,
Bulletin of the American Meteorological Society **79**.1, p. 61 (cit. on p. 13).

Treble, Pauline et al. (2007), *High resolution Secondary Ionisation Mass Spectrometry (SIMS) δ^{18}O analyses of Hulu Cave speleothem at the time of Heinrich Event 1*, Chemical Geology **238**, p. 197 (cit. on p. 15).

Tremaine, Darrel M., Philip N. Froelich and Yang Wang (2011),
Speleothem calcite farmed in situ: Modern calibration of δ^{18}O and δ^{13}C paleoclimate proxies in a continuously-monitored natural cave system, Geochimica et Cosmochimica Acta **75**.17, p. 4929 (cit. on pp. 3, 14).

Venables, W. N. and B. D. Ripley (2002), *Modern Applied Statistics with S*, Fourth,
ISBN 0-387-95457-0, New York: Springer (cit. on pp. 43, 49).

Wagner, Bernd et al. (2019),
Mediterranean winter rainfall in phase with African monsoons during the past 1.36 million years,
Nature **573**, p. 256 (cit. on pp. 58, 80, 116, 117).

Wagner, Jennifer D. et al. (2010),
Moisture variability in the southwestern United States linked to abrupt glacial climate change,
Nature Geoscience **3**, p. 110 (cit. on p. 38).

Waltgenbach, Sarah et al. (2020),
Climate and structure of the 8.2 ka event reconstructed from three speleothems from Germany,
Global and Planetary Change **193**, p. 103266 (cit. on pp. 16, 24).

Wan, Nai-Jung et al. (2011), *Spatial variations of monsoonal rain in eastern China: Instrumental, historic and speleothem records*, Journal of Asian Earth Sciences **40**.6, p. 1139, Quaternary

Paleoclimate of the Western Pacific and East Asia: State of the Art and New Discovery (cit. on p. 16).

Wang, Pinxian, Jun Tian and Lucas J. Lourens (2010), *Obscuring of long eccentricity cyclicity in Pleistocene oceanic carbon isotope records*, Earth and Planetary Science Letters **290**.3, p. 319 (cit. on p. 10).

Weaver, John Ernest (1965), "Native vegetation of Nebraska", Agronomy & Horticulture – Faculty Publications (cit. on p. 99).

Webb, Thompson and Reid A. Bryson (1972), *Late- and Postglacial Climatic Change in the Northern Midwest, USA: Quantitative Estimates Derived from Fossil Pollen Spectra by Multivariate Statistical Analysis*, Quaternary Research **2**.1, p. 70 (cit. on p. 30).

Weitzel, N., A. Hense and C. Ohlwein (2019), *Combining a pollen and macrofossil synthesis with climate simulations for spatial reconstructions of European climate using Bayesian filtering*, Climate of the Past **15**.4, p. 1275 (cit. on pp. 4, 5, 22, 30, 40–42, 45–48, 123, 134, 139–141).

Weitzel, Nils (2020), *Climate field reconstructions from pollen and macrofossil syntheses using Bayesian hierarchical models*, Bonner Meteorologische Abhandlungen 89, p. 153 (cit. on pp. 47, 48).

Zhang, Haiwei et al. (2019), *The Asian Summer Monsoon: Teleconnections and Forcing Mechanisms—A Review from Chinese Speleothem δ^{18}O Records*, Quaternary **2**.3 (cit. on p. 15).

Appendix

APPENDIX A

Age-depth models

In the following, Bayesian age-depth models using the Bacon algorithm are presented. The order corresponds to that of the local climate reconstructions in this thesis.

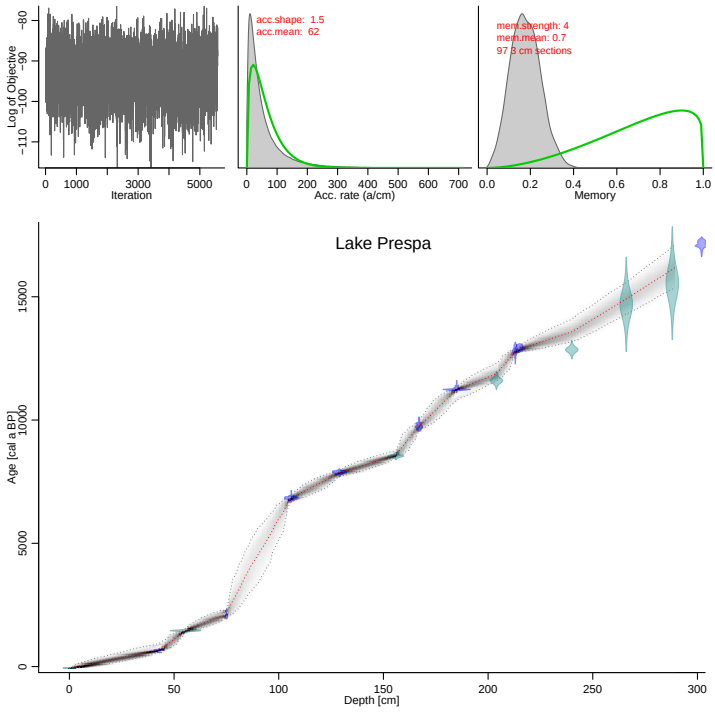


Figure A.1: The bottom panel shows an age-depth model calculated with Bacon from a sediment core from Lake Prespa. The top panels depict, from left to right, a trace plot of the posterior MCMC samples, prior (green lines) and posterior (gray area) distributions of accumulation rate and memory.

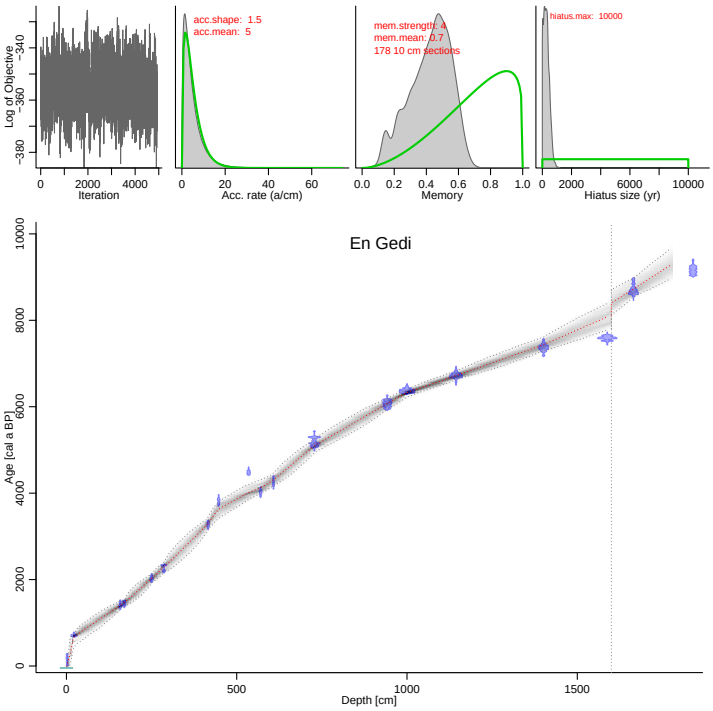


Figure A.2: As Fig. A.1, but for En Gedi. In addition, prior and posterior distributions of hiatus size are presented.

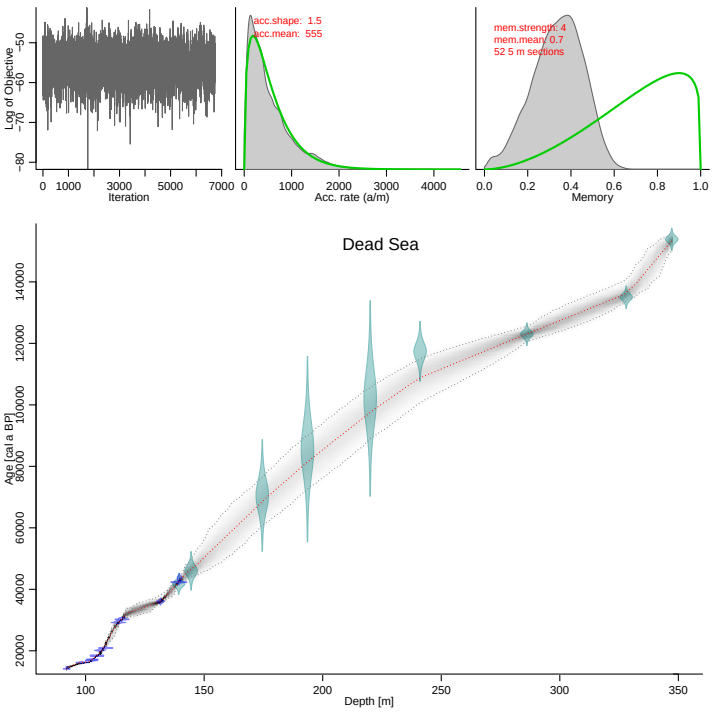


Figure A.3: As Fig. A.1, but for Dead Sea.

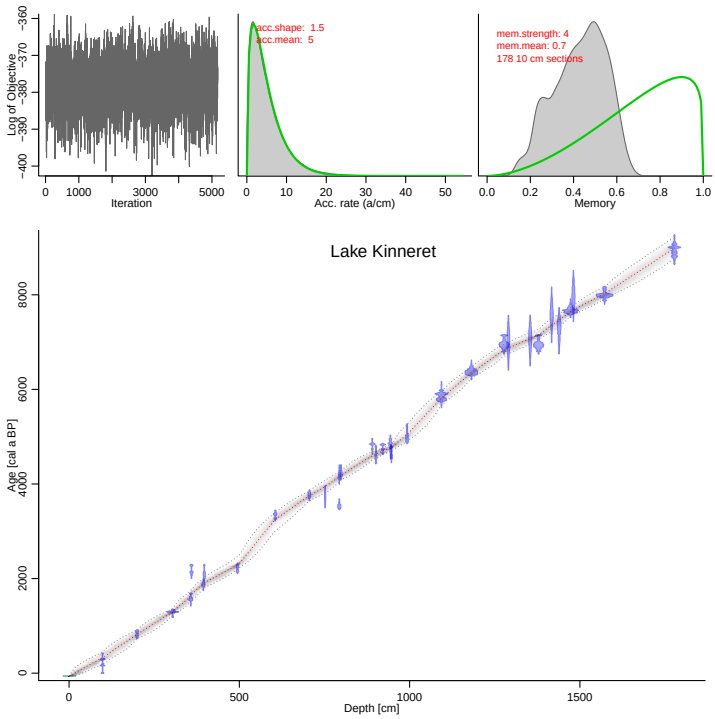


Figure A.4: As Fig. A.1, but for Lake Kinneret.

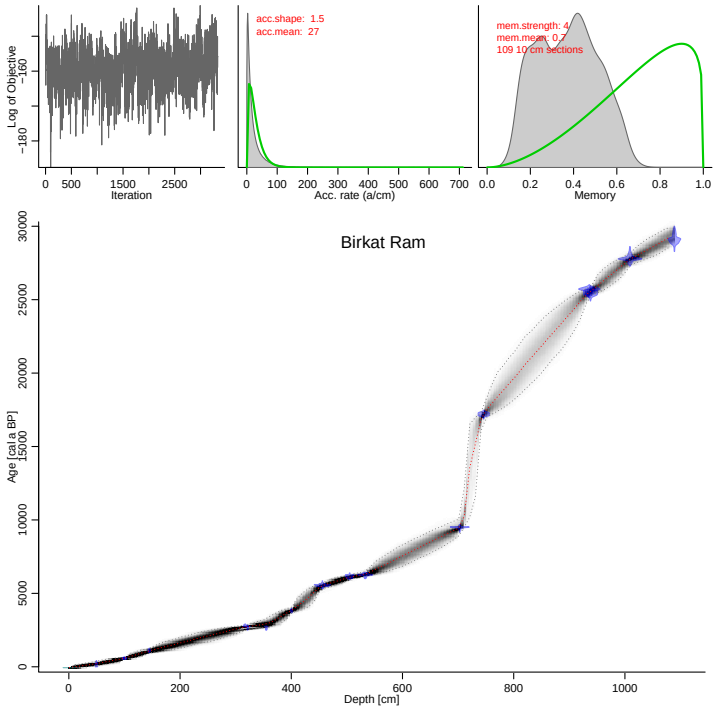


Figure A.5: As Fig. A.1, but for Birkat Ram.

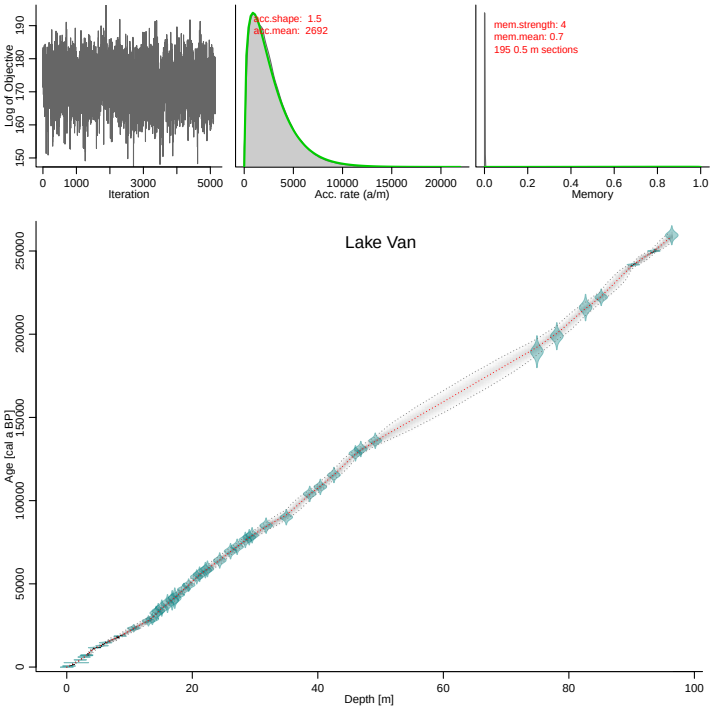


Figure A.6: As Fig. A.1, but for Lake Van.

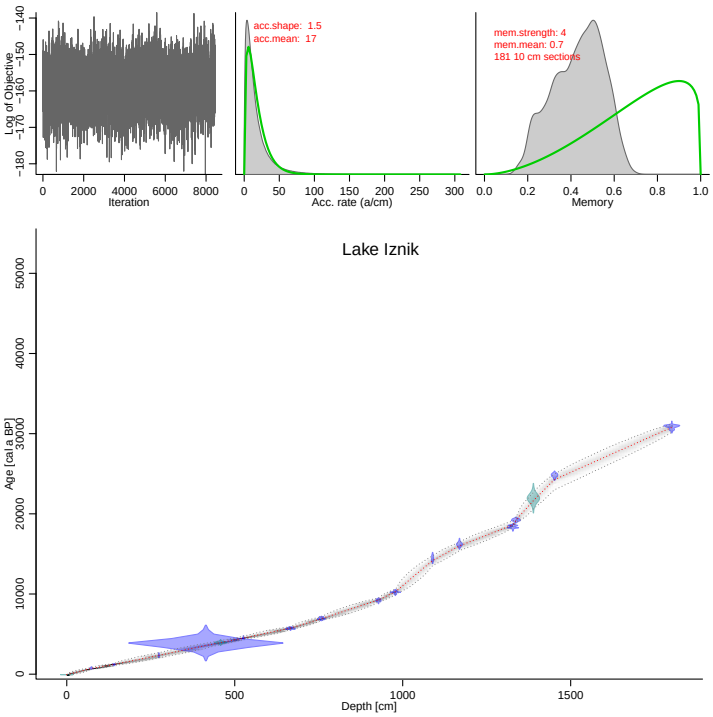


Figure A.7: As Fig. A.1, but for Lake Iznik.

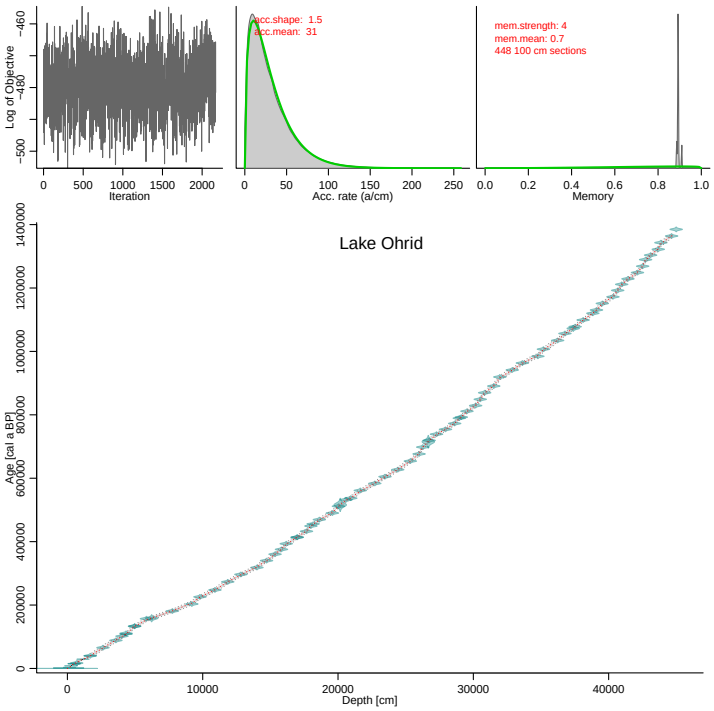


Figure A.8: As Fig. A.1, but for Lake Ohrid.

APPENDIX B

Spatial climate reconstructions

B.1 Middle Holocene results incorporating the Simonis dataset

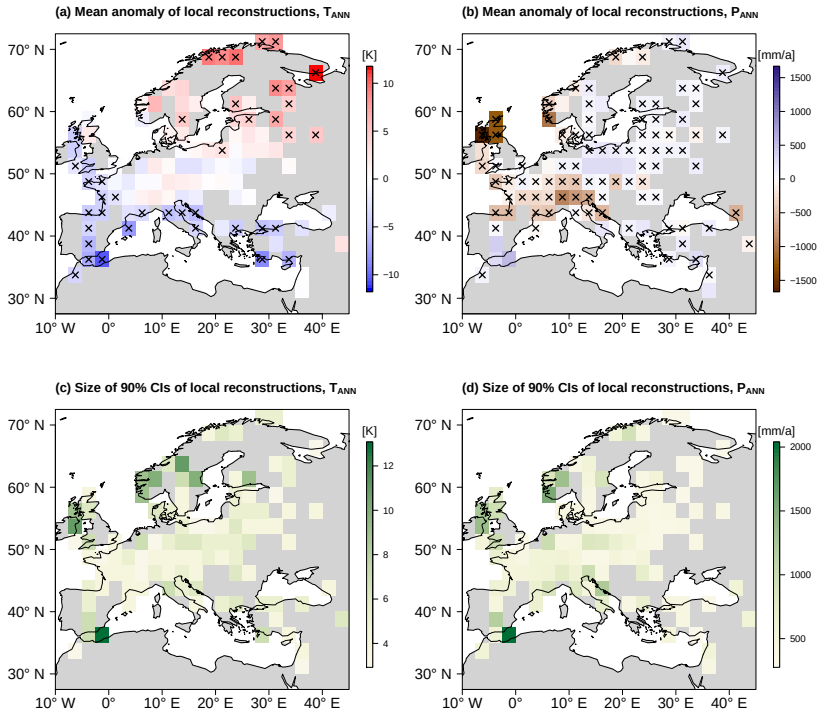


Figure B.1: Panels (a) and (b) show the mean T_{ANN} and P_{ANN} anomaly with respect to the CMIP ensemble for the grid points where the new local information from this work and from the Simonis dataset are available for the MH. The grid points with crosses indicate significant anomalies. In (c) and (d), the corresponding 90 % credible intervals are depicted.

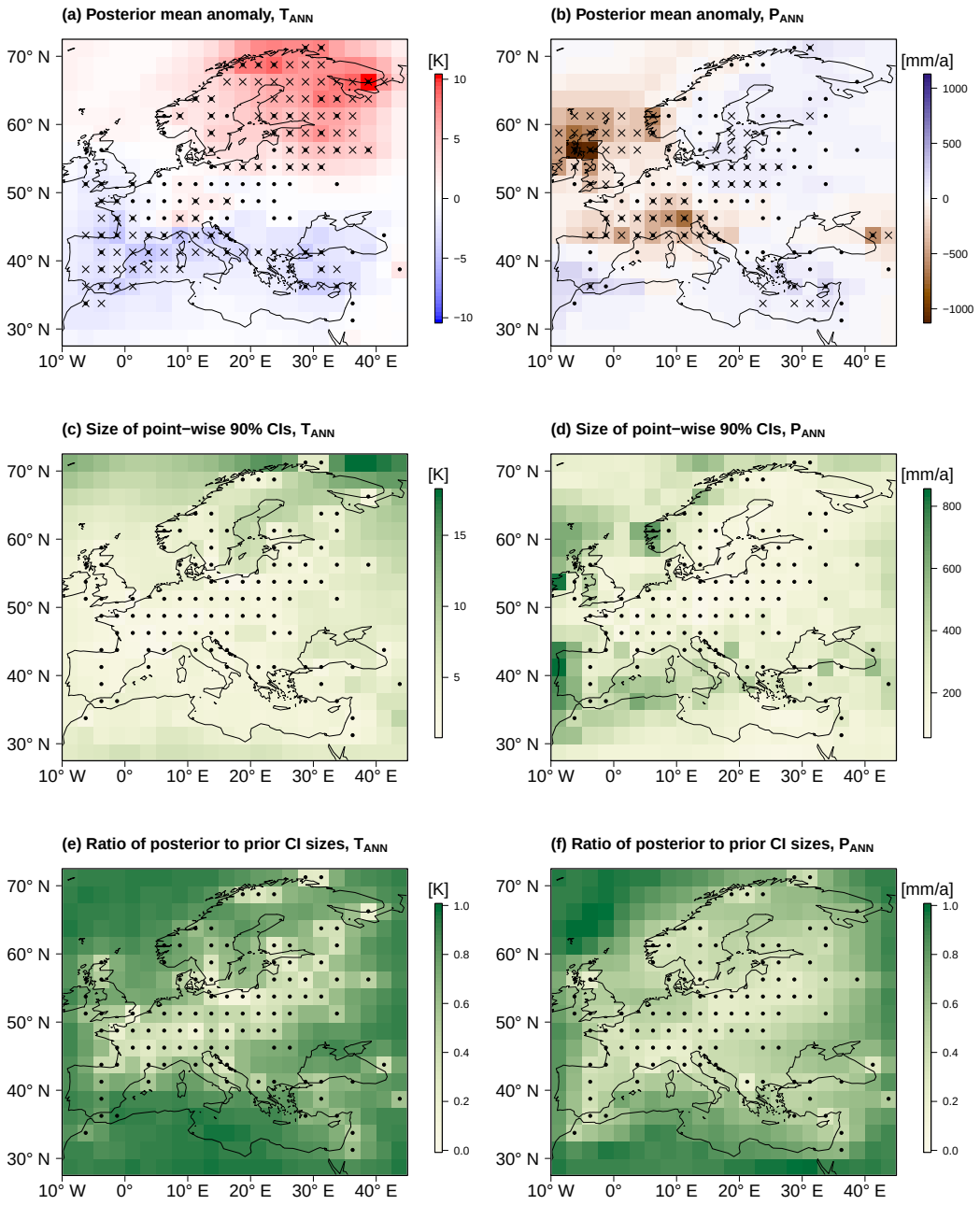


Figure B.2: Panels (a) and (b) show the mean posterior anomaly of T_{ANN} and P_{ANN} with respect to the CMIP ensemble. The grid points with black dots contain local reconstruction information shown in Fig. B.1, and those with crosses indicate a significant anomaly. In (c) and (d), the corresponding 90 % credible intervals are depicted. The ratios between the 90 % credible interval of the posterior and prior simulations are given in panels (e) and (f).

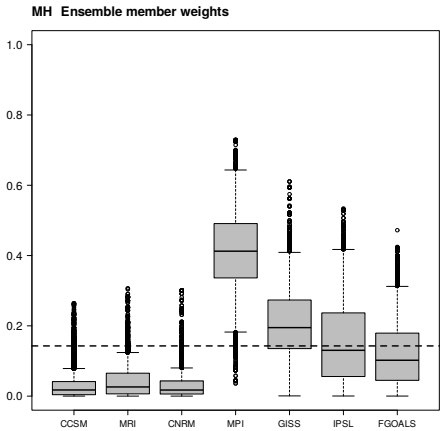


Figure B.3: The boxplots represent the selected MCMC samples of the posterior weights of each ensemble member of the MH. The corresponding prior weight is depicted by the horizontal dashed line.

B.2 Spatial fields of the lower and upper boundary of the 90 % CI

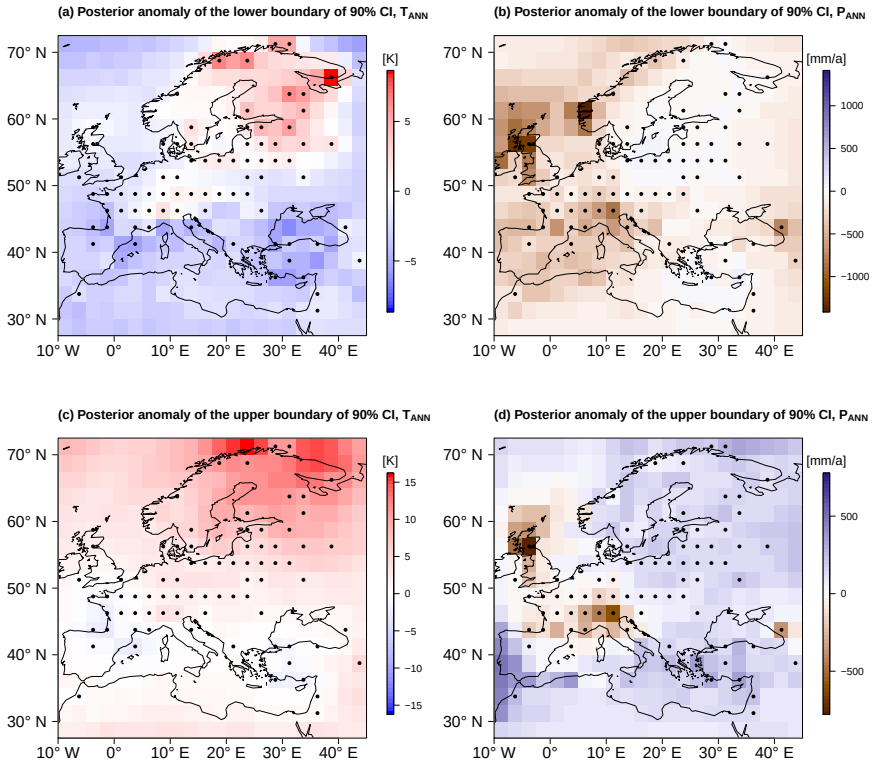


Figure B.4: In (a) and (b), the lower bound of the 90 % credible interval of annual temperature and annual precipitation of the Middle Holocene is shown. Panels (c) and (d) present the same, but for the upper boundary.

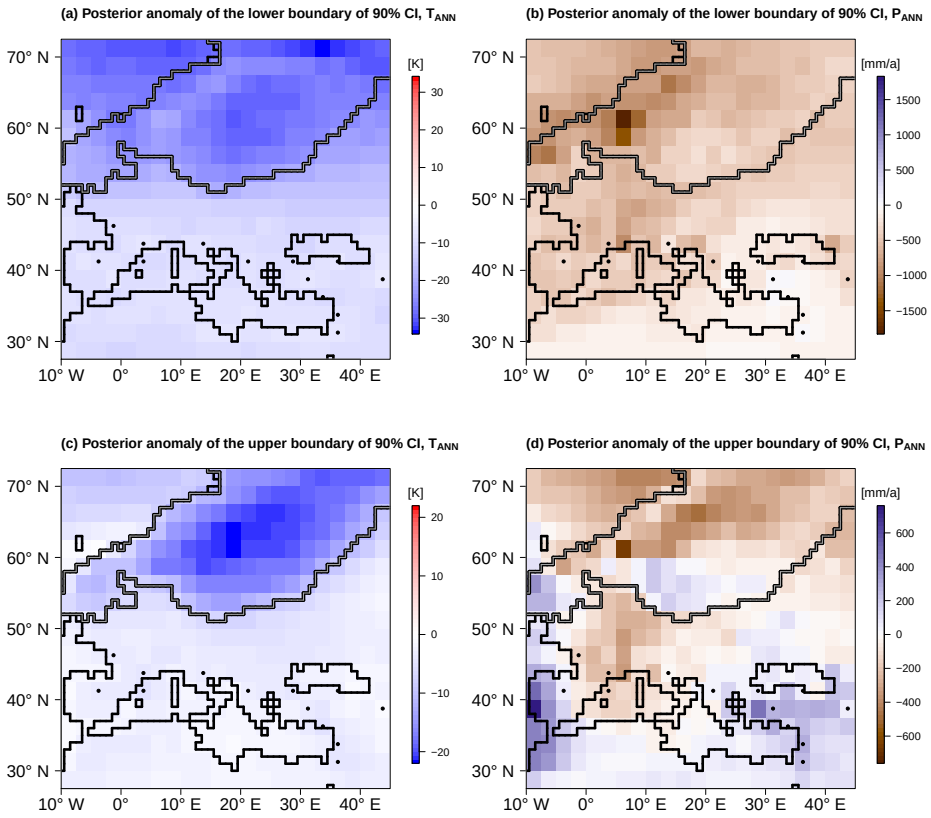


Figure B.5: As Fig. B.4, but for the Last Glacial Maximum.

Acronyms

$\mathbf{P_{ANN}}$ Mean annual precipitation.

$\mathbf{T_{ANN}}$ Mean annual temperature.

AC Autocorrelation.

AdaBoost Adaptive Boosting.

AMOC Atlantic Meridional Overturning Circulation.

amsl Above mean sea level.

ANN Artificial Neural Network.

ANOVA Analysis of variance.

AP Arboreal pollen.

ASM Asian summer monsoon.

BA Bølling-Allerød.

Bacon **B**ayesian **ac**cumulati**on**.

BBM Bayesian Biome Model.

Bchron A Bayesian based age-depth model.

BHM Bayesian Hierarchical Model.

BIC Bayesian information criterion.

BITM Bayesian Indicator Taxa Model.

bmsl Below mean sea level.

CDF Cumulative distribution function.

CFR Climate field reconstructions.

CI Credible interval.

CMIP Coupled Model Intercomparison Project.

COI Cone of influence.

CRPS Continuous Ranked Probability Score.

CRPSS Continuous Ranked Probability Skill Score.

CRU Climate Research Unit.

DAG Directed acyclic graph.

DO Dansgaard-Oeschger.

EMD Earth Mover's Distance.

ENSO El Niño Southern Oscillation.

EPD European Pollen Database.

ESM Earth System Model.

FNN Feedforward neural network.

GBM Gradient Boosting Machine.

Glasso Graphical lasso.

GLM Generalized Linear Model.

GM Gaussian Model.

H Heinrich.

HIST Historical.

IAEA International Atomic Energy Agency.

IAT Integrated autocorrelation time.

IRD Ice rafted debris.

KM Kernel Model.

LDA Linear Discriminant Analysis.

LGM Last Glacial Maximum.

LIG Last Interglacial.

LIS Laurentide Ice Sheet.

LPAZ Local pollen assemblage zone.

LR04 Globally distributed benthic δ^{18}O stack.

LUT Lookup table.

MAT Modern Analogue Technique.

MCMC Markov chain Monte Carlo.

MCR Mutual Climate Range.

MDA Mixture Discriminant Analysis.

Medstack Mediterranean stack from planktonic δ^{18}O.

MG Metropolis-within-Gibbs.

MH Middle Holocene.

MIS Marine Isotope Stage.

ML Machine learning.

NAO North Atlantic Oscillation.

NAP Non-arboreal pollen.

NGRIP North Greenland Ice Core Project.

PCA Principal component analysis.

PDF Probability density function.

PMIP Paleoclimate Modelling Intercomparison Project.

QDA Quadratic Discriminant Analysis.

RF Random Forest.

RM Regression Model.

rwMH Random walk Metropolis-Hastingsn.

SISAL **S**peleothem **I**sotopes **S**ynthesis and **A**na**l**ysis.

SMOTE Synthetic Minority Oversampling Technique.

SST Sea Surface Temperature.

StalAge Speleothem age modelling.

SVM Support Vector Machine.

VPDB Vienna Pee Dee Belemnite.

VSMOW Vienna Standard Mean Ocean Water.

WPS Wavelet power spectrum.

WRF Regional weather research and forecast model.

YD Younger Dryas.

Acknowledgements

At this point I would like to thank some people without whom this work would not have been possible. First and foremost, I would like to mention Andreas Hense, who made it possible for me to work on the topic and always had time for questions and ideas. In addition, I would like to thank Thomas Litt, Andrea Miebach, and Konstantinos Panagiotopoulos for their support and expertise. Moreover, my first steps into paleoclimatology were made possible by Sophie Stolzenberger, Benno Thoma, and Nils Weitzel. I would also like to thank Sebastian Buschow, who always gave me advice and proofread my work. Another thank you goes to Thalea Rehan who took the time to spell check my work. The German Research Foundation (DFG) funded my work within the CRC 806 "Our Way to Europe".

BONNER METEOROLOGISCHE ABHANDLUNGEN

Herausgegeben vom Institut für Geowissenschaften der Universität Bonn, Abteilung Meteorologie, durch Prof. Dr. H. FLOHN (Hefte 1-25), Prof. Dr. M. HANTEL (Hefte 26-35), Prof. Dr. H.-D. SCHILLING (Hefte 36-39), Prof. Dr. H. KRAUS (Hefte 40-49), ab Heft 50 durch Prof. Dr. A. HENSE.

Heft 1-79: siehe `https://www.ifgeo.uni-bonn.de/abteilungen/meteorologie/bibliothek/bonner-meteorologische-abhandlungen-bma`

80-94: open access, verfügbar unter `https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/1627`

Heft 80: ***Tanja Zerenner***: Atmospheric downscaling using multi-objective genetic programming, 2016, [erschienen] 2017, X, 191 S.

Heft 81: ***Sophie Stolzenberger***: On the probabilistic evaluation of decadal and paleoclimate model predictions, 2017, IV, 122 S.

Heft 82: ***Insa Thiele-Eich***: Flooding in Dhaka, Bangladesh, and the challenge of climate change, 2017, V, 158 S.

Heft 83: ***Liselotte Bach***: Towards a probabilistic regional reanalysis for Europe, 2017 [erschienen] 2018, VI, 114 S.

Heft 84: ***Yen-Sen Lu***: Propagation of land surface model uncertainties in terrestrial system states, 2017, [erschienen] 2018, X, 120 S.

Heft 85: ***Rüdiger Hewer***: Stochastic physical models for wind fields and precipitation extremes, 2018, 99 S.

Heft 86: ***Sebastian Knist***: Land-atmosphere interactions in multiscale regional climate change simulations over Europe, 2018, VIII, 147 S.

Heft 87: ***Jessica Keune***: Integrated terrestrial simulations at the continental scale: Impact of groundwater dynamics and human water use on groundwater-to-atmosphere feedbacks during the European heatwave in 2003, 2019, IX, 172 S.

Heft 88: ***Christoph Beekmans***: 3-D Cloud Morphology and Evolution Derived from Hemispheric Stereo Cameras, 2019, [erschienen] 2020, VIII, 118 S.

Heft 89: ***Nils Weitzel***: Climate field reconstructions from pollen and macrofossil syntheses using Bayesian hierarchical models, 2019, [erschienen] 2020, XII, 153 S.

Heft 90: ***Alexander Kelbch***: Investigations to quantify individual exposure to solar ultraviolet erythemal radiation including cloud meteorological impact, 2020, III, 107 S.

Heft 91: ***Mari L. Schmidt***: Improvement of hail detection and nowcasting by synergistic combination of information from polarimetric radar, model predictions, and in-situ observations, 2020, VI, 136 S.

Heft 92: ***Sebastian Brune***: Der Wavelet-basierte Organisationsindex als Maß der konvektiven Organisation über Deutschland und dem tropischen Atlantik, 2021, IV, 121 S.

Heft 93: ***Sebastian Buschow***: Spatial Verification with Wavelets, 2022, V, 195 S.

Heft 94: ***Michael Langguth***: Representation of deep convection at gray-zone resolutions - Implementing and testing the HYbrid MAss flux Convection Scheme (HYMACS) in the ICON model, 2022, VI, 173 S.

Heft 95: ***Timon Netzel***: Quantitative paleoclimate reconstructions in the European region based on multiple proxies, 2023, VI, 179 S.

Institut für Geowissenschaften
Abteilung Meteorologie
Mathematisch Naturwissenschaftliche Fakultät
Universität Bonn

institut für geowissenschaften
METEOROLOGIE