BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 78 (2017) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Benno Michael Thoma

PALAEOCLIMATE RECONSTRUCTION IN THE LEVANT AND ON THE BALKANS

BONNER METEOROLOGISCHE ABHANDLUNGEN
Heft 78 (2017) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Benno Michael Thoma

PALAEOCLIMATE RECONSTRUCTION IN THE LEVANT AND ON THE BALKANS

Palaeoclimate Reconstruction in the Levant and on the Balkans

DISSERTATION
ZUR
ERLANGUNG DES DOKTORGRADES (DR. RER. NAT.)
DER
MATHEMATISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT
DER
RHEINISCHEN FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

vorgelegt von
Diplom-Meteorologe **Benno Michael Thoma**
aus
Illertissen

Bonn, Oktober 2016

Diese Arbeit ist die ungekürzte Fassung einer der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn im Jahr 2016 vorgelegten Dissertation von Benno Michael Thoma aus Illertissen.

This paper is the unabridged version of a dissertation thesis submitted by Benno Michael Thoma born in Illertissen to the Faculty of Mathematical and Natural Sciences of the Rheinische Friedrich-Wilhelms-Universität Bonn in 2016.

Anschrift des Verfassers: Address of the author:

Benno Michael Thoma
Meteorologisches Institut der
Universität Bonn
Auf dem Hügel 20
D-53121 Bonn

1. Gutachter: **Prof. Dr. Andreas Hense**, Universität Bonn
2. Gutachter: **Prof. Dr. Thomas Litt**, Universität Bonn

Tag der Promotion: 30. Mai 2017
Erscheinungsjahr: 2017

Für Christina und Aaron.

Abstract

For an understanding of the climate system's variability, knowledge of the past climate is essential. Continuous observations are not available for longer than the last 100 years, a period that is insufficient to understand the variability and sensitivity of the climate system. Since information of both is necessary to build good climate models, the reconstruction of the past climate is fundamental. With palaeoclimate reconstructions it is possible to get information about the past climate and the climate changes in the period and region of interest. In this work these are the Levant region and the Balkans. The Levant region is situated around the Jordan Valley in Israel.

In this work the presence of pollen and macrofossils is used as a proxy. In detail the method bases on the assumption that the presence of a plant or more general of biometypes in a certain area is addicted to the climate. This connection between the occurrence of the plants and the climate is described by transfer functions. The nature of these transfer functions has to be probabilistic because the climate-biosphere system is a stochastical system. In the Collaborative Research Centre (CRC) 806 project B3 "Our way to Europe" high-resolution lacustrine sediment cores were drilled in March 2010 by Thomas Litt and his working group at Lake Kinneret and Birkat-Ram. The sediment core situated on the Balkans was retrieved at Lake Prespa in November 2009 as part of the CRC project B2.

This study presents the results of local palaeoclimate reconstructions based on methods which are a statistical extension of the concept of biomisation, plant functional types and mutual climatic range (MCR). In more detail the Bayesian Biome Model (BBM) is applied for Lake Kinneret and Ein Gedi and the Bayesian Indicator Taxa Model (BITM) for Lake Prespa. For Birkat Ram the Bayesian Indicator Taxa and Biome Model (BITBM) is newly developed and applied. This method combines the BBM and BITM. Reconstructed are the near surface temperatures, middle troposphere temperatures (850 hPa level), the annual climatic water deficit CWD_{ANN} and the annual precipitation amount P_{ANN}.

All presented palaeoclimate temperature reconstructions except Lake Kinneret share that the surface and the middle troposphere temperature reconstructions are in accordance. It is also shown that the CWD_{ANN} palaeoclimate reconstruction works and does not contradict P_{ANN}. The marginal distribution for CWD_{ANN} is for example for Lake Prespa a reconstructed palaeoclimate variable which allows more identifiable variation than in P_{ANN}. For Lake Prespa there are four, for Birkat Ram three, for Ein Gedi also three and for Lake Kinneret no identifiable time ranges with different climate in the marginal probability density function (pdf)s. In the case that there are time ranges with different climate they are clearly identifiable in the palaeoclimate reconstructions since the marginal distribution profiles before and after differ more or less. Some of these time ranges are compared by application of a Student's t-test for a significance test.

Also presented is an interpolation of local reconstructions situated in the Levant or more precise the Jordan Valley which allows a better assessment of climate changes. The Jordan Valley climate field reconstruction (CFR) results is a dryer palaeoclimate than the modern climate for P_{ANN} and no climate change for the considered temperatures for all fossil sites and considered time slices. This result remains uncertain since there are some difficulties with the climate database.

Contents

Introduction

Today the human influenced climate change and the effects on humankind play an important role in the public and of course also in political discussions. Movies, such as ”An Inconvenient Truth” produced by Al Gore, the Winner of the Nobel Peace Prize 2007, showed these effects to a broad public. For answering questions connected to this topic it is very important to understand the variability of the climate system today and in the past. Nowadays measurements allow an objective quantification in terms of climatological variables. On a global scale these continuous measurements are only available for the last 100 years, for example the dataset provided by the Climate Research Unit (CRU). This time period is too short to examine and understand the variability of the climate system. Historical and palaeo climatology tries to fill this information gap with quantitative climate reconstructions. Historical climatology in this context means the application of all kind of written data as proxy data. They are based on indicators which are available on long time scales and which are influenced essentially by climatic conditions. The nature of these so called proxies could be, for example, biological, chemical or also historical reports. In this work pollen counts are used as proxy. The chapter directly following gives an overview of the science of climate reconstructions focused on this proxy.

One important question connected with climate change is the influence on the migration of humans in the past. The study presented here is a contribution to the Collaborative Research Centre (CRC) 806 ‚Our Way To Europe‘ (`http://www.sfb806.uni-koeln.de/`), funded by the German Research Foundation (DFG). The title of the CRC indicates that this project tries to answer the question of origin of modern man in Europe (which is one part of migration of the humans in the past) and the corresponding driving factors. In more detail, this study presents the results from palaeoclimate reconstructions in the southern Levant and the Balkan regions (CRC subprojects B3 and B2), illustrated by Figure I. It bases on the CRC 806-logo which indicates the two most probable routes of migration of the modern man into Europe. In Figure I the location of the bodies of water where the sediment core are drilled (s. chap. 2.3) are marked. They form the most important database of the palaeoclimate reconstructions presented in this work: Lake Prespa, Lake Kinneret, Ein Gedi and Birkat Ram. Since climate change is one possible driver (s. e. g. CRC homepage or Müller *et al.* 2011) for human migration and all Lakes are situated in the corridor of the eastern migration route this study tries to answer if there is a palaeoclimate change. All required basic mathematical principles for these palaeoclimate reconstructions are introduced in chapter 1. Chapter 2 presents all

Figure I.: Migration routes of the modern human into Europe (based on the CRC 806-logo). Marked are the lake position used in this work.

necessary and used data in this work.

For Lake Prespa (situated on the Balkan) the Bayesian Indicator Taxa Model (BITM) is the selected reconstruction method as presented in chapter 3. Previous applications of the BITM reconstruct T_{Jan}, T_{Jul} and P_{ANN} (s. e. g. Simonis *et al.* 2012) or T_{Jan}^{850hpa} and T_{Jul}^{850hpa} (s. e. g. Stolzenberger 2011). This study introduces a new possible climate variable for the real water availability CWD (s. chap. 2.1.1, eq. 2.1 for details) and tries to show if it is possible to reconstruct palaeoclimate described by CWD. For Lake Prespa this means a palaeoclimate reconstruction based on a three dimensional climate state vector $\vec{C}$ with four different kinds of definition for $\vec{C}$ (s. eq. 3.1). The definition comprises the 2 m surface climate variables T_{Jan}, T_{Jul} and P_{ANN} as well as the middle troposphere variables T_{Jan}^{850hpa} and T_{Jul}^{850hpa}. It tries to show if the results are in accordance.

Chapter 4 presents the result of an application of the Bayesian Biome Model (BBM) on the Lake Kinneret sediment core data which is developed by Litt *et al.* (2012) for the Ein Gedi sediment core. This chapter also tries to answer the same questions as for Lake Prespa: Is it possible to reconstruct CWD or more precisely the annual amount CWD_{ANN}, are the results of different definitions of the climate state vector in accordance and is there a climate change identifiable? Due to the fact that for the BBM different settings are possible the different setups were tested and compared. For Lake Kinneret is also tested if it is possible to use another method for palaeoclimate reconstruction, the pollen-ratio model (PRM) (s. chap. 4.4 for details). Since later on in chapter 6 the Ein Gedi palaeoclimate reconstruction with other settings as in Litt *et al.* (2012) is required chapter 4 presents the adaptations and the result for CWD_{ANN}.

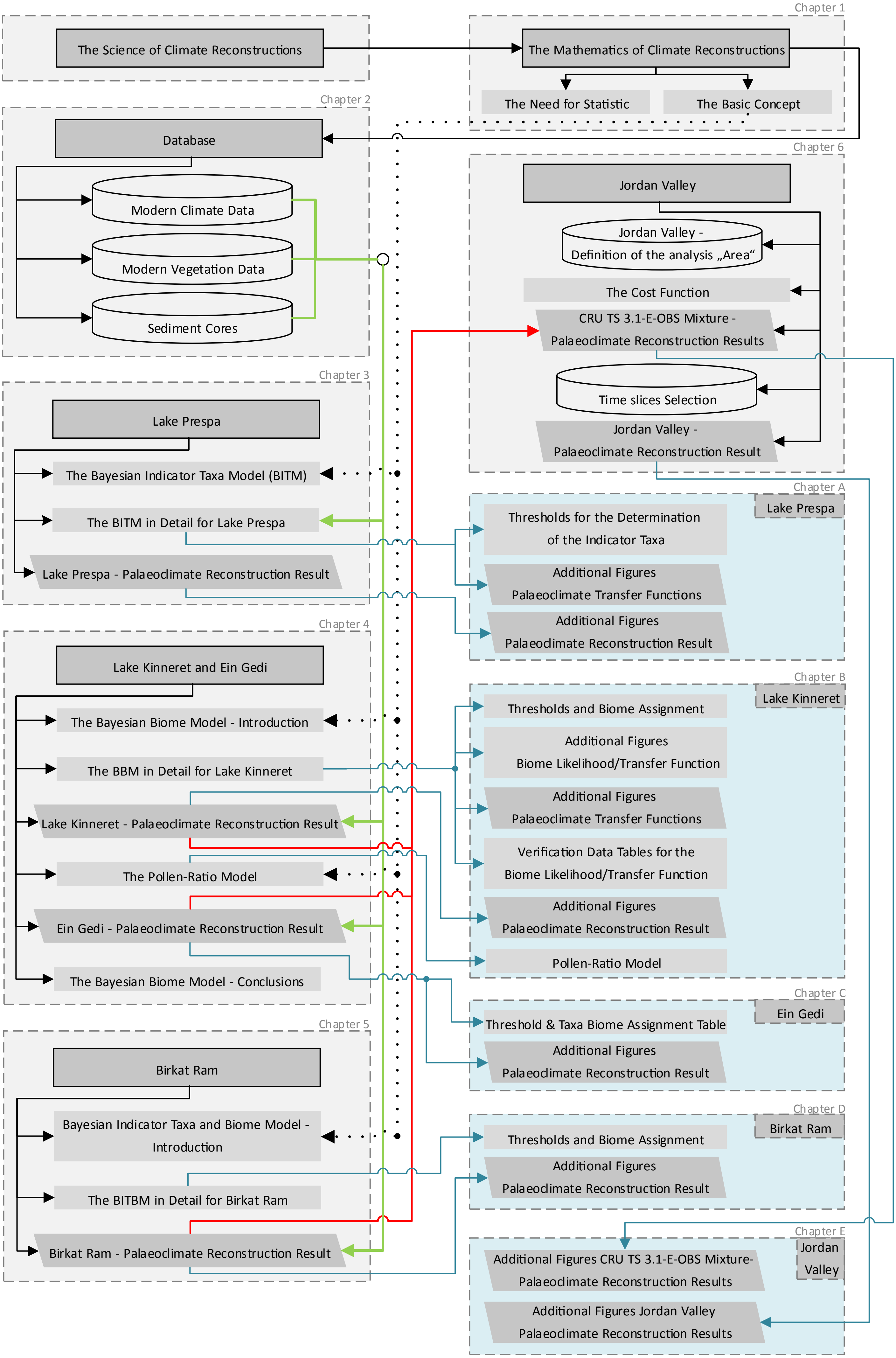
In chapter 5 is investigated for the Birkat Ram sediment core data if it is possible to reconstruct palaeoclimate with CWD. Additionally a new method is developed which is situated between the BBM and BITM since the requirements for an application of these methods are not exactly given (s. chap. 5.1 for details): The Bayesian Indicator Taxa and Biome Model (BITBM) combines biome (BBM) and taxon (BITM) information. It is analysed if this is possible and the reconstructed palaeoclimate is discussed in chapter 5.3.

Due to the fact that Lake Kinneret, Birkat Ram and Ein Gedi are all situated in the Jordan Valley chapter 6 presents an approach summarizing the local palaeoclimate reconstructions in a simple climate field reconstruction (CFR). Simple in this context means that no model like e. g. the one used in Gebhardt (2003); Simonis (2009) is applied. In this work their approach is simplified and applied to try to answer: if there is palaeoclimate change indicated in one of the local palaeoclimate reconstructions, is it also identifiable in the whole Jordan Valley?

The following page shows the structure of this thesis. It is a flowchart showing the relation of the different chapters and the appendix. The flowchart includes the appendix which is highlighted with a blue colour as background. The dotted lines show where to find the mathematics for each palaeoclimate reconstruction. The green coloured lines connect the database to the palaeoclimate reconstructions. The cyan coloured lines indicate the connection of the appendix chapters to the rest of the work. The red lines show how the Jordan Valley palaeoclimate reconstruction is embedded in the work.

Directly after the appendix follows the Register part. It includes:

- List of Figures and Tables (p. 227),
- List of Acronyms,
- List of Symbols which lists the mathematical symbols divided in three parts dependent on the mathematical dimension (scalar quantity, vectorial quantities and matrices),
- List of Taxa with all used taxa listed in alphabetical order. Additionally the shortcuts (if defined), the sediment core and the assigned biome type (if definded, s. 4.2.2 for details) are listed.
- a Glossary and
- the bibliography. The bibliography is divided in parts:
  - Scientific Books
  - Parts of Scientific Collections
  - Scientific Articles
  - Scientific Thesis (PhD thesis, Master thesis and diploma thesis)
  - Electronic Data Sources (databases, websites and programs)
  - Other Sources (the rest)


Chapter 1
The Science of Climate Reconstructions
The Mathematics of Climate Reconstructions
The Need for Statistic
The Basic Concept
Chapter 2
Database
Modern Climate Data
Modern Vegetation Data
Sediment Cores
Chapter 6
Jordan Valley
Jordan Valley - Definition of the analysis „Area“
The Cost Function
CRU TS 3.1-E-OBS Mixture - Palaeoclimate Reconstruction Results
Time slices Selection
Jordan Valley - Palaeoclimate Reconstruction Result
Chapter 3
Lake Prespa
The Bayesian Indicator Taxa Model (BITM)
The BITM in Detail for Lake Prespa
Lake Prespa - Palaeoclimate Reconstruction Result
Chapter A
Lake Prespa
Thresholds for the Determination of the Indicator Taxa
Additional Figures Palaeoclimate Transfer Functions
Additional Figures Palaeoclimate Reconstruction Result
Chapter 4
Lake Kinneret and Ein Gedi
The Bayesian Biome Model - Introduction
The BBM in Detail for Lake Kinneret
Lake Kinneret - Palaeoclimate Reconstruction Result
The Pollen-Ratio Model
Ein Gedi - Palaeoclimate Reconstruction Result
The Bayesian Biome Model - Conclusions
Chapter B
Lake Kinneret
Thresholds and Biome Assignment
Additional Figures Biome Likelihood/Transfer Function
Additional Figures Palaeoclimate Transfer Functions
Verification Data Tables for the Biome Likelihood/Transfer Function
Additional Figures Palaeoclimate Reconstruction Result
Pollen-Ratio Model
Chapter C
Ein Gedi
Threshold & Taxa Biome Assignment Table
Additional Figures Palaeoclimate Reconstruction Result
Chapter 5
Birkat Ram
Bayesian Indicator Taxa and Biome Model - Introduction
The BITBM in Detail for Birkat Ram
Birkat Ram - Palaeoclimate Reconstruction Result
Chapter D
Birkat Ram
Thresholds and Biome Assignment
Additional Figures Palaeoclimate Reconstruction Result
Chapter E
Jordan Valley
Additional Figures CRU TS 3.1-E-OBS Mixture- Palaeoclimate Reconstruction Results
Additional Figures Jordan Valley Palaeoclimate Reconstruction Results

The Science of Climate Reconstructions

This chapter gives an overview over the current state of art in the science of palaeoclimate reconstructions, focused on pollen, since this is the selected proxy in this work. The basic assumptions and mathematical tools for statistical palaeoclimate reconstruction as applied in chapter 3, 4, 5 are introduced in the next chapter.

The science of palaeoclimate reconstructions has made a big progress in the last two decades. Especially the need for statistics for palaeoclimate reconstructions has become evident in the palaeo-science community, visible in the increasing number of publications in the past ten years, e. g. Li *et al.* (2010); McShane, Wyner (2011); Ohlwein, Wahl (2012); Tingley *et al.* (2012); Tingley, Huybers (2010); Tolwinski-Ward *et al.* (2014); Werner *et al.* (2013). The Intergovernmental Panel on Climate Change (IPCC) also emphasizes this in the Fifth Assessment Report (AR5) (chapter 5, s. Masson-Delmotte *et al.* (2013)). In detail, the need of assessing the uncertainty of palaeoclimate reconstructions and the relevance of Bayesian Hierarchical Model (BHM) for this aim are discussed in the AR5.

The application of pollen as a proxy for palaeoclimate reconstructions has a long tradition in palaeo science. Ohlwein, Wahl (2012) and Bartlein *et al.* (2011) maintain the fact that pollen-climate transfer methods are used for at least 65 years. Ohlwein, Wahl (2012) also give an overview of the development of probabilistic pollen-climate transfer methods and identify the first approaches using mathematics for describing the relationship between pollen and climate (Bernabo (1981), Howe, Webb III (1983) and Bartlein *et al.* (1984)). The connection between pollen and climate in these publications is described with a deterministic multiple linear model and stochastic errors. Multiple linear in this context implies that different climate variables are incorporated in the model (Ohlwein, Wahl 2012). As explained in Ohlwein, Wahl (2012) "*these models could exhibit high explained variability in calibration, but were found to be non general enough for reconstructions*". Calibration, in this context, means the fitting of the pollen-climate connection in the modern time frame.

The next step in the development of proxy-climate transfer functions was the expansion of the perspective from local reconstruction methods to CFR methods. Local reconstructions imply that the reconstruction is applied for one point at the earth's surface and usually only representative and interpretable for this geographical point. This locality refers to the place where the proxy time series is received. In this study, the locality is the drilling location of the sediment cores. One of the first steps from local reconstructions to CFR were explored by Cook *et al.* (1994), who apply a canonical correlation analysis (CCA) on tree rings. As Ohlwein, Wahl (2012) pointed out: *"This is not a systematic field reconstruction method per se*", but it is a "*closely-related method*" which "*forms a truncated basis for a empirical orthogonal function (EOF)*". As also mentioned in Ohlwein, Wahl (2012) EOFs are the basis of CFR. EOFs are an important tool in the science of climatology to analyse the most important spatial and temporal structures of a climatic field.

After these first efforts with CFR the perception that the application of the Bayes theorem (for details s. chapter 1) and the consideration of the full probability density function (pdf) at least for local reconstruction could enhance the analysis of uncertainty of the climate of the past was an important step forward. The use of the Bayes theorem in palaeoclimatology was first mentioned, as far as we know, in Birks (1998). Birks (1998) also gives a good overview of the regression based methods and presents a good discussion of the respective advantages and disadvantages of the methods. A more actual publication in this context is Bartlein *et al.* (2011). Actually the Bayes theorem and pdfs were applied in Robertson *et al.* (1999), a publication also mentioned in Birks (1998). Robertson *et al.* (1999) used the results of an application of the Bayes theorem and a kernel density estimator to estimate the joint pdf for carbon isotope index taken from tree rings and mean temperature of July and August (T_{JA}^{mean}). Robertson *et al.* (1999) also compared this "*novel*" method with regression based methods.

One step further in the development of probabilistic transfer function was the first detailed introduction of the BHM by a Finnish research group (Schölzel (2006), section 2.2.1, p. 16). In general, BHM allows the separation of uncertainty in different parts called levels. The Finnish research group used the abundance of Chironomidaes (non-biting midges) as proxy in their Bayesian response model (Vasko *et al.* (2000), Toivonen *et al.* (2001), Korhola *et al.* (2002)).

A collaboration of climatologists and palaeontologists, initialized by Prof. Dr. Andreas Hense and Prof. Dr. Thomas Litt, developed pollen-climate transfer functions. A number of publications are the result of this collaboration from now on called "*Bonner Group*". The first result was the statistical extension of the mutual climatic range (MCR) method (s. chapter 3, p. 23 for details) to the so called pdf method (Kühl 2002; Kühl *et al.* 2002). The method presented in Kühl *et al.* (2002) bases on a bivariate normal distribution for the description of the transfer function for January temperature (T_{Jan}) and July temperature (T_{Jul}). Gebhardt (2003); Gebhardt *et al.* (2008) compared two probability distributions as basis for the transfer functions: a kernel density estimator and a Gaussian mixture model. He also presents an approach for CFR, which differs from those mentioned above, based on data assimilation usually used in numerical weather prediction: a physical coherent spatial interpolation of the local pdf reconstructions (for more detailed information, s. chap. 6).

Another non-probabilistic pollen based reconstruction method, which fits into the context of BHM, is presented in Haslett *et al.* (2006): the modern analogue technique (MAT), which is in detail described in chapter 4. In a nutshell the MAT uses the similarity of the structure of modern pollen samples and those of the past. The reconstructed climate variables in

Haslett *et al.* (2006) are the mean temperature of the coldest month (MTCO) and the growing degree days above 5 °C (GDD5).

The final implementation of the pdf-method in BHM was developed by the ”*Bonner Group*” and is presented in the PhD thesis of Schölzel (2006) who applied the so called BITM in the Near East. Schölzel (2006) also studies in detail how the aforementioned classical regression based methods fit into the context of BHM. He also introduced the BBM based on plant functional types (PFT) for the palaeoclimate reconstruction of Ein Gedi (Israel), which was enhanced by Litt *et al.* (2012). The considered transfer function describe a multivariate pdf for mean winter temperature (T_{DJF}), mean summer temperature (T_{JJA}) and annual precipitation amount (P_{ANN}). The BBM is also applied for the sediment core drilled at Lake Kinneret (Israel) and in detail described in chapter 4. Motivated by the success of reconstructing P_{ANN}, Simonis (2009) explores the potential of combining the BITM for T_{Jan}, T_{Jul} and P_{ANN} with the variational method introduced by Gebhardt (2003).

Pollen are not the only possible proxy for which probabilistic reconstructions should be developed, e. g. Robertson *et al.* (1999). BHM offer the possibility to consider/handle different kinds of proxies in a mathematically consistent way for reconstructing a conditional pdf of the past climate given a mixture of different proxy data, often called multi-proxy approach. Hughes, Ammann (2009) classify the different approaches for reconstructions of the past climate, which is shown in Figure II. This scheme classifies the palaeoclimate reconstruction methods on the basis of two characteristics: the understanding of the forming processes of the proxy archives and the amount of available data required for each method. The classical multi-proxy cross-validation and also the CFR are classified compared to BHMs as methods with less processes information but all three approaches need a lot of data to achieve results.

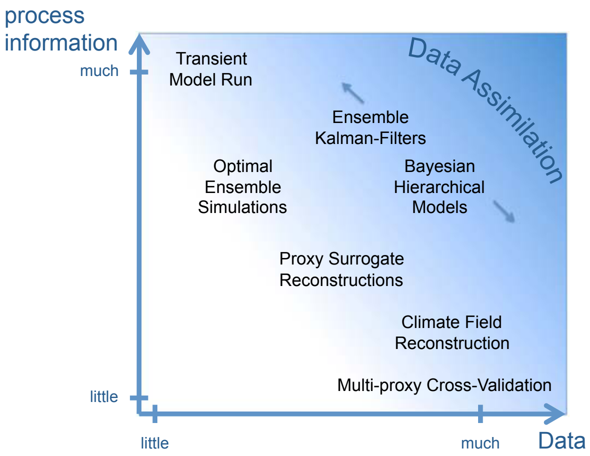


Figure II.: Scheme for the classification of palaeoclimate reconstruction methods, redrawn after Hughes, Ammann (2009)

Garreta (2010) and Garreta *et al.* (2010) provide the next important contribution to the development of palaeoclimate reconstructions. The study presents the capabilities of estimating a palaeoclimate reconstruction with the help of a Bayesian framework with a process-based vegetation model.

Li *et al.* (2010) present a BHM for reconstructing the past climate based on tree rings, borehole temperature and pollen abundance. Since they use synthetic proxies generated from the output of a climate model it is also possible to quantify the reconstructed uncertainty of the reconstructed annual mean temperature (T_{ANN}^{mean}). Tingley, Huybers (2010) present an extension of the approach from Li *et al.* (2010): The reconstruction of the "*temporal evolution of climate field from incomplete data*".

If we switch back to the reconstruction methods that incorporate the knowledge of the process understanding, Birks *et al.* (2010) published an extensive discussion of the different reconstruction methods available for pollen as proxy, in detail these are "*the indicator-species approach involving bioclimate envelope modelling, the assemblage approach involving modern analogue techniques and response surfaces and the multivariate calibration-function approach*". Bartlein *et al.* (2011) also give an overview of reconstruction methods but more focused on the global available reconstructions for the mid-Holocene (around 6000 years before present (y. BP)) and the Last Glacial Maximum (around 21000 y. BP). In detail, they present a short discussion of the methods and then focus on a synthesis of the reconstruction results for four temperature and two moisture climate variables.

The next contribution of the "*Bonner Group*" was a new description for transfer functions based on the BITM for T_{Jan} and T_{Jul}. In the diploma thesis of Stolzenberger (2011) transfer functions are estimated with a generalized linear model (GLM) for different taxa and different climate datasets. The results are compared with statistical verification methods (in detail ROC-curves, Brier skill score and relative entropy). For the modern time frame, field reconstructions based on transfer function estimated with the GLM, the Gaussian mixture model and the normal distribution are compared. Also a GLM-BITM reconstruction for Lake Holzmaar is presented.

Tingley *et al.* (2012) (together with the publication Tingley, Huybers (2010) is one of the most general BHM description in connection with palaeoclimate reconstructions. The BHM in Tingley *et al.* (2012) is based on the joint probability of six random variables and leads, after an extensive discussion of the assumptions which are used for a simplification, to five different levels of uncertainty. These levels allow to quantify the uncertainty contribution for the palaeoclimate reconstruction of multiple proxy data, multiple instrumental data, the reconstructed climate state and physical process(es) which form the proxy data and the instrumental data. How the separation into level is done mathematically in general is described in chapter 1.2. Tingley *et al.* (2012) also show how the classical methods like e. g. classical linear regression fit in this complex BHM.

In comparison with Tingley *et al.* (2012) Ohlwein, Wahl (2012) set the focus on one proxy named pollen. Ohlwein, Wahl (2012) give an overview of the development of "*probabilistic pollen-climate transfer methods*" till 2012 and show how MCR resp. the BITM, MAT, PFT and biomisation resp. the BBM, and process-based vegetation models like e. g. described in Garreta (2010) fit in the framework of BHM for palaeoclimate reconstructions. Also the "*selection of prior distributions*" is discussed. As an example a case study with the PRM is presented, also used in this study and therefore in detail described in chapter 4, which is a simplification of the MAT for the ratio of two pollen taxa.

Parnell *et al.* (2012) also used a BHM for "*modelling the uncertainties in palaeoclimate reconstructions*". They analyse real pollen proxy data with an extension of the BHM introduced in Tingley *et al.* (2012). They incorporate the uncertainty of the chronological dating process of the pollen samples of sediment cores. The special focus of this publication is the estimation of the BHM level called likelihood, which is the probability for the pollen proxy data given the climate data, in situation where "*repeated calls to the likelihood are costly*".

Garreta *et al.* (2012) present an extension of the reconstruction method presented in Garreta *et al.* (2010). Garreta *et al.* (2012) study by using the vegetation model LPJ-GUESS, imbedded in a BHM, the link between modern pollen samples and modern climate in Europe.

The publication of Crucifix (2012) is a very general review of "*traditional and novel approaches to palaeoclimate modelling*". The three main topics of the review are the process-based, simulator-based and statistical modelling. Simulator-based modelling stands for the evaluation of numerical general circulation model (GCM) developed and used by climatologists "*which encapsulate available knowledge on a huge variety of climate processes at different spatial and temporal scales*" (Crucifix 2012). An example for a GCM is mentioned in chapter 1.1. Crucifix (2012) compares all three approaches, shows the advantages and disadvantages and discusses how to connect them with BHM.

In 2013 some studies which apply the BHM developed by Tingley, Huybers (2010) and Tingley *et al.* (2012) were published. First, Werner *et al.* (2013) evaluate a simplified version of the just mentioned BHM and also a CCA for CFR over Europe. The evaluation is performed by using proxy data generated by the National Center for Atmospheric Research (NCAR) Community Climate System Model, version 1.4 (CCSM1v4). Secondly, Tingley, Huybers (2013) use "*instrumental, tree-ring, ice-core and lake-sediment records*" for analysing "*temperature extremes at high northern latitudes*".

Mairesse *et al.* (2013) applied the above mentioned data assimilation to estimate which climatic field shows the best agreement with the information from a climate model with its implemented forcings and local proxy based palaeoclimate reconstructions. For the Mid Holocene, a period roughly from 7 000 y. BP - 5 000 y. BP, they compare the results of one model with the opportunity to include or exclude data assimilation named LOVECLIM1.2 and three other GCMs.

Guillot *et al.* (2013) present a new method for CFR named GraphEM. In detail, it is an adopted ”*approach based on multivariate linear regression*” which applies an expectation maximization (EM) algorithm. Briefly summarized, all available proxy data and instrumental temperature data are regarded as realisation of a multivariate random vector with a Gaussian distribution. The reconstruction of the pre-instrumental temperature data in this approach is treated as missing value problem, in which Gaussian Markov random fields are used to estimate the covariance matrix of the regarded Gaussian distribution.

Mukhopadhyay, Bhattacharya (2013) present an BHM approach called Bayesian Semiparametric Model. Their work is based on the above mentioned work of Haslett *et al.* (2006). In detail Mukhopadhyay, Bhattacharya (2013) present a BHM for Chironomidaes and Pollen whereat they ”*model the species assemblages*” with a ”*zero-inflated multinomial distribution*” and ”*the species response functions*” with ”*Dirichlet-process based Gaussian mixtures*”. The publication presents no palaeoclimate reconstruction but focuses on assessment of the model adequacy via cross-validation for the aforementioned BHM.

Janson, Rajaratnam (2014) introduce another methodology for statistical palaeoclimate reconstructions. The main focus of their publication is the development of a method ”*which combines quantile regression (QR) with autoregressive (AR) structures in the residuals*”. As a case study a palaeoclimate reconstruction for the mean global northern hemisphere temperature of the last millennium is presented. Additionally, the method is evaluated with the so called ”forward modelling” approach in order to get an impression how good the connection between the proxies and the temperatures is reflected.

Tolwinski-Ward *et al.* (2014) explore the possibilities of probabilistic BHM in order to reconstruct climate with the proxy tree-ring width. In order to examine the sensitivity of the method, numerical experiments with synthetical proxies are performed. The last part of the article demonstrates again the flexibility of BHMs with a multiproxy based reconstruction of local temperature and moisture of the Methuselah Walk, a locality in the White Mountains of California. The second incorporated proxy is the isotope ratio of $\delta^{13}C$.

1

The Mathematics of Climate Reconstructions

In the introduction the need of palaeoclimate reconstructions was revealed as the lack of quantitative information for the climate of the past. In detail, this means that measurements which quantify the climate with physically measurable and comparable quantities like for example the air temperature, wind speed or precipitation amount are only available for a maximum timeframe of about 200 years on some sparse locations. The longest continuous weather record of this kind is available from 1780 until present at the Meteorological Observatory Hohenpeissenberg in Southern Germany (Kapala 2003; *Das Meteorologische Observatorium Hohenpeißenberg*). On a global scale, continuous measurements are only available for the last 100 years, for example the datasets provided by the CRU which are used in this study (s. chap. 2.1, p. 7).

If quantitative climatic information is needed for longer time slices or other locations it requires another type of data source. The science of palaeoclimatology tries to fill this information gap with statistical palaeoclimate reconstructions. The basis are indicators that are available on long time scales and which are influenced essentially by climatic conditions. These so called *proxies* or *Climate Proxy*, could be for example biological, chemical or also historical reports.

This chapter introduces the basic idea and principles for the palaeoclimate reconstructions presented in chapter 3, 4 and 5.

1.1. The Need for Statistic

Climate reconstructions base on transfer functions which connect the climate information and the proxy information. The mathematical formulation of these transfer functions is a challenging task and will be described in the following.

First of all one has to consider the general nature of the transfer function formulation. The climate system is a complex system (Hense 2002, 2005). In the considerations presented here, the climate system is defined as a coupled system of several connected subsystems. In detail, these are the atmosphere, the hydrosphere, the cryosphere, the biosphere and the lithosphere. Every subsystem includes many processes. Some of these processes are connected and some not, but every process has uncertainties and many processes are non-linear. Some processes are not known and some are not completely understood. These complexities are illustrated for example in Figure 1.1, which shows all incorporated subsystems and phenomena in the CCSM3 developed from the NCAR.

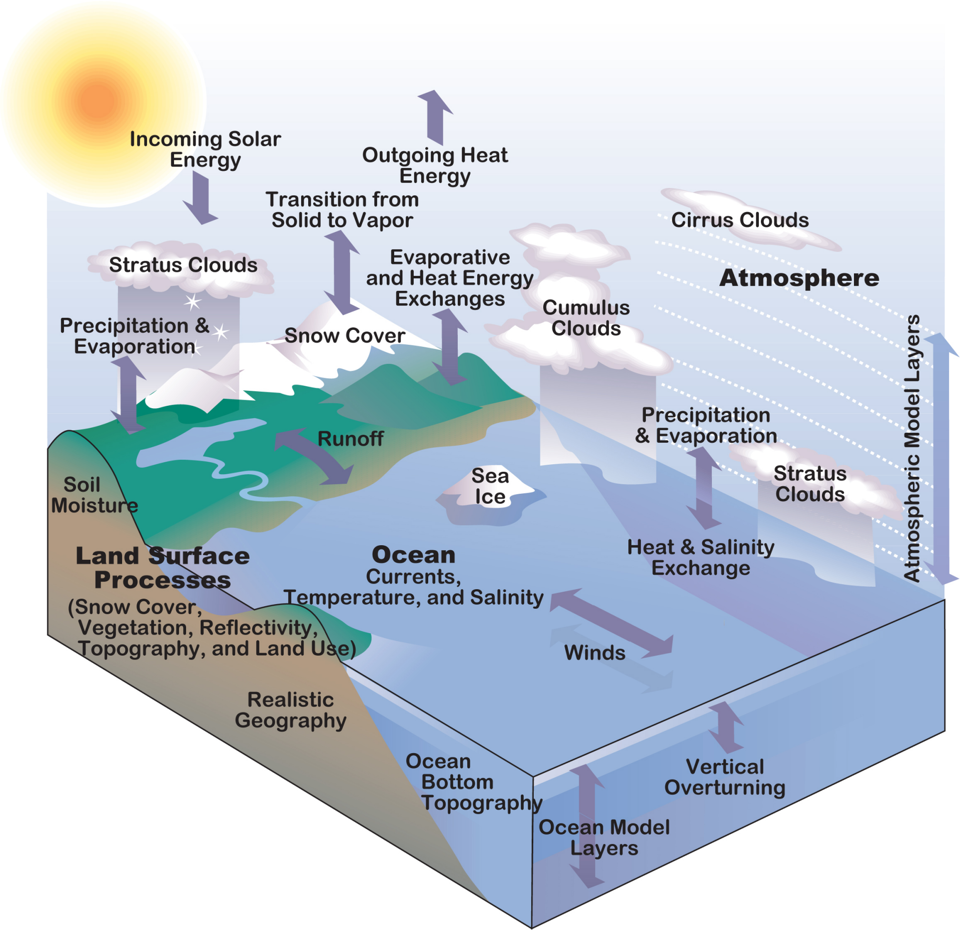


Figure 1.1.: The complex earth climate system and related processes incorporated in the CCSM3. **source:** Atmospheric Research (UCAR) (2004)

Alone from this point of view it has to be pointed out, that every calculation that bases on the climate system or parts of it has to regard these complexities and non-linearities. These facts are extensively discussed in Gebhardt (2003); Li *et al.* (2010); Ohlwein, Wahl (2012); Schölzel (2006); Simonis (2009); Stolzenberger (2011); Tingley *et al.* (2012). In science, these facts are considered by the use of statistics, more precise by regarding all used variables as random variables.

The definition of random variables Z, denoted with capital letters, is an important step for the mathematical description of statistical quantities (s. e. g. Wilks 2011). The general definition combines all important information of the considered quantity, which are

- the realisation z, e. g. the measurement of the summer temperature, denoted with small letters,
- the sample space $\mathcal{S}$ that is the range of all theoretically possible realisations,
- the probability $\mathbb{P}(\cdot)$ for each realisation. If the sample space is continuous and the probability function is differentiable it is also possible to use the pdf $f(\cdot)$.

This leads to the definition of a univariate random variable (eq. 1.1a). In the case of a vectorial realisation the random variable is a multivariate random variable (eq. 1.1b):

$$Z := \left\{ \left(z,\ \mathbb{P}(z) \right),\ z\ \in \mathcal{S} \right\} = \left\{ \left(z,\ f(z) \right),\ z\ \in \mathcal{S} \right\}, \tag{1.1a}$$

$$\vec{Z} := \left\{ \left(\vec{z}, \mathbb{P}(\vec{z}) \right), \vec{z} \in \mathcal{S} \right\} = \left\{ \left(\vec{z}, f(\vec{z}) \right), \vec{z} \in \mathcal{S} \right\}. \tag{1.1b}$$

Specially the consideration about the sample space is important for the specification of the used cumulative distribution function (CDF) as probability.

1.2. The Basic Concept

The definition of statistical transfer functions requires a definition of the regarded random variables. Two essential groups of random variables can be distinguished: the group of the recent and the past ones. This division is based on the fact that different scales in space and time are regarded (Schölzel 2006; Tingley *et al.* 2012). Specially Tingley *et al.* (2012) discussed extensively difficulties which arise from the different scales that are reflected in the underlying processes.

As mentioned in the last section, transfer functions connect climate information and proxy information. Due to this and the arguments mentioned in this section, the minimum number of necessary random variables are:

- the recent climate variables or more general environmental variables $\vec{C} = \left\{ \left(\vec{c}, f\left(\vec{c} \right) \right), \vec{c} \in \mathcal{K} \right\}$,
- the recent proxy variables $\vec{P} = \left\{ \left(\vec{p}, f\left(\vec{p} \right) \right), \vec{p} \in \mathcal{P} \right\}$,
- the past climate environmental variables $\vec{C}_0 = \left\{ \left(\vec{c}_0, f\left(\vec{c}_0 \right) \right), \vec{c}_0 \in \mathcal{K}_0 \right\}$,
- and the past proxy variables $\vec{P}_0 = \left\{ \left(\vec{p}_0, f\left(\vec{p}_0 \right) \right), \vec{p}_0 \in \mathcal{P}_0 \right\}$,

In this notation, all random variables that belong to the group called past are marked with an index zero. Past, in detail, means that the random variables $\vec{C}_0$ and $\vec{P}_0$ are given at historic time t_0 as concluded by Schölzel (2006). Transfer functions in this statistical perception are expressed as the joint probability or joint pdf of the involved random variables:

$$f_{\vec{C},\vec{P}}\left(\vec{c},\vec{p}\right), \tag{1.2a}$$

$$f_{\vec{C}_0,\vec{P}_0}\left(\vec{c}_0,\vec{p}_0\right). \tag{1.2b}$$

This interpretation is equivalent to a consideration of a random variable for the past time frame $\vec{R}_0=\left\{\left(\vec{c}_0,\vec{p}_0,f_{\vec{C}_0,\vec{P}_0}\left(\vec{c}_0,\vec{p}_0\right)\right),\vec{c}_0\in\mathcal{K}_0,\vec{p}_0\in\mathcal{P}_0\right\}$ and the recent time frame $\vec{R}=\left\{\left(\vec{c},\vec{p},f_{\vec{C},\vec{P}}\left(\vec{c},\vec{p}\right)\right),\vec{c}\in\mathcal{K},\vec{p}\in\mathcal{P}\right\}$. The used pdf for this random variables are those defined in equation 1.2. The notation for pdfs in equation 1.2 and in the following chapters summarizes two facts: The subscript indicates which random variable is considered and in brackets the dedicated realisation. As mentioned in the definition of a random variable in the last section, pdfs are only available for continuous and differentiable random variables. Since small letters denote pdfs, it has to be mentioned that the connection between discrete and continuous random variables is straightforward and that the application of pdfs from now on could be made w. l. o. g. (Schölzel 2006).

The joined pdfs in equation 1.2 or more general probabilities are the origin for a BHM as defined in e. g. in Tingley *et al.* (2012) or Ohlwein, Wahl (2012). In more detail, BHMs apply the Bayes theorem (Bayes 1763) on the joint probability or joint pdf and split it into a product of conditional probabilities called levels. The complexity of the BHM depends on the number of included random variables in the joint probability. Tingley *et al.* (2012) for example include 5 levels in the BHM and Ohlwein, Wahl (2012) four.

For the minimum number of involved random variables in equation 1.2 the Bayes theorem leads to

$$f_{\vec{C}|\vec{P}}\left(\vec{c}|\vec{p}\right)\cdot f_{\vec{P}}\left(\vec{p}\right)=f_{\vec{P}|\vec{C}}\left(\vec{p}|\vec{c}\right)\cdot f_{\vec{C}}\left(\vec{c}\right), \tag{1.3a}$$

$$f_{\vec{C}_0|\vec{P}_0}\left(\vec{c}_0|\vec{p}_0\right)\cdot f_{\vec{P}_0}\left(\vec{p}_0\right)=f_{\vec{P}_0|\vec{C}_0}\left(\vec{p}_0|\vec{c}_0\right)\cdot f_{\vec{C}_0}\left(\vec{c}_0\right). \tag{1.3b}$$

In palaeoclimate reconstructions the main interest is the conditional probability climate given proxy data. Equation 1.4 is the rearranged equation 1.3 with the commonly used notation for BHM, which is also used e. g. in Simonis (2009) or Schölzel (2006):

$$f_{\vec{C}|\vec{P}}\left(\vec{c}|\vec{p}\right)=\frac{f_{\vec{P}|\vec{C}}\left(\vec{p}|\vec{c}\right)\cdot\pi_{\vec{C}}\left(\vec{c}\right)}{m_{\vec{P}}\left(\vec{p}\right)}, \tag{1.4a}$$

$$f_{\vec{C}_0|\vec{P}_0}\left(\vec{c}_0|\vec{p}_0\right)=\frac{f_{\vec{P}_0|\vec{C}_0}\left(\vec{p}_0|\vec{c}_0\right)\cdot\pi_{\vec{C}_0}\left(\vec{c}_0\right)}{m_{\vec{P}_0}\left(\vec{p}_0\right)}. \tag{1.4b}$$

In this notation $\pi_{\dots}(\cdot)$ generally denotes the prior distribution and $m_{\dots}(\cdot)$ the marginal distribution. The remaining part of the right side of the equations $f_{\vec{P}|\vec{C}}\left(\vec{p}_0|\vec{c}_0\right)$, is called the likelihood and the left side of the equations is called the posterior probability/pdf.

The main assumption in this work and also in Gebhardt (2003); Schölzel (2006); Simonis (2009); Tingley *et al.* (2012) is that the connection between environment and proxy in the recent time frame is the same as in the past. This is expressed by

$$f_{\vec{C}_0|\vec{P}_0}\left(\vec{c}_0|\vec{p}_0\right) = f_{\vec{C}|\vec{P}}\left(\vec{c}_0|\vec{p}_0\right) = \frac{f_{\vec{P}|\vec{C}}\left(\vec{p}_0|\vec{c}_0\right) \cdot \pi_{\vec{C}}\left(\vec{c}_0\right)}{m_{\vec{P}}\left(\vec{p}_0\right)}. \tag{1.5}$$

The interpretation of this equation is that the connection between proxy and climate is still stochastic, but the joint distribution of recent and proxy variables is identical (Schölzel 2006). In more detail, this means that the transfer function is established on the modern proxy and environmental random variables but evaluated with the realisations of the past.

The modern point of view, as shown by Schölzel (2006), is to consider the pdf with all involved random variables including an additional random variable $\vec{\Theta}$. This random variable describes the statistical model and it is often introduced in BHM (e. g. Ohlwein, Wahl 2012; Schölzel 2006; Tingley *et al.* 2012). The joint probability without this random variable $\vec{\Theta}$ is obtained by integration over the complete parameter space $\mathcal{V}_\Theta$ which is called marginalisation and done in equation 1.6 and 1.7. In this work (chap. 3.2, 4.2 and 5.2) the transfer functions are estimated via maximum likelihood estimation (MLE) as in Gebhardt (2003); Kühl *et al.* (2002); Schölzel (2006); Simonis (2009). Schölzel (2006) shows how this method is "*embedded in the general concept of statistical transfer functions*". The results are summarized in the following.

The analogue pdf to $f_{\vec{C}_0|\vec{P}_0}\left(\vec{c}_0|\vec{p}_0\right)$ in equation 1.5 in this point of view is

$$f_{\vec{C}_0|\vec{P}_0,\vec{C},\vec{P}}\left(\vec{c}_0|\vec{p}_0,\vec{c},\vec{p}\right) = \int\limits_{\mathcal{V}_\Theta} \frac{f_{\vec{C}_0,\vec{\Theta},\vec{P}_0,\vec{C},\vec{P}}\left(\vec{c}_0,\vec{\theta},\vec{p}_0,\vec{c},\vec{p}\right)}{f_{\vec{P}_0,\vec{C},\vec{P}}\left(\vec{p}_0,\vec{c},\vec{p}\right)} d\vec{\theta}. \tag{1.6}$$

The aforementioned marginalisation and the Bayes theorem are applied to obtain this pdf. Schölzel (2006) applies again the Bayes theorem and uses the BHM notation to get

$$\begin{aligned} & f_{\vec{C}_0|\vec{P}_0,\vec{C},\vec{P}}\left(\vec{c}_0|\vec{p}_0,\vec{c},\vec{p}\right) \\ & = \int\limits_{\mathcal{V}_\Theta} f_{\vec{P}_0|\vec{C}_0,\vec{\Theta},\vec{C},\vec{P}}\left(\vec{p}_0|\vec{c}_0,\vec{\theta},\vec{c},\vec{p}\right) \frac{\pi_{\vec{C}_0|\vec{\Theta},\vec{C},\vec{P}}\left(\vec{c}_0|\vec{\theta},\vec{c},\vec{p}\right)}{m_{\vec{P}_0|\vec{C},\vec{P}}\left(\vec{p}_0|\vec{c},\vec{p}\right)} \pi_{\vec{\Theta}|\vec{C},\vec{P}}\left(\vec{\theta}|\vec{c},\vec{p}\right) d\vec{\theta} \end{aligned} \tag{1.7}$$

Then he assumes that

1.) ”$\vec{C}_0$ *is prior independent of* $\vec{\theta}$, $\vec{C}$ *and* $\vec{P}$”, which reduces $\pi_{\vec{C}_0|\vec{\Theta},\vec{C},\vec{P}}\left(\vec{c}_0|\vec{\theta},\vec{c},\vec{p}\right)$ to $\pi_{\vec{C}_0}\left(\vec{c}_0\right)$ and

2.) ”$\vec{P}_0$ *independent of* $\vec{C}$ *and* $\vec{P}$” which reduces $m_{\vec{P}_0|\vec{C},\vec{P}}\left(\vec{p}_0|\vec{c},\vec{p}\right)$ to $m_{\vec{P}_0}\left(\vec{p}_0\right)$.

Schölzel (2006) also concludes that MLE of the ”*optimal parameters* $\vec{\theta}_{opt}$” is equivalent to a substitution of $\pi_{\vec{\Theta}|\vec{C},\vec{P}}\left(\vec{\theta}|\vec{c},\vec{p}\right)$ with Dirac's delta function $\delta\left(\vec{\theta}-\vec{\theta}_{opt}\right)$. This function ”*satisfies the criteria for probability density functions*” (Schölzel 2006). All these considerations allow to evaluate the integral in equation 1.7. The result is that the first pdf in the integral is then independent of $\vec{C}$ and $\vec{P}$ given $\vec{\theta}_{opt}$:

$$f_{\vec{C}_0|\vec{P}_0,\vec{C},\vec{P}}\left(\vec{c}_0|\vec{p}_0,\vec{c},\vec{p}\right) = f_{\vec{P}_0|\vec{C}_0,\vec{\Theta}_{opt}}\left(\vec{p}_0|\vec{c}_0,\vec{\theta}_{opt}\right)\frac{\pi_{\vec{C}_0}\left(\vec{c}_0\right)}{m_{\vec{P}_0}\left(\vec{p}_0\right)} \tag{1.8a}$$

$$= f_{D(\vec{\theta}_{opt})}\left(\vec{p}_0|\vec{c}_0\right)\frac{\pi_{\vec{C}_0}\left(\vec{c}_0\right)}{m_{\vec{P}_0}\left(\vec{p}_0\right)} \tag{1.8b}$$

The last step in equation 1.8 from 1.8a to 1.8b reflects the aforementioned fact that the transfer function or more precise the corresponding parameters $\vec{\theta}_{opt}$ are estimated on realisations of modern proxy and environmental random variables but evaluated with the realisations of the past. Compared to equation 1.5 here it is transferred to the modern point of view with the more elaborated pdf on the left side. With that knowledge in background it is possible to use the notation of equation 1.5 as shortcut in the remaining part of this work since equation 1.8b shows that the right sides are equivalent in this approach.

2

Database

This chapter gives an overview of the used climatological, botanical and palaeobotanical data which are the basis for the palaeoclimate reconstructions presented in the following chapters.

2.1. Modern Climate Data

The palaeoclimate reconstructions presented in this thesis need a database of modern climatological data. In this work datasets developed from the CRU, the European Climate Assessment & Dataset project (ECA&D) and the European Centre for Medium-Range Weather Forecasts (ECMWF) are used. This chapter describes the used datasets. The selection of environmental parameters for the palaeoclimate reconstructions depends on the detailed definition of the climate state vector $\vec{C}$ and can be found in chapter 3.1.2 for Lake Prespa and 4.2.3 for Israel.

2.1.1. CRU TS 3.1

The paleoclimate reconstructions of Lake Kinneret and Birkat Ram presented in chapter 4 and 5 are based on previous works of Schölzel (2006), Neumann *et al.* (2007) and Litt *et al.* (2012). They used the CRU Time Series (CRU TS) dataset version 1.1 for the reconstruction of near surface climatic values. The British Atmospheric Data Centre (BADC) provides version 3.1 which includes nine monthly climatological variables: mean temperature, diurnal temperature range, maximum and minimum temperature, precipitation total, vapour pressure, cloud cover, rain day counts and potential evapotranspiration (PET). PET is included for the first time and calculated from a variant of the Penman–Monteith formula (Harris *et al.* 2014).

In more detail the CRU TS 3.1 data are quality checked meteorological station observations for the time slice from 1901 until 2009 interpolated on a regular $0.5^\circ \times 0.5^\circ$ latitude-longitude grid. The calculations presented in the following chapters are applied to

- T_{DJF}, an average of the months December, January and February (Fig. 2.1),
- T_{JJA}, an average of the months June, July and August (Fig. 2.2),
- P_{ANN}, the sum of the precipitation amount of the months from January to December (Fig. 2.3),

- annual climatic water deficit (CWD_{ANN}), described in the following (Fig. 2.4),
- T_{Jan},
- and T_{Jul}.

It is an ongoing discussion which combination of environmental parameters is the best to be reconstructed (s. e. g. Telford, Birks 2011). Schölzel (2006) suggested and discussed the usage of winter and summer temperatures instead of the coldest and warmest months. In paleoclimate reconstruction often January and July are used for these months (s. e. g. Andreev, Klimanov 2000; Cheddadi *et al.* 1998; Gebhardt 2003; Kühl *et al.* 2002; Seppä, Birks 2001; Simonis 2009).

Simonis (2009) showed problems in reconstruction of the annual precipitation amount in Europe and suggested to try another type of proxy data. An alternative is to reconstruct another limiting factor of plants, like for example the climatic water deficit (CWD), which is in general defined as the difference between PET and precipitation (P):

$$CWD := PET - P. \tag{2.1}$$

Crimmins *et al.* (2011) among others analysed the connection between CWD and the distribution of plants and found out that it is a limiting factor. Additionally, CWD or the difference $P - PET$ is often utilized in the computation of climatic drought indices, like for instance the Standardized Precipitation Evapotranspiration Index ($SPEI$) developed by Vicente-Serrano *et al.* (2010).

Figure 2.1 to 2.4 show the geographical distribution of the climatic mean for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} for the time slice from 1961 till 1990 based on CRU TS 3.1 data. This thirty year average is the classical definition of the term climate (s. Schölzel 2006, chap. 3.2, p. 33). Further, the distribution maps for the taxa described in the next section (2.2) are also defined for that time slice (T. Litt pers. comm.). The shown area is the largest available distribution area of one taxon.

The displayed winter temperature in Figure 2.1 ranges from $-50\,^\circ$C in Siberia to $25\,^\circ$C in the deserts in North Africa and India. The direction of the gradient is southwest oriented. The gradient of the summer temperatures in contrast (Figure 2.2) is almost meridional. The range of T_{JJA} covers an interval from $-2.5\,^\circ$C in Spitsbergen to $35\,^\circ$C in the Western Sahara. Clearly visible in the distribution of these 2 m surface temperatures are the orographical effects from the Himalaya and the Alps.

The climatic mean of the annual precipitation amount (P_{ANN}) presented in Figure 2.3 is shown with square root transformation in order to improve the visibility of small values. The regions with the lowest amount of precipitation are the deserts in North Africa, the Arabian Peninsula and the Tibetan highland with values from zero up to $100\,\frac{\text{mm}}{\text{year}}$. The monsoon areas in Western India have the largest amounts of precipitation with values up to $5000\,\frac{\text{mm}}{\text{year}}$.

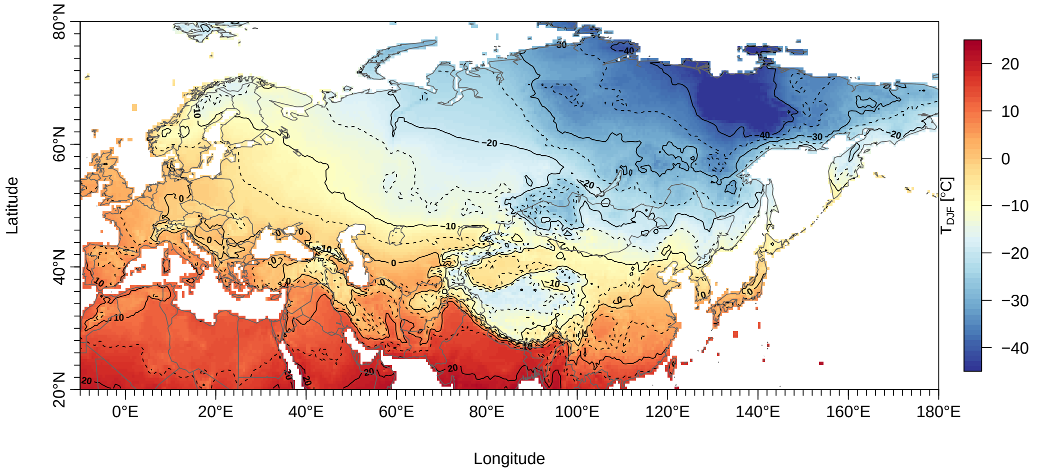


Figure 2.1.: The geographical distribution of the 2 m surface winter temperature (T_{DJF}). In detail this is the field of the climatic mean for the time slice 1961-1990 based on CRU TS 3.1 data.

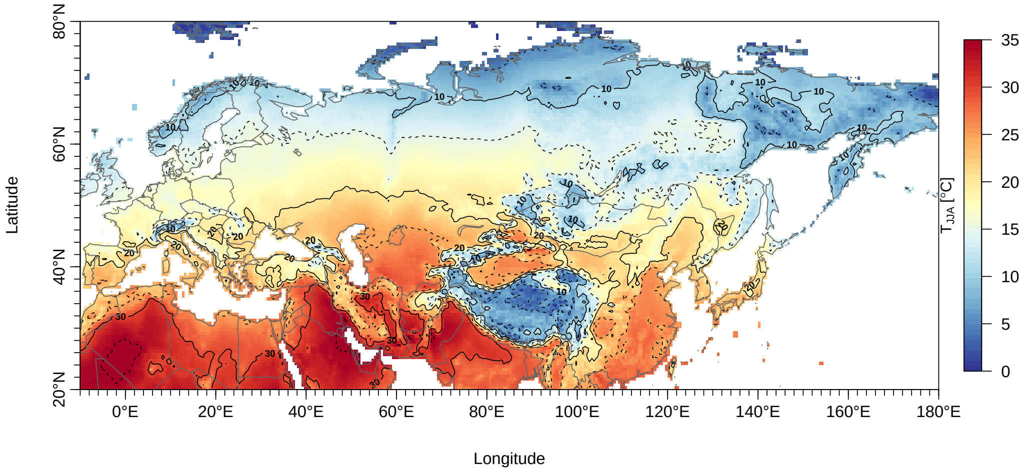


Figure 2.2.: The geographical distribution of the 2 m surface summer temperature (T_{JJA}). In detail this is the field of the climatic mean for the time slice 1961-1990 based on CRU TS 3.1 data.

The direction of the gradient in the precipitation field is not as distinctive as in the temperature fields of T_{DJF} and T_{JJA}.

The geographical distribution of the climatic mean of the annual aggregated CWD_{ANN} is shown in Figure 2.4 and it is, for the same reasons as for precipitation, transformed with the square root. The most humid areas ($CWD < 0$, blue colour scale in Fig. 2.4) are the monsoon areas at the Ganges basin with values lower than $-2000\,\frac{\text{mm}}{\text{year}}$ and also the Alps, the east coasts of the British Isles and the Scandinavian Peninsula, the Islands of Japan and the coast of Indochina are very humid regions with values up to $-1000\,\frac{\text{mm}}{\text{year}}$. Also important to note is that

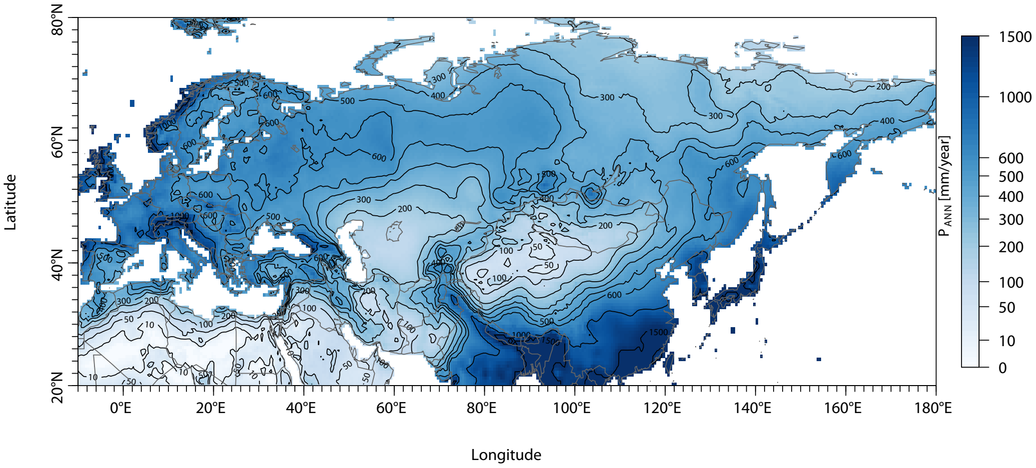


Figure 2.3.: The geographical distribution of the annual precipitation amount (P_{ANN}). In detail this is the field of the climatic mean for the time slice 1961-1990 based on CRU TS 3.1 data and transformed with square root in order to improve the visibility of small values.

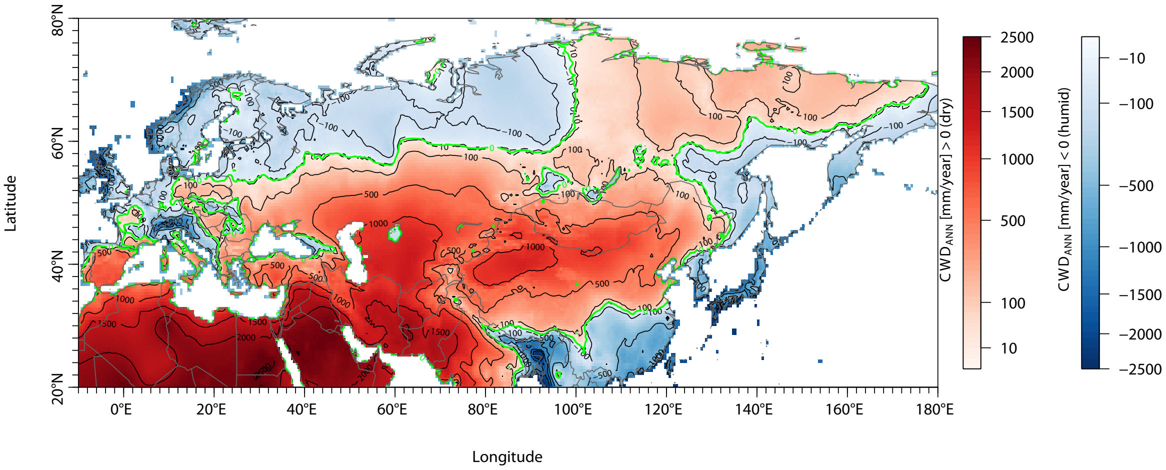
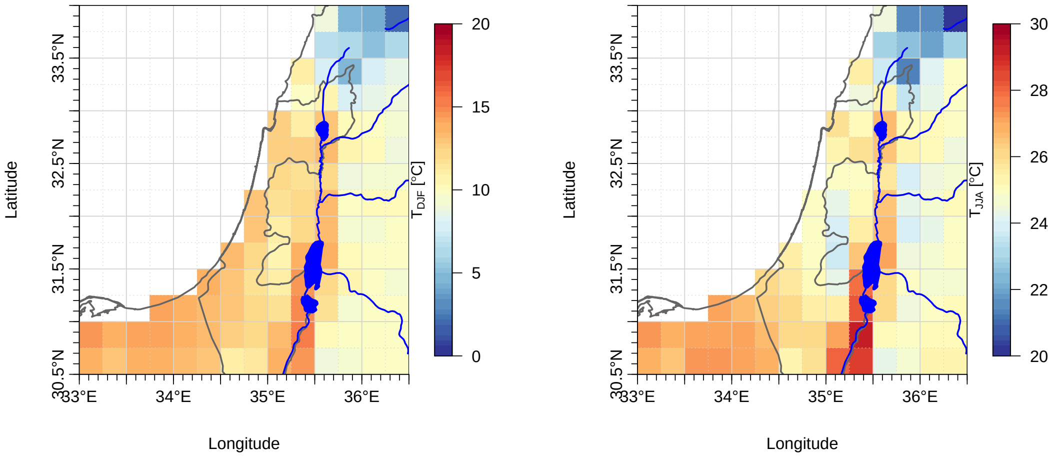


Figure 2.4.: The geographical distribution of the annual climatic water deficit (CWD_{ANN}). In detail this is the field of the climatic mean for the time slice 1961-1990 based on CRU TS 3.1 data and transformed with square root in order to improve the visibility of small values and two separate scales for a clearly visible differentiation between dry (red) and humid (blue) areas. The green line separates the dry and humid areas and defines the zero level.

almost all coasts have CWD_{ANN} values below zero even if they border directly on a desert which are indicated with $CWD_{ANN} \gg 0$ (red colour scale in Fig. 2.4) like for example in North Africa. It could be assumed that this is an effect resulting from the Sea breeze. The Sahara Desert in North Africa and the deserts on the Arabian Peninsula are the driest areas shown in Figure 2.4 with CWD_{ANN} values greater than 2000 $\frac{\text{mm}}{\text{year}}$. Beside these deserts almost the rest of the dry areas with values between green zero level line in Figure 2.4 and up to about 1000 $\frac{\text{mm}}{\text{year}}$ are continental areas. The spatial structure of the gradient is almost the same as the one for precipitation due to the fact that the structure of the geographical distribution of annual potential evapotranspiration (PET_{ANN}) (not shown) and P_{ANN} are very similar.

2.1.2. E-OBS

In this study another climatological dataset is used for the calculation of paleoclimate reconstructions in chapter 4.4 and 6: E-OBS. Geographical E-OBS covers Europe, parts of North Africa, the Anatolian Peninsula and parts of the Levant with a resolution in space of 0.5 or 0.25 degrees in longitude/latitude. The maximum available time interval ranges from 1950 to 2006 with a resolution of one day. As well as the CRU TS 3.1 dataset, it is estimated by interpolation of station data on a regular grid. Haylock *et al.* (2008) and van den Besselaar *et al.* (2011) describe the interpolation process in detail, which differs from the procedure utilized for CRU TS 3.1 data. The ECA&D provides the dataset which comprises the five climatological variables daily mean temperature, daily minimum temperature, daily maximum temperature, daily precipitation amount and daily averaged sea level pressure.



(a) winter temperature T_{DJF} **(b)** summer temperature T_{JJA}

Figure 2.5.: (a) and (b) are the geographical distribution of the climatic mean fields of T_{DJF} and T_{JJA} based on E-OBS data in Israel. In detail these are the field of the climatic mean for the time slice 1961-1990.

The calculation in chapter 4.4 and 6 use a mixture of the CRU TS and E-OBS dataset. Hence the temporal scale is first adopted by averaging on the CRU TS time scale, in detail the monthly mean temperatures and the monthly precipitations sum are calculated. In the next step, T_{DJF}, T_{JJA} and P_{ANN} are generated in the same way as described in the last section for the CRU TS 3.1

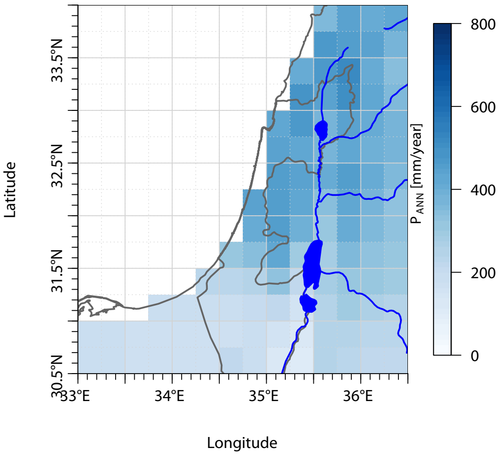


Figure 2.6.: The geographical distribution of the climatic mean fields of the annual precipitation amount P_{ANN} based on E-OBS data in Israel. In detail this is the field of the climatic mean for the time slice 1961-1990.

data. The last step in the preparation of the E-OBS data as input for the reconstruction is the calculation of the climatic mean for 1961 to 1990. The results are shown in Figure 2.5 and 2.6 only for Israel due to the fact that only in this region a higher spatial resolution of climate data is needed. The Jordan Valley is identifiable in all four figures, most clearly in the summer temperature. The temperature distribution shows decreasing values with increasing height.

2.1.3. ERA-Interim

Motivated by the results of Stolzenberger (2011), also reconstructions with air temperatures at the 850 hPa level are carried out. The selected database is the latest reanalysis product of the ECMWF: ERA-Interim with a resolution of also 0.5 degrees in longitude/latitude. The datasets described in the last two sections result from a simple interpolation on grid. The ERA-Interim dataset is a reanalysis. In addition to the station data, this method includes a physical model as a constraint for the interpolation that includes many other meteorological variables. The result are physically consistent fields. Dee *et al.* (2011) describe the process in detail.

Figure 2.7 and 2.8 show the climatic mean of winter and summer temperature at the 850 hPa-level (T_{DJF}^{850hpa}, T_{JJA}^{850hpa}). The underlying time interval is the maximum available time range from 1989 till 2009. Compared to the 2 m surface summer temperature based on CRU TS 3.1 shown in Figure 2.1 and 2.2 the effect of orography is reduced, only an effect of the Himalaya is identifiable. The direction of the temperature gradient matches with the direction of the gradient of the surface temperature. T_{DJF}^{850hpa} ranges from $-28.6\,°C$ in Eastern Siberia to $17.7\,°C$ in almost the complete southern part shown in the map. The T_{JJA}^{850hpa}-interval covers $-2.3\,°C$ in Spitsbergen to $32.3\,°C$ over the Arabian Peninsula.

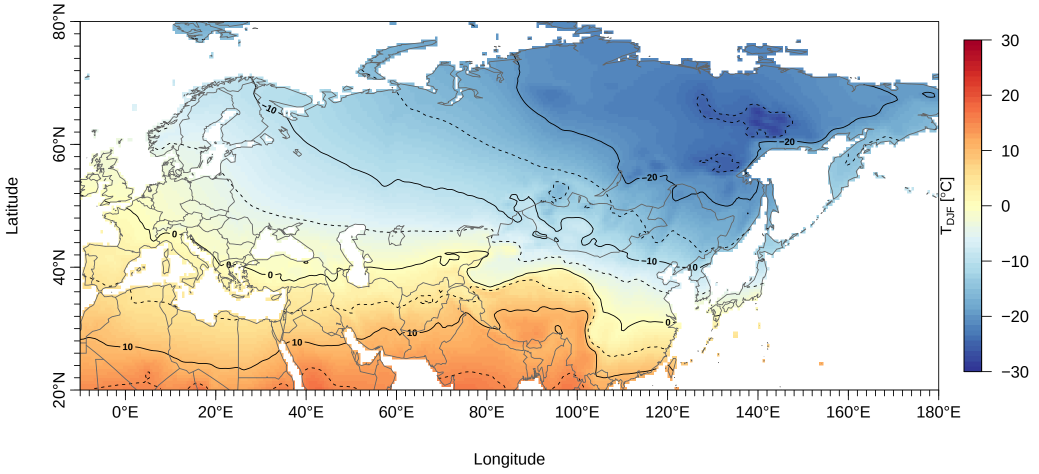


Figure 2.7.: The geographical distribution of the climatic mean fields of the winter temperatures at the 850 hPa-level (T_{DJF}^{850hpa}) based on ERA-Interim data. In detail this is the field of the climatic mean for the time slice 1989-2009.

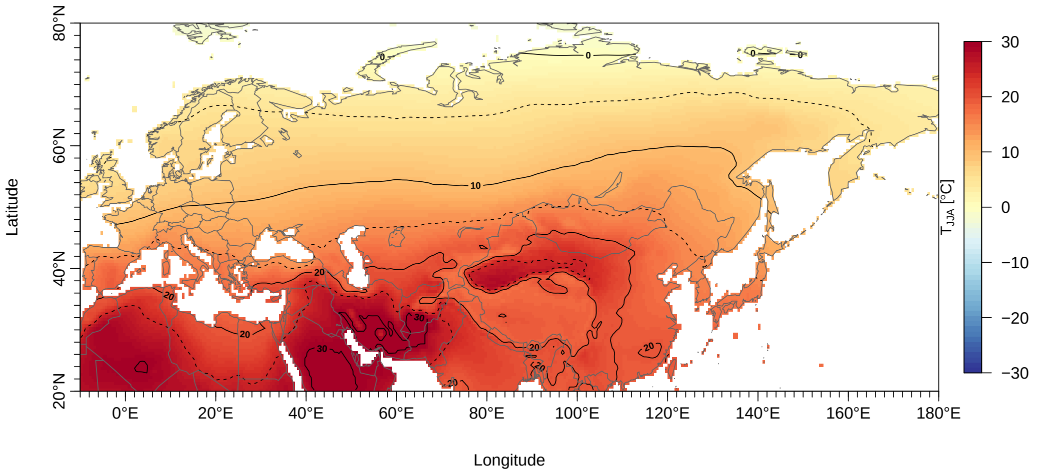


Figure 2.8.: The geographical distribution of the climatic mean fields of the summer temperatures at the 850 hPa-level (T_{JJA}^{850hpa}) based on ERA-Interim data. In detail this is the field of the climatic mean for the time slice 1989-2009.

2.2. Modern Vegetation Data

2.2.1. Distribution-Maps

The three methods BITM (Lake Prespa, chap. 3 and Neumann *et al.* (2007)), BBM (Lake Kinneret and Ein Gedi, chap. 4) and BITBM (Birkat Ram, chap. 5.1) use the geographical distribution of vegetation areas as modern vegetation data. The areas that are used here base on printed maps which are digitized on the same grid as the CRU TS 3.1 data with a software developed by Schölzel *et al.* (2002). This software handles many different types of map projections of the printed maps as well as undefined map projections. Currently about 300 maps are digitized.

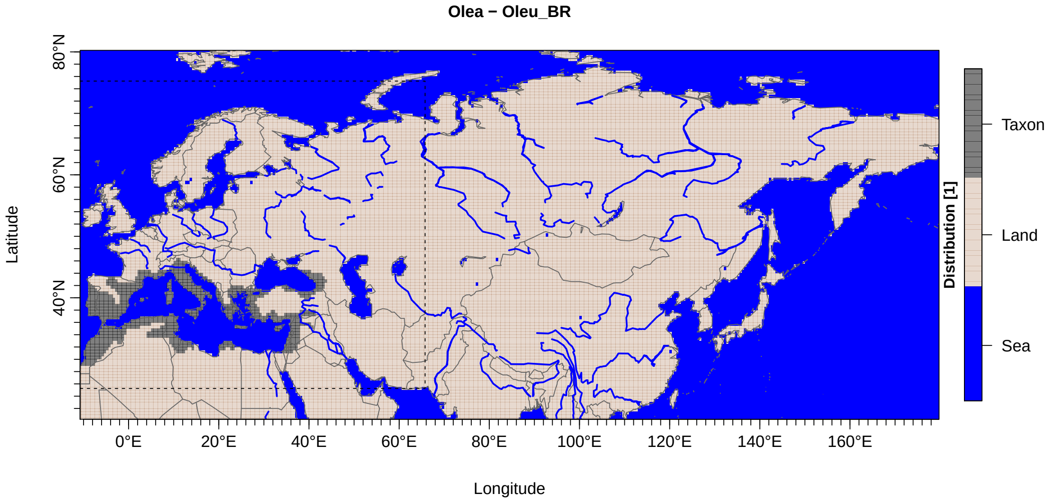


Figure 2.9.: Geographical distribution area of *Olea* after Meusel *et al.* (1974). The dark grey colour indicates the occurrence of the *Olea*, light grey represents Land and no presence of *Olea* and blue is the Sea. The dotted rectangle is the considered maximum area for *Olea*.

BITM as well as BITBM use the geographical extent of a taxon, more specific the distribution of a single species. Figure 2.9 shows for example the map of the geographical distribution of *Olea* after Meusel *et al.* (1974) as used for reconstruction of Lake Prespa described in chapter 3. The information presented on this map is subdivided in three different levels: occurrence of *Olea* in dark grey, no occurrence of *Olea* light grey (land surface) and sea surface in blue.

One important point has to be kept in mind before using this vegetational dataset: Is the taxon distribution only sensitive to climate? In detail every single taxon has to be checked before using for reconstructions (Gebhardt 2003; Kühl 2002; Simonis 2009; Stolzenberger 2011). An additional important point to consider that the resolution of the digitalized maps (0.5° × 0.5°) is too coarse to include microclimate effects. These effects play an important role in grid boxes with strong gradients in the orography, like for example in the alps. Hence the information about the presence or absence of the regarded taxon in these grid boxes is not representative. Gebhardt (2003); Kühl (2002); Simonis (2009); Stolzenberger (2011) just as this study take this into account by filtering out those grid boxes where the difference between the mean and the minimum height AMSL is larger than 400 m. *Digital Terrain Data* (2013) derived from the National Geophysical Data Center (NGDC) is used for filtering.

BBM in contrast to BITM uses the distribution of vegetation zones. Later on this vegetation zones are referred as biomes or biome areas. Litt *et al.* (2012) define and use the biomes

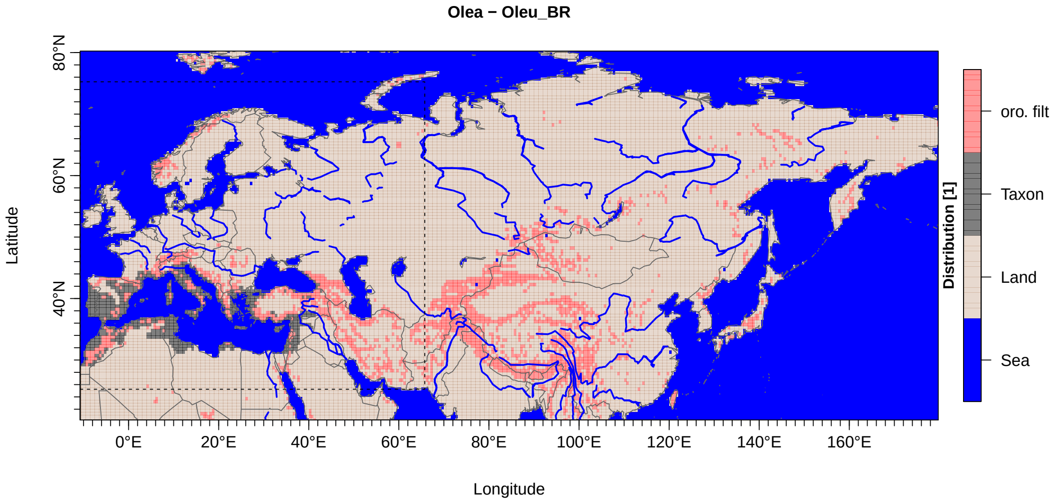
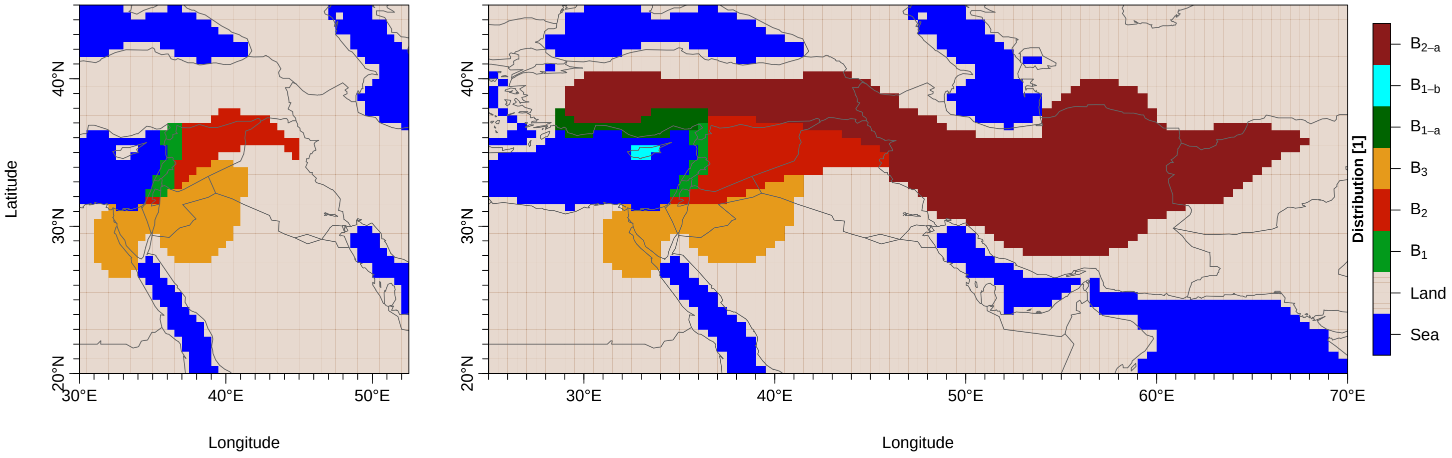


Figure 2.10.: Geographical distribution area of *Olea* after Meusel *et al.* (1974). The colours indicate the same as in Figure 2.9. Additionally all grid boxes with an orography difference between the mean and minimum height AMSL larger than 400 m are marked in red.



(a) Original biome vegetation areas

(b) Extended biome vegetation areas

Figure 2.11.: The maps of the vegetation zones in (a) and (b) in green define the Mediterranean biome B_1, in red define the Mediterranean Irano-Turanian subbiome B_2 and in orange define the Saharo-Arabian biome B_3. (a) is redrawn after and used in Litt *et al.* (2012) and (b) is an extension after Meusel *et al.* (1965). The dark green area B_{1-a} and the cyan area B_{1-b} are the extension of the Mediterranean biome. More precise the extension B_{1-a} is the South Anatolian Mediterranean subbiome. The dark red area B_{2-b} together with the red area B_2 is the complete extend of the Irano-Turanian biome B_2.

as shown in Figure 2.11a. They define three different biomes: The Mediterranean biome B_1 shown in green, the Mediterranean Irano-Turanian subbiome B_2 and the Saharo-Arabian biome B_3. For this study the complete extent of the Irano-Turanian (Fig. 2.11b, dark red area) biome as well as the South Anatolian Mediterranean subbiome (Fig. 2.11b, dark green) has been digitized. This is done for a better consideration of the continental climate (T. Litt pers. comm.) and the Meditrranean climate in the reconstructions. The definition of the biome areas B_1, B_2 and B_3 differs due to a suspected error in the digitalisation. Thereby the most changes are apparent for the Irano-Turanian biome area B_2 and the Saharo-Arabian biome B_3 area: a slightly shift south of B_2 and modification of the common border. The area of the Mediterranean biome B_1 is only slightly reduced on the eastern boundary. A detailed discussion of the resulting effect can be found in chapter 4.2.3.

2.2.2. Modern Pollen Spectra

The PRM as applied in chapter 4.4 requires another kind of modern vegetation data: pollen spectra at specific locations in the neighbourhood of the coring site for which the palaeoclimate reconstruction is estimated. For this purpose the spectra which are described in Table 6.1 in Weinstein (1979) are used. Both commonly used databases, *PANGAEA* (2012) and *European Pollen Database (EPD)* (2012), and also no other database, as known by the author, have entries with modern pollen spectra in the considered region. Therefore the tables in Weinstein (1979) are selected. More precisely the spectra presented in these tables are annual averages of airborne pollen spectra (s. Tab. B.7 and B.8). Weinstein (1979) defines the localities only by names. For the connection to the climate dataset geographical coordinates are required.

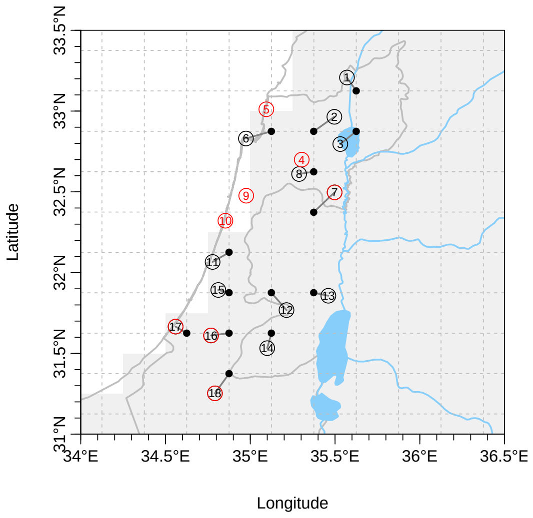


Figure 2.12.: Localities of modern pollen spectra in Israel based on Weinstein (1979). Every number refers to the site number and the geographical coordinate fixed with *Google Maps* (2012) (Table B.6). Sites with black numbers are used and sites with red numbers are rejected. Black number and red circle indicates localities marked with a comment in Weinstein (1979). The vertical and horizontal dotted lines mark the centre of the grid points of the E-OBS dataset. The grey shaded background indicates the availability of climate data.

Hence, *Google Maps* (2012) is applied to get the coordinates (red or black number in Fig. 2.12 s. Tab. B.6 for the data). In order to connect these spectra with the E-OBS-climate-dataset every locality with available pollen spectra is assigned to one E-OBS-grid point. This is done by calculating the distance between one E-OBS-grid point (black point in Fig. 2.12) and the one or two nearest localities. Figure 2.12 shows the resulting 13 localities with modern pollen spectra and modern climate information with black numbers and a line for visualizing the assignment to the E-OBS-grid. In the original data table in Weinstein (1979) all localities are annual averages of airborne pollen spectra except for four which are marked with a black number and a red circle in Figure 2.12.

2.3. Sediment Cores

The essential part for palaeoclimate reconstructions of the past are proxy data of the past. The pollen counts are obtained from pollen diagrams which base on sediment cores. Figure 2.13 shows such a pollen diagram for Birkat Ram. Sediment cores are drilled in different kind of bodies of water. These bodies of water collect all pollen which reach the water surface area or are transported in by any kind of inflow. Because of the last mentioned point the selection of the used body of water by the geoscientists and/or palaeontologist is essential. Over the time the pollen and also macrofossils are deposited on the ground of the body of water and are then recoverable by drilling a core in the sediments.

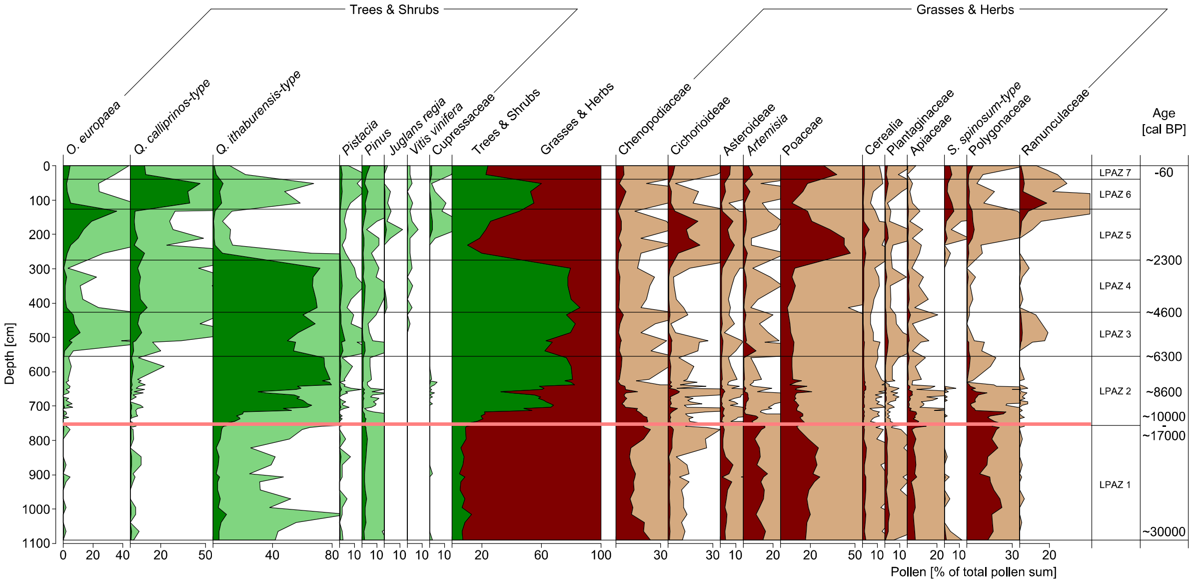


Figure 2.13.: Pollen diagram for Birkat Ram generated by Schiebel (2013) showing most relevant taxa. For more information see Schiebel (2013, chap. 5.2, p. 47)

The reconstructions presented in the next chapters base on sediment cores, or more precise on composite cores that were obtained in different drilling campaigns. The term "*composite core*" reflects the fact that it is not possible to obtain a sediment core of a length of about 10 m as a whole. A core of this length is maintained by drilling two holes in a very short distance. This is e. g. for Lake Kinneret and Birkat Ram a distance of 2 m (Schiebel 2013, chap. 4.2 and 4.3). In the next step the received sediments of every hole are cutted in such portions of about 2 m that the point of cutting in both core is not the same, which means that the individual segments have a short overlapping. By measuring the magnetic susceptibility of each segment and then correlating the values it is possible to compose a complete profile of unlimited length.

The available data that is based on these cores and the coring sites are shortly described in this section. In general, for every sediment core used in this study an age-to-depth model based on calibrated radiocarbon data is available. Age-to-depth models assign every depth to a calibrated age usually measured in calibrated years before present (y. cal BP)

2.3.1. Balkan peninsula - Lake Prespa

The sediment core Co1215 of Lake Megali Prespa was retrieved at $40^\circ 57' 50'' N$, $20^\circ 58' 41'' E$ in November 2009 as part of the CRC project B2. It has a length of 320 cm. Lake Megali Prespa from now on Lake Prespa, is situated on the Balkan Peninsula in south-west Macedonia.

Panagiotopoulos *et al.* (2013) performed the pollen analysis and characterises the Lake and the environment shown in Figure 2.14: "*The transboundary Prespa catchment* 1300 km^2 *comprises two lakes: Megali and Mikri Prespa, with an approximate surface area of* 253.6 km^2 *and* 47.4 km^2 *respectively*. . .*Lake Prespa is situated at an altitude of* 849 m *AMSL and is surrounded by mountains with the highest peak at* 2601 m *AMSL to the east, and several other peaks around or above* 2000 m *AMSL to the west and the north*. . . .*It has a mean water depth of* 14 m."

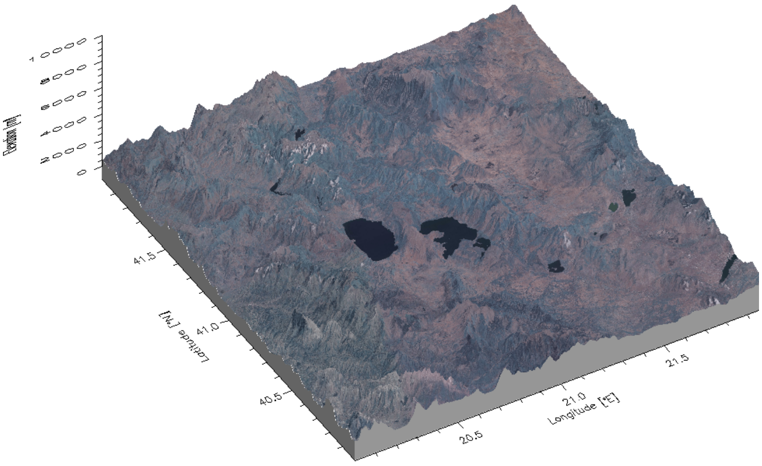


Figure 2.14.: The Lake Prespa area: A projected satellite image (*CRC806-Database Satellite WMS*) on the NASA digital elevation model. The two lakes in the centre of the satellite image are Lake Megali Prespa and Lake Mikri Prespa.

60 samples were taken out of 318 cm long composite core for the palynological analysis of the pollen with an interval ranging from 2 – 8 cm. From the 82 identified and counted taxa 26 are used for the palaeoclimate reconstruction presented in chapter 3. For these 26 it is assumed that they are not anthropogenic influenced and only sensitive to climatic conditions.

Aufgebauer *et al.* (2012) established and discussed the age-to-depth model that is used in this study and concluded that the core covers a time period of about 17 000 y. cal BP.

2.3.2. Southern Levant - Jordan Valley

The sediment cores described in this section are all gained at locations in the southern Levant or, to be more precise, at locations in the Jordan Valley or in the direct neighbourhood. All palaeoclimate reconstructions in this region which are presented in this study base on cores analysed at or in cooperation with the Steinmann Institute for Geology, Mineralogy and Palaeontology of the University of Bonn. The two latest sediment cores, namely Lake Kinneret and Birkat Ram, were drilled as part of the CRC project B3 in March 2010 and analysed by Schiebel (2013). Figure 2.15 shows the Jordan Valley, the Dead Sea in the South and Lake Kinneret in the North.

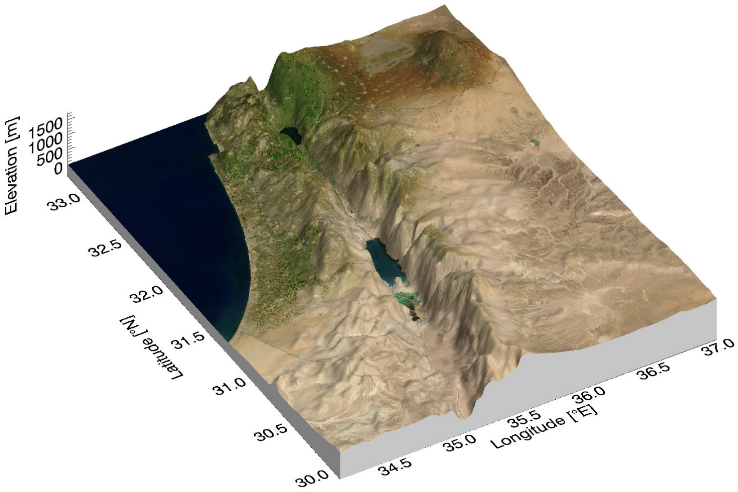


Figure 2.15.: Southern Levant Area: A projected satellite image on a digital elevation model. Image courtesy of Schölzel (2006)

2.3.2.1. Lake Kinneret

Lake Kinneret, in the literature also called Sea of Galilee or Lake Tiberias, has a surface area of 166 km^2 and a catchment area of 2760 km^2 (Schiebel 2013). This lake has a maximum depth of 43 m and the lake surface lies 211 m below mean sea level (BMSL). Schiebel (2013) also described the origin of the water: two-thirds derive form the Jordan River and one-third originates from other streams, seasonal floods, direct rainfall and subaqueous springs. The core was drilled near the centre of the Lake at $32°49'13.8''N$ and $35°35'19.7''E$ at a water depth of 38.8 m. The composite core is 17.8 m long and sampled in 25 cm intervals for the palynological analysis. The results are 73 samples and 63 identified and counted taxa. The palaeoclimate reconstruction in chapter 4 uses 42 of them. Schiebel (2013) also established two different age-to-depth models which differ in the assumptions. The differences between the age-to-depth models turn out in the lower part of the core below 10 m. Depending on the selected model the record spans an time range between 8 600 y. cal BP and 9 200 y. cal BP. In this study the age model with increasing interpolated reservoir correction (V. Schiebel pers. comm. and Schiebel 2013) is used.

2.3.2.2. Birkat-Ram

For the maar lake in the northern Golan Heights two composite cores are available:

1.) The first one was drilled in 1999 and palynological analysed by Schwab *et al.* (2004). Schwab *et al.* (2004) also established an age-to-depth model for the 543 cm long composite core which covers about 6 500 y. cal BP from the present to the past. The 140 samples on which the pollen counts are based are gained in an interval ranging from $1-8$ cm. For this sediment core the palaeoclimate reconstructions for T_{DJF}, T_{JJA} and P_{ANN} are available (s. Neumann *et al.* 2007; Schölzel 2006). Therefore nine of 153 identified taxa are used.

2.) The second one was drilled, as mentioned above, within the CRC in March 2010 at $33°13'54.3''N$, $35°46'1.4''E$ at water depth 14.5 m. The composite core is 10.96 m long. Schiebel (2013) determined and discussed the age-to-depth model in connection with the first drilled core, describe before, with the result that the covered time period ranges from 0 y. cal BP to 30 000 y. cal BP. Between 703 cm and 746 cm (10 000 and $\approx$ 17 000 y. cal BP) Schiebel (2013) detected a lower sedimentation rate and assumed a desiccation of Birkat Ram. Schiebel (2013) analysed all 44 available sample layers and identified 59 taxa of which six are used in chapter 5.

Birkat Ram lies at 940 m AMSL and has an average surface area of 0.45 km^2 with a water depth that depends on the season and ranges between 6 m and 12 m (Schiebel 2013). The catchment area is quite small, about 1.5 km^2 (Schwab *et al.* 2004) and the main water source is precipitation.

2.3.2.3. Ein Gedi

For the Dead Sea the analysed core was drilled 1997 at Ein Gedi spa at 415 m BMSL ($31°30'0''N$, $35°24'0''E$). The sediment core is extensively analysed and discussed in Migowski (2001), Migowski *et al.* (2004), Migowski *et al.* (2006), Litt *et al.* (2012). The composite sediment core is 21 m long and the pollen analysis is carried out on 58 sample layers (Litt *et al.* 2012). 26 taxa are selected for the palaeoclimate reconstruction developed in Schölzel (2006) and enhanced in Litt *et al.* (2012). The Ein Gedi core covers the last 10 000 y. cal BP (Migowski *et al.* 2004, 2006).

In contrast to the already described sediment cores the Ein Gedi core has recorded the history of the Dead Sea basin which is characterized by a large catchment area (Litt *et al.* 2012). It includes very distinctive vegetation zones, which are highly sensitive to climate changes (Litt *et al.* 2012). This point is important for the selection of the used method for palaeoclimate reconstruction (s. chap. 4)

3

Lake Prespa

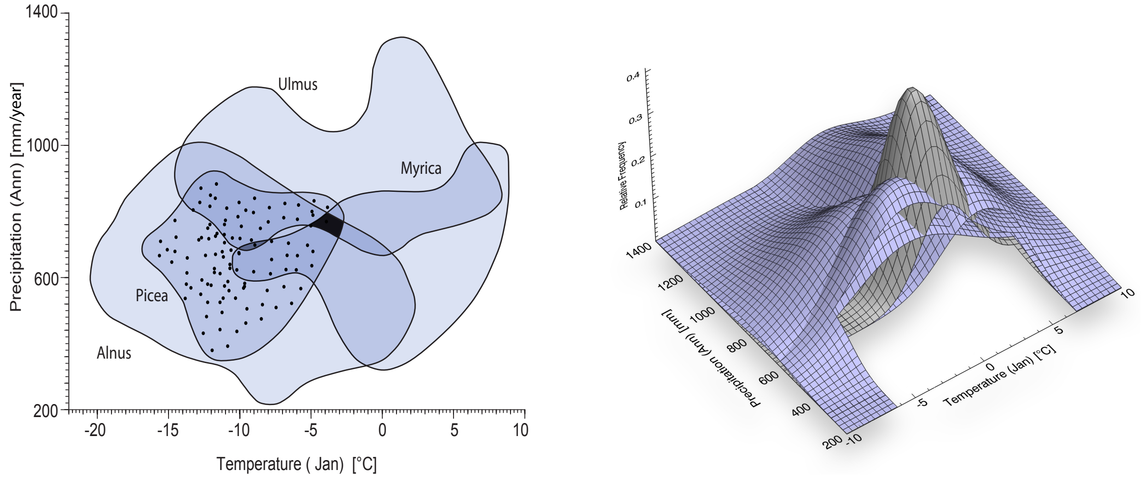
This chapter presents the result and the mathematical tools that are used for the palaeoclimate reconstruction which bases on the sediment core data of Lake Prespa. The sediment core and the location are described in chapter 2.3.1. The basic mathematical principles for this purpose are introduced in chapter 1.2 and are extended in the first section of this chapter to the Bayesian Indicator Taxa Model (BITM) which is applied here. The second and third section of this chapter presents the palaeoclimate reconstruction for Lake Prespa. In more detail each contributing part of the BITM is explained with respect to Lake Prespa. The main interest is the posterior probability density function (pdf) and therefore it is presented in a separate section.

3.1. The Bayesian Indicator Taxa Model (BITM)

3.1.1. The development of the BITM

The BITM is an approach which was developed within the ”*Bonner Group*”. The latest version, which is also used here, was developed and published by Schölzel (2006) for the sediment core of Birkat Ram. It is a result of many years of research. Ohlwein, Wahl (2012) described this development extensively. The milestones are summarized here.

Iversen (1944) and Hintikka (1963) published a graphical approach: bioclimatic ranges for different kind of taxa. The idea for the BITM bases on the mutual climatic range (MCR) method from Grichuk (1969). MCR is an refinement and improvement of the methods from Iversen (1944) and Hintikka (1963). The MCR method combines the information of climate maps (s. Fig. 2.1 to 2.8) and geographical distribution maps of taxa (s. e. g. Fig. 2.9 for *Olea*). Figure 3.1a shows this principle, in detail for taxon *Picea* and the climate variables January temperature (T_{Jan}), defined as the coldest month, and annual precipitation amount (P_{ANN}). In detail each black point for *Picea* in Figure 3.1a represents one point of the geographical distribution where the information of presence is available given the plotted coordinate combination of T_{Jan} and P_{ANN} of the point. Due to the fact that MCR is a graphical approach a line which encloses all these points is drawn. This enclosed line defines the climate phase space for each taxon: Inside of this defined area the probability for a taxon to occur for the given combination of climate variables is equal to one and outside of this area it is zero. This process



(a) MCR method

(b) BITM

Figure 3.1.: Principles of the MCR method (a) and the BITM (b):
(a) shows the palaeoclimate reconstruction principle of the MCR method. The black intersection area of all taxa indicates the reconstructed climate. The figure is redrawn after Schölzel (2006) and Grichuk (1969).
(b) shows the palaeoclimate reconstruction principle of the BITM. The blue bivariate pdfs in (b) replace the blue areas in (a). The reconstructed climate is indicated by the grey bivariate pdf. The cut in Figure (b) is only for better presentation of individual normal distributions. Images courtesy of Ohlwein (2012).

of defining an area for the climatic range of a taxon is repeated for all taxa present in one pollen sample layer of a sediment core. In Figure 3.1a this is shown for *Ulmus* and *Myrica*. The intersection of all areas is then interpreted as the reconstructed climate.

The MCR approach has two main problems as pointed out by Ohlwein, Wahl (2012). First, no uncertainty estimation is possible due to the sharp boundaries. As described in the last paragraph the area in the climate phase space for one taxon describes only the two discrete probability values zero and one. Additionally the shape is graphical defined which allows not traceable and arbitrary shapes like for example for *Ulmus* in Figure 3.1a. This problem is called overfitting. Secondly, if many taxa are included it is possible that more than one intersection area exists, like for example the dark grey area in Figure 3.1a. Within the MCR approach there is no statement possible to exclude or interpret this result. Pross *et al.* (2000) and Klotz *et al.* (2004) enhance the MCR concept to the so called probability mutual climatic sphere (PCS) to solve this problem.

The "*Bonner Group*" enhance the MCR method to the so called probabilistic indicator taxa approach or pdf method (Kühl *et al.* (2002)). Also Gebhardt (2003)/Gebhardt *et al.* (2008) used this method. In detail this method interprets the areas above described for one taxon as pdf or more precisely as conditional pdf of climate given the appearance of the taxon. In Figure 3.1b these are the blue bivariate pdfs. The reconstruction of the climate is estimated by multiplication of the taxon individual pdfs and is indicated by the grey bivariate pdf in

Figure 3.1b. The cut in Figure 3.1b is only for better presentation of individual normal distributions. How these pdfs are estimated and the mathematics are discussed in the following section.

The last step in the development was the integration of the pdf method from Kühl *et al.* (2002) in the context of Bayesian Hierarchical Model (BHM) by Schölzel (2006) as presented in the next section. The BITM is also used in Neumann *et al.* (2007), Simonis (2009)/Simonis *et al.* (2012), Stolzenberger (2011).

3.1.2. The BITM - Introduction

The BITM as developed by Schölzel (2006) and applied here uses a climate state vector defined as

$$\begin{aligned} \vec{C} &:= (T_{Jan}, T_{Jul}, P_{ANN})^T && \text{or} && (3.1a) \\ &:= (T_{Jan}, T_{Jul}, CWD_{ANN})^T && \text{or} && (3.1b) \\ &:= \left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, P_{ANN}\right)^T && \text{or} && (3.1c) \\ &:= \left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, CWD_{ANN}\right)^T && \text{all with the realisations} \quad \vec{c} \in \mathbb{R}^3. && (3.1d) \end{aligned}$$

The modern climate state data, are the climate datasets already presented in chapter 2.1 and the past climate state is the unknown property in this study. In equation 3.1 and later on ^{T} reflects the fact that from now on if a vector is defined via its components it is always regarded as a column vector.

The proxy used in the BITM are, as already mentioned, pollen data. On the one hand the modern part of the proxy data are the distribution maps for the taxa as presented in chapter 2.2.1 with the geographical information of presence and absence. On the other hand these are the past proxy data, the data which has recorded the past: Here these are the pollen counts of N_k different taxa in each pollen sample layer of the sediment core, which are presented in chapter 2.3. How these counts are transferred into the information of presence and absence for each taxon is described later in this section. The index of N indicates that the corresponding mathematical counting index is in general k for taxa. From these considerations follows the definition of the corresponding random vector analogously to Schölzel (2006):

$$\vec{T} := (T_1, \ldots, T_{N_k})^T \quad \text{with the realisations} \quad \vec{t} \in \{0,1\}^{N_k}. \tag{3.2}$$

Here $t_k = 0$ stands for absence and $t_k = 1$ for presence of taxon k as mentioned in chapter 3.1.1.

The corresponding joint probability for the considered random variables is

$$\mathbb{P}_{\vec{C},\vec{T}}\left(\vec{c}, \vec{t}\right). \tag{3.3}$$

Compared to equation 1.2 (chap. 1.2, p. 4) the random variable $\vec{P}$ is replaced by $\vec{T}$. The application of the Bayes theorem leads analogously to chapter 1.2 and equation 1.4 to

$$\mathbb{P}_{\vec{C}|\vec{T}}\left(\vec{c}\,|t_1,\dots,t_{N_k}\right)=\frac{\mathbb{P}_{\vec{T}|\vec{C}}\left(t_1,\dots,t_{N_k}|\,\vec{c}\right)\cdot\pi_{\vec{C}}\left(\vec{c}\right)}{m_{\vec{T}}\left(t_1,\dots,t_{N_k}\right)}. \tag{3.4}$$

One assumption inherent to the MCR method is that connection between the individual taxa and the climate is independent from the connection to the other taxa. This is expressed by a ”*pairwise conditional independence of all taxa given a certain climate state*” (Schölzel (2006)). Hence the so called likelihood splits as follows into

$$\mathbb{P}_{T_1,\dots,T_{N_k}|\vec{C}}\left(t_1,\dots,t_{N_k}|\,\vec{c}\right)=\prod_{k=1}^{N_k}\mathbb{P}_{T_k|\vec{C}}\left(t_k|\,\vec{c}\right). \tag{3.5}$$

The $\mathbb{P}_{T_k|\vec{C}}\left(t_k|\,\vec{c}\right)$ are the individual taxon specific part of the likelihood, and are the so called transfer functions for each taxon. As mentioned in Ohlwein, Wahl (2012) ”*these conditional probabilities can be estimated directly e. g. via generalized linear model (GLM) with polynomials of continuous covariates or indirectly*”. The estimation via GLM is described in chapter 3.2.2.

The mentioned indirect method applies again the Bayes theorem on $\mathbb{P}_{T_k|\vec{C}}\left(t_k|\,\vec{c}\right)$ and results together with equation 3.4 and 3.5 in

$$\mathbb{P}_{\vec{C}|\vec{T}}\left(\vec{c}\,|t_1,\dots,t_{N_k}\right)=\frac{\pi_{\vec{C}}\left(\vec{c}\right)}{m_{\vec{T}}\left(t_1,\dots,t_{N_k}\right)}\cdot\prod_{k=1}^{N_k}\frac{\mathbb{P}_{\vec{C}|T_k}\left(\vec{c}\,|t_k\right)\cdot\pi_{T_k}\left(t_k\right)}{m_{\vec{C}}\left(\vec{c}\right)}. \tag{3.6}$$

Equation 3.6 is essentially the BITM used by Schölzel (2006) but it includes presence and absence of taxa in the likelihood. The original concept of the MCR method, described in the last section, and the deduced pdf method, the precursor method of the BITM, use only the presence information for estimating the transfer functions. Schölzel (2006) solves this problem by the assumption that the absence events $t_k = 0$ are ”*conditionally independent of the climate state vector given further random variables like e.g. unknown soil properties, denoted by a general condition*” G (eq. 3.7). Finally Schölzel (2006) neglects all terms in the resulting equation which don’t depend on the climate because they include no necessary information for the palaeoclimate reconstruction:

$$\mathbb{P}_{\vec{C}|T_{Ind(1),\dots,Ind(N_{\tilde{k}})}}\left(\vec{c}\,|1,\dots,1,G\right)\propto\pi_{\vec{C}}\left(\vec{c}\right)\cdot\prod_{\tilde{k}=1}^{N_{\tilde{k}}}\frac{\mathbb{P}_{\vec{C}|T_{Ind(\tilde{k})}}\left(\vec{c}\,|1,G\right)}{m_{\vec{C}}\left(\vec{c}\right)} \tag{3.7}$$

The new index for T in equation 3.7 reflects that

$$\forall t_{Ind(\tilde{k})}=1\quad\text{with}\quad Ind\left(\tilde{k}\right)\in\{1,\dots,N_k\}\quad\wedge\quad\tilde{k}=1,\dots,N_{\tilde{k}} \tag{3.8}$$

This neglecting step is also possible for the combination of equation 3.4 and 3.5 without the application of the Bayes theorem on the likelihood and results in

$$\mathbb{P}_{\vec{C}|\vec{T}}\left(\vec{c}\,|t_1,\ldots,t_{N_k}\right) \propto \pi_{\vec{C}}\left(\vec{c}\right)\cdot\prod_{k=1}^{N_k}\mathbb{P}_{T_k|\vec{C}}\left(t_k|\vec{c}\right). \tag{3.9}$$

Until now probabilities are considered in the equations of this chapter, but it is possible to switch to pdf w. l. o. g. which leads to the equations 3.10. Additionally $\propto$ is replaced by $=$ and equation 1.5, the main assumption for palaeoclimate reconstructions as presented and described in chapter 1.2, is used. Equation 3.1 to 3.9 do not distinguish between the past and the recent random variables and realisations due to the fact that the considerations are the same. Equation 3.10a is the BITM-equation used in Schölzel (2006), Neumann *et al.* (2007), Kühl *et al.* (2007) and Litt *et al.* (2009). Stolzenberger (2011) used only the likelihood part of equation 3.10b for the palaeoclimate reconstruction. Summarized there are two different but closely related possibilities to estimate the posterior pdf

$$f_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0|1_0,\ldots,1_0,G_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right)\cdot\prod_{\tilde{k}=1}^{N_{\tilde{k}}}\frac{f_{\vec{C}|T_{Ind(\tilde{k})}}\left(\vec{c}_0|1_0,G_0\right)}{f_{\vec{C}}\left(\vec{c}_0\right)}, \tag{3.10a}$$

$$f_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0|\vec{t}_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right)\cdot\frac{\prod_{k=1}^{N_k}\mathbb{P}_{T_k|\vec{C}}\left(t_{k,0}|\vec{c}_0\right)}{\sum_{k=1}^{N_k}\prod_{k=1}^{N_k}\mathbb{P}_{T_k|\vec{C}}\left(t_{k,0}|\vec{c}_0\right)}. \tag{3.10b}$$

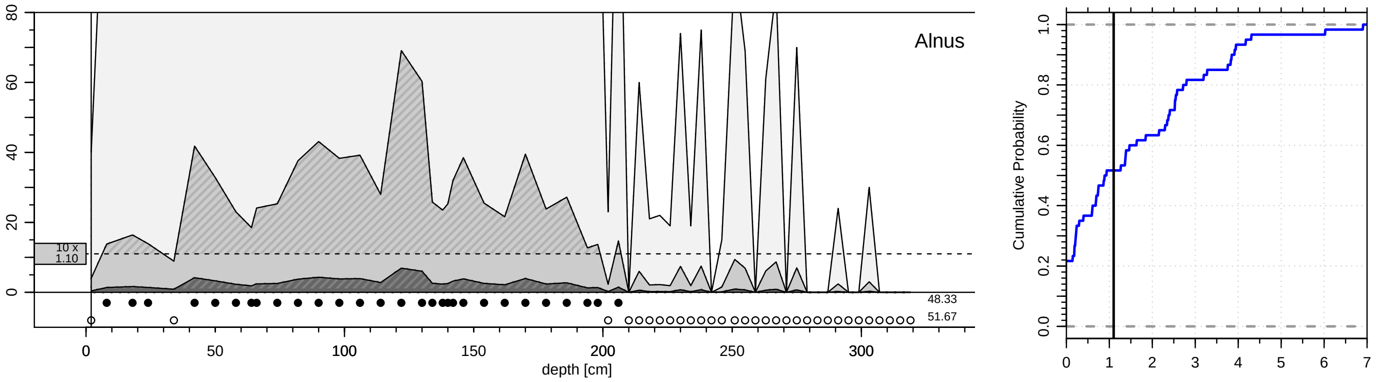
The denominator in equation 3.10b is a direct consequence of this pdf approach. It is the normalisation constant which ensures that the integral over the numerator is equal to one, the essential requirement for a pdf.

3.2. The BITM in Detail for Lake Prespa

3.2.1. Determination of the Indicator Taxa

In BITM the word indicator represents the fact that presence of a certain taxon in one pollen sample layer $\imath$ of the sediment core represents a certain climate state. For this aim the aforementioned transformation of pollen counts to presence and absence of each climatic sensitive taxon k_0 of one sample layer $\imath$ needs to be estimated. Climatic sensitive implies that only taxa which are only influenced by climate are selected for the palaeoclimate reconstruction . In this study the same approach as in Neumann *et al.* (2007) and Litt *et al.* (2009) by exceeding a threshold ϑ_{k_0} is used.

The ϑ_{k_0} estimation bases on finding a gap in the empirical cumulative distribution function (ECDF) for all samples of the regarded taxon of the sediment core. Figure 3.2 demonstrates this approach for Lake Prespa and the taxon *Alnus*. Figure 3.2 shows the depth profile of the relative abundance $\vec{\widetilde{\omega}}_{k_0}$ and the corresponding ECDF (Fig. 3.2b). One gap in the ECDF could be examined at 1.1 % relative abundance. As described in Neumann *et al.* (2007) all values below this threshold are "*interpreted as background noise*" and hence interpreted as absence of the taxon. The figures for the remaining taxa can be found in appendix A.1.2 on page 132 and all defined threshold values are summarized in Table A.1 (p. 131).



(a) depth profile

(b) ECDF

Figure 3.2.: *Alnus*: (a) Depth profile of the relative abundance $\widetilde{\omega}_{\iota\, k_0}$ in each pollen sample ι in dark grey. The light grey area is the ten times and the very light grey the 100 times enlarged profile. (b) is the corresponding ECDF. The threshold value ϑ_{k_0} is visualized in (a) by the different shaded areas and in (b) by the vertical bold black line. The presence/absence information is visualized in (a) with black dots (=presence) and circles (=absence).

Figure 3.2b also demonstrates the disadvantage of this determination: there are more gaps detectable than the selected one. The advantage is that this approach also offers the possibility to include the background knowledge of the palaeontologists by selection of the best ϑ_{k_0}. Additionally Litt *et al.* (2009) pointed out that "*choosing different thresholds for a certain taxon within a span of several percent leads to similar results regarding colonization and population expansion*". And Neumann *et al.* (2007) pointed out "*that every taxon that is excluded due to a high threshold value does not influence the reconstruction in an erroneous way. The model is robust enough to react to a lack of taxon information with an increase of uncertainty, not with misleading reconstructions.*"

In cooperation with Dr. Kostas Panagiotopoulos M. Sc. the palaeontologist who analysed the sediment core of Lake Prespa and counted the taxa and Dipl. Bio. PD Dr. Norbert Kühl, $N_{k_0} = 26$ taxa are selected for the palaeoclimate reconstruction. All resulting relative abundances $\widetilde{\omega}_{\iota\, k_0}$

of the $N_\iota = 60$ sample layers are summarized in one matrix

$$\mathbf{\Omega}^{N_\iota \times N_{k_0}} := \begin{pmatrix} \widetilde{\omega}_{1\,1} & \cdots & \widetilde{\omega}_{1\,k_0} & \cdots & \widetilde{\omega}_{1\,N_{k_0}} \\ \vdots & \ddots & & & \vdots \\ \widetilde{\omega}_{\iota\,1} & & \widetilde{\omega}_{\iota\,k_0} & & \widetilde{\omega}_{\iota\,N_{k_0}} \\ \vdots & & & \ddots & \vdots \\ \widetilde{\omega}_{N_\iota\,1} & \cdots & \widetilde{\omega}_{N_\iota\,k_0} & \cdots & \widetilde{\omega}_{N_\iota\,N_{k_0}} \end{pmatrix} \tag{3.11}$$

and then transferred with the above described approach into presence and absence result in the indicator matrix

$$\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}} := \begin{pmatrix} 0_{1\,1} & \cdots & 1_{1\,k_0} & \cdots & 0_{1\,N_{k_0}} \\ \vdots & \ddots & & & \vdots \\ 1_{\iota\,1} & & 1_{\iota\,k_0} & & 0_{\iota\,N_{k_0}} \\ \vdots & & & \ddots & \vdots \\ 0_{N_\iota\,1} & \cdots & 1_{N_\iota\,k_0} & \cdots & 0_{N_\iota\,N_{k_0}} \end{pmatrix} = \begin{pmatrix} t_{1\,1} & \cdots & t_{1\,k_0} & \cdots & t_{1\,N_{k_0}} \\ \vdots & \ddots & & & \vdots \\ t_{\iota\,1} & & t_{\iota\,k_0} & & t_{\iota\,N_{k_0}} \\ \vdots & & & \ddots & \vdots \\ t_{N_\iota\,1} & \cdots & t_{N_\iota\,k_0} & \cdots & t_{N_\iota\,N_{k_0}} \end{pmatrix} \tag{3.12}$$

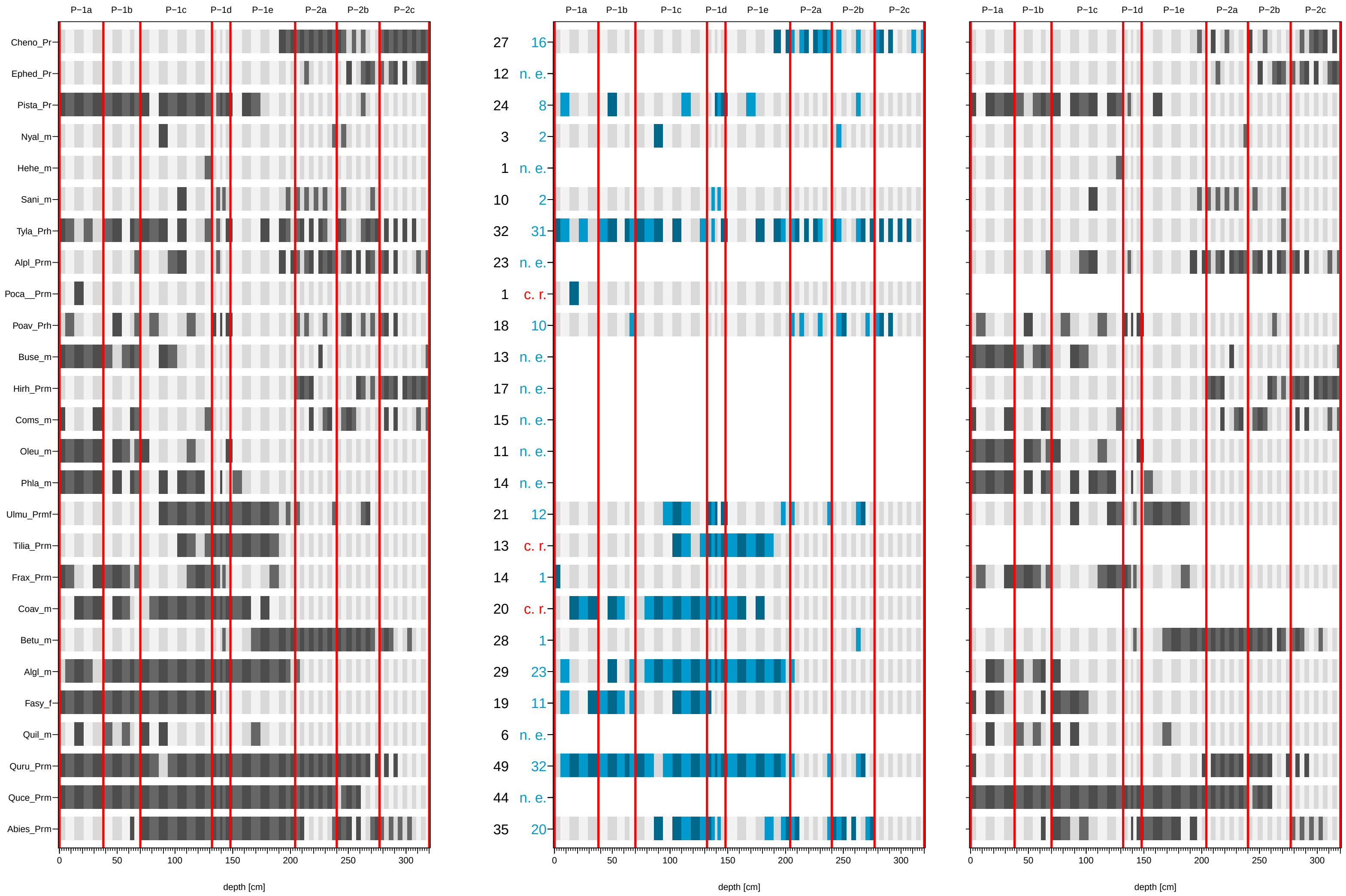
which is shown in Figure 3.3a. This matrix contains only zeros (light grey colours) and ones (dark grey colours) and is the summarized past proxy realisation $t_{\iota\,k_0}$ and the database for the palaeoclimate reconstruction of Lake Prespa.

One problem could arise in the BITM approach if many indicator taxa are included: A reduction of uncertainty due to the multiplication of taxon transfer functions (eq. 3.10) of taxa which occur under too similar climatic conditions. Kühl *et al.* (2002)/Kühl (2002) developed an algorithm for avoiding this. Gebhardt (2003)/Gebhardt *et al.* (2008), Simonis (2009)/Simonis *et al.* (2012) and Stolzenberger (2011) also used this approach which is also used in this study for the palaeoclimate reconstruction of Lake Prespa. Due to the fact that the procedure is not changed the important parts are summarized below.

The main part of this algorithm is the squared mahalanobis distance (Mahalanobis (1936)) estimated for the random variables T_k of two taxa a and b

$$\mathcal{D}^2\left(T_a, T_b\right) = \left(\vec{\mu}_a - \vec{\mu}_b\right)^T \mathbf{\Sigma}_a^{-1} \left(\vec{\mu}_a - \vec{\mu}_b\right). \tag{3.13}$$

In detail $\vec{\mu}_{k=a,b}$ and $\mathbf{\Sigma}_{k=a,b}$ are the parameters of a multivariate Normal distribution, the mean vector and the covariance matrix, estimated with the standard procedure via maximum likelihood estimation (MLE) for a multivariate Normal distribution. The algorithm filters the above described estimated presence/absence information $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ for every sample layer of the sediment core. The remaining taxa with presence $t_{\iota\,k_0} = 1$ of each sample layer fulfil $\mathcal{D} > \Delta_{\mathcal{D}} = 0.2$, a value already used in Simonis (2009) for three-dimensional climate state vector (eq. 3.1). The interpretation follows from the defintion of the mahalanobis distance $\mathcal{D}\left(T_a, T_b\right)$ of two



(a) $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ **(b)** $\mathcal{I}_{diff}^{N_\iota \times N_{k_0}} = \mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}} - \mathcal{I}_{filtered}^{N_\iota \times N_{k_0}}$ **(c)** $\mathcal{I}_{filtered}^{N_\iota \times N_{k_0}}$

Figure 3.3.: Indicator Taxa Matrices for Lake Prespa. All three figures show the absence information ($t_{\iota\, k_0} = 0$) for each regarded taxon k_0 in two light grey colours and the PAZ as defined in Panagiotopoulos *et al.* (2013) in red and labelled on the top. The presence information ($t_{\iota\, k_0} = 1$) is shown in two dark grey colours in (a) and (c) and in two blue colours in (b). The shortcuts on the vertical axis of (a) refer to the full taxon name in Table A.1 (p. 131). Figure (a) shows the original Indicator Taxa Matrix, a result of the ECDF approach, (c) shows the filtered Indicator Taxa Matrix, a result of the filtering algorithm with the mahalanobis distance and (b) shows the difference between them. The black numbers between (a) and (b) are the result of the sum over all sample layer ι for each taxon k_0 with $t_{\iota\, k_0} = 1$ in (a). The information next to this numbers are the corresponding results of the same summation for (b) with two exceptions: First if the complete taxon is removed by the filtering c. r. is written. Secondly if there is no filtering effect n. e. is written.

taxa: If two taxa are compared the algorithm ensures that the probability to observe taxon b (with $\vec{\mu}_b$ and $\mathbf{\Sigma}_b$) given the presence of taxon a is lower than 90 %.

Figure 3.3c shows the result for all sediment layers summarized in the filtered indicator taxa matrix $\mathcal{I}_{filtered}^{N_\iota \times N_{k_0}}$. The regarded climate state vector in the algorithm is $\vec{C} = (T_{Jan}, T_{Jul}, P_{ANN})^T$. Also shown in Figure 3.3 are the PAZ as defined in Panagiotopoulos *et al.* (2013). Visible in the difference $\mathcal{I}_{diff}^{N_\iota \times N_{k_0}}$ (Fig. 3.3b) is that the filtering has no effect (n. e.) on ten taxa of the selected taxa : *Quercus cerris* (Quce_Prm), *Quercus ilex* (Quil_m), *Phillyrea* (Phla_m), *Olea* (Oleu_m), *Cornus mas* (Coms_m), *Hippophae* (Hirh_Prm), *Buxus* (Buse_m), *Alisma cf. A. plantago-aquatica* (Alpl_Prm), *Hedera* (Hehe_m) and *Ephedra* (Ephed_Pr). Also three taxon are completely removed by the filtering: *Corylus* (Coav_m), *Tilia* (Tilia_Prm) and *Polemoniaceae* (Poca__Prm).

The result for $\mathcal{I}_{filtered}^{N_\iota \times N_{k_0}}$ presented in Figure 3.3c is used for all palaeoclimate reconstructions described in chapter 3.3. This is done to allow a comparison of the palaeoclimate reconstruction results for $\vec{C} = (T_{Jan}, T_{Jul}, P_{ANN})^T$ and $\vec{C} = (T_{Jan}, T_{Jul}, CWD_{ANN})^T$ with respect to the temperatures and to assess the uncertainty arranging from the different regarded water variables. Additionally CWD_{ANN} already includes P_{ANN} and therefore it is assumed that the extra information which could be obtained by a second filtering based on $\vec{C} = (T_{Jan}, T_{Jul}, CWD_{ANN})^T$ is not large.

The index notation in equation 3.10 does not reflect the described threshold exceedance and the mahalanobis filtering. If this is included the final equations for BITM which are applied on every pollen sample layer ι of a sediment core are:

$$f_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0 | 1_0, \ldots, 1_0, G_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right) \cdot \prod_{\substack{k \\ \forall t_{\iota k_0}^{filtered}=1}} \frac{f_{\vec{C}|T_k}\left(\vec{c}_0 | 1_0, G_0\right)}{f_{\vec{C}}\left(\vec{c}_0\right)}, \tag{3.14a}$$

$$f_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0 | \vec{t}_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right) \cdot \frac{\prod\limits_{\substack{k \\ \forall t_{\iota k_0}^{filtered}=1}} \mathbb{P}_{T_k|\vec{C}}\left(t_{k,0} | \vec{c}_0\right)}{\sum\limits_{\substack{k \\ \forall t_{\iota k_0}^{filtered}=1}} \prod\limits_{\substack{k \\ \forall t_{\iota k_0}^{filtered}=1}} \mathbb{P}_{T_k|\vec{C}}\left(t_{k,0} | \vec{c}_0\right)}. \tag{3.14b}$$

3.2.2. Palaeoclimatic Transfer Functions

Transfer function are the second contribution of taxa for palaeoclimate reconstructions. As aforementioned two methods are used for estimating the transfer functions: Gaussian mixture models and GLMs. It is important to point out that all estimations of the climate transfer functions have to regard the special character of precipitation: only positive values are allowed. Therefore the inverse CDF method is applied on the precipitation data. This approach uses the Gamma distribution $\mathcal{G}$ for a coordinate transformation that ensures that only positive values are regarded. The inverse CDF method is a straightforward method and can be found in the literature e. g. Gentle (2003). It is already used by Schölzel (2006) and Simonis (2009) for palaeoclimate reconstructions. If not explicitly mentioned every estimation in the following which involves precipitation uses this approach.

Estimation of Model Parameters - Maximum Likelihood Estimation

First MLE for the model parameters $\vec{\theta}$ is presented. This method is used in both transfer function concepts which are presented afterwards. Let $Y_1,\ldots,Y_{N_i}$ denote a sample of size N_i and each Y_i is independent and identically distributed (iid) with unknown pdf $f_{Y_i|\vec{\Theta}}\left(y_i|\,\vec{\theta}\right)$ which is described by the parameter vector $\vec{\theta}$. The joint probability density of the sample given the parameter is then given as:

$$f_{\vec{Y}|\vec{\Theta}}\left(\vec{y}\,\middle|\,\vec{\theta}\right)=\prod_{i=1}^{N_i} f_{Y_i|\vec{\Theta}}\left(y_i|\,\vec{\theta}\right). \tag{3.15}$$

The likelihood function is defined with fixed $\vec{y}$ and free variable $\vec{\theta}$:

$$\mathscr{L}\left(\vec{\theta}\,\middle|\,\vec{y}\right):=\prod_{i=1}^{N_i} f_{\vec{\Theta}|Y_i}\left(\vec{\theta}\,\middle|\,y_i\right). \tag{3.16}$$

The maximum likelihood estimator of the parameter vector $\vec{\theta}$ is the vector $\vec{\theta}_{opt}$ which maximizes $\mathscr{L}$. Since the logarithm is a strictly monotonic function, $\vec{\theta}_{opt}$ maximizes also the log-likelihood function

$$\ell\left(\vec{\theta}\,\middle|\,\vec{y}\right)=\ln\mathscr{L}\left(\vec{\theta}\,\middle|\,\vec{y}\right)=\sum_{i=1}^{N_i}\ln f_{\vec{\Theta}|Y_i}\left(\vec{\theta}\,\middle|\,y_i\right). \tag{3.17}$$

In the case of a continuous and differentiable pdf it is possible to calculate the maximization with the mathematical tool of curve sketching by calculating partial derivative of ℓ with respect to the θ_i. The most important step for estimating the coefficients is then to define the corresponding likelihood.

GLM and Palaeoclimatic Transfer Functions

GLMs for estimating the transfer functions were introduced by Schölzel (2006) and is one topic of the Diploma thesis of Stolzenberger (2011) and the Master thesis of Schult (2013).

First a brief introduction into the theory of GLM and its application to palaeoclimate transfer functions is given followed by the presentation of the result for *Olea*.

The theory presented here is a summary of the chapters 2, 3 and 7 of *An Introduction to Generalized Linear Models* written by Dobson, Barnett (2008). The model fitting process described in Dobson, Barnett (2008, begin of chapter 2) is divided in four steps: model specification, estimation of the parameters of the model, checking the adequacy of the model and inference. The last two parts are not part of this thesis and therefore not described here.

Model specification - a model is specified by two parts: the probability distribution of the response variable Y and an equation linking the response Y and explanatory variables X. For GLMs the probability distributions or pdf all belong to the family of exponential distributions, which includes the Normal, Binomial, Poisson and many other distribution (Dobson, Barnett 2008, chapter 2.3.2). The distribution of the exponential family can be written in the form of equation 3.18. The functions a and b are called the natural parameter, c and d are for a specific exponential pdf known (Dobson, Barnett 2008, chapter 3.2) :

$$f_{Y,\vec{\Theta}}\left(y,\vec{\theta}\right)=\exp\left[a(y)\,b\left(\vec{\theta}\right)+c\left(\vec{\theta}\right)+d(y)\right] \tag{3.18}$$

Table 3.1 summarizes the properties a, b, c and d for the most important pdfs of the exponential family.

Table 3.1.: Poisson, Normal and Binomial distribution as members of the exponential family with $a(y)=y$, after Dobson, Barnett (2008). Note that π in the Normal distribution $\mathcal{N}$ is the circle constant and not the prior probability as defined in chapter 1.2

distribution	pdf	Natural parameter $b(\theta)$	$c\left(\vec{\theta}\right)$	$d(y)$
Normal $\mathcal{N}$	$\dfrac{\exp\left[-\dfrac{(y-\mu)^2}{2\sigma^2}\right]}{(2\pi\sigma^2)^{\frac{1}{2}}}$	$\dfrac{\mu}{\sigma^2}$	$-\dfrac{\mu^2}{2\sigma^2}-\dfrac{1}{2}\ln\left(2\pi\sigma^2\right)$	$-\dfrac{y^2}{2\sigma^2}$
Binomial Bin	$\binom{n}{y}p^y(1-p)^{n-y}$	$\ln\left(\dfrac{p}{1-p}\right)$	$n\ln(1-p)$	$\ln\binom{n}{y}$
Bernoulli Bern	$p^y(1-p)^{n-y}$	$\ln\left(\dfrac{p}{1-p}\right)$	$\ln(1-p)$	$\ln\binom{1}{y}$

The aforementioned monotone and differentiable linking equation g has the form (following Dobson, Barnett (2008))

$$g\left[\mathrm{E}(Y)\right]=\underbrace{\alpha_0+\alpha_1x_1+\ldots+\alpha_{N_j}x_{N_j}=\alpha_0+\sum_{j=1}^{N_j}\alpha_jx_j\,.}_{\text{linear component}} \tag{3.19}$$

at which $\mathrm{E}(\cdot)$ is the general notation for the expectation value. With the following definitions for the vector of responses $\vec{Y}$, the vector of parameters $\vec{\alpha}$ and the design Matrix $\boldsymbol{X}$ which defines the linear model,

$$\vec{Y} := \begin{pmatrix} Y_1 \\ \vdots \\ Y_{N_i} \end{pmatrix}, \tag{3.20a}$$

$$g\left[\mathrm{E}\left(\vec{Y}\right)\right] := \begin{pmatrix} g\left[\mathrm{E}\left(Y_1\right)\right] \\ \vdots \\ g\left[\mathrm{E}\left(Y_{N_i}\right)\right] \end{pmatrix}, \tag{3.20b}$$

$$\vec{\alpha} := \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{N_j} \end{pmatrix}, \tag{3.21a}$$

$$\boldsymbol{X} := \begin{pmatrix} 1 & x_{11} & \cdots & x_{1j} & \cdots & x_{1N_j} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \vdots & \vdots & & x_{ij} & & \vdots \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 1 & x_{N_i1} & \cdots & x_{N_ij} & \cdots & x_{N_iN_j} \end{pmatrix}, \tag{3.21b}$$

it is possible to write equation 3.19 in matrix notation as:

$$g\left[\mathrm{E}\left(\vec{Y}\right)\right] = \boldsymbol{X}\vec{\alpha}. \tag{3.22}$$

The link function g is described in more detail later in the context of palaeoclimatic transfer functions when it is used.

For the **estimation of the model parameters** of GLMs two methods exist: *MLE* and *least squares*. The estimation of the model parameter $\vec{\alpha}$ of the GLMs in this thesis are performed with the statistical software R[1]. This software applies the MLE as described at the beginning of this section. The required definition for the log-likelihood function ℓ based on equation 3.18 are

$$\ell\left(\theta|y\right) = a\left(y\right)b\left(\theta\right) + c\left(\theta\right) + d\left(\theta\right), \tag{3.23a}$$

$$U\left(\theta|y\right) = \frac{\partial \ell\left(\theta|y\right)}{\partial \theta} = a\left(y\right)b'\left(\theta\right) + c'\left(\theta\right), \tag{3.23b}$$

The partial derivation $U(\cdot)$ of the log-likelihood is called the *score statistic*. With that tools it is now possible to estimate the GLM for the different kind of distribution for the response variables shown in Table 3.1.

The application of **GLM as palaeoclimatic transfer functions** is now straight forward. The input data is discussed and presented in detail in chapter 2. The transfer functions considered here connects the geographical distribution of a taxon (e. g. Figure 2.10 for *Olea*, p. 14) and the geographical climate maps (e. g. Figure 2.4 for CWD_{ANN}, p. 10). From this follows that the explanatory variable X in the case considered here is the modern climatic state $\vec{C}$ (eq. 3.1, p. 25, chap. 3.1.2) at all grid points and the response variable Y is the modern occurrence

---

[1] in more detail with the base package *stats* of R version 3.0.1, s. Team (2013) for more details

of the taxa $\vec{T}$ at (eq. 3.2, p. 25) all grid points. It is important to note here again that both information are used here (presence and absence) in contrast to the mixture model applied for example in Simonis (2009).

This means that realisation of the explanatory variable $\vec{x}_{\lambda\varphi}$ of the climate state $\vec{c}_{\lambda\varphi}$ at each grid point $\lambda\varphi$ is defined as

$$\begin{aligned} \left(\vec{x}_{\lambda\varphi}\right)^T &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}_{\lambda\varphi} = \begin{pmatrix} x_{\lambda\varphi 1} & x_{\lambda\varphi 2} & x_{\lambda\varphi 3} \end{pmatrix} && (3.24a)\\ \equiv \left(\vec{c}_{\lambda\varphi}\right)^T &= \begin{pmatrix} T_{Jan} & T_{Jul} & P_{ANN} \end{pmatrix}_{\lambda\varphi} && (3.24b)\\ &\overset{or}{=} \begin{pmatrix} T_{Jan} & T_{Jul} & CWD_{ANN} \end{pmatrix}_{\lambda\varphi}. && (3.24c) \end{aligned}$$

In this work a GLM with a quadratic term is applied. These last two considerations lead too the specification of the design matrix $\boldsymbol{X}$

$$\boldsymbol{X} := \left(\begin{array}{c|ccc|ccc|ccc} & \multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad\qquad}^{linear}} & & & & \multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad\qquad}^{quadratic}} \\ 1 & x_{11} & x_{12} & x_{13} & x_{11}x_{12} & x_{11}x_{13} & x_{12}x_{13} & x_{11}^2 & x_{12}^2 & x_{13}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{\lambda\varphi 1} & x_{\lambda\varphi 2} & x_{\lambda\varphi 3} & x_{\lambda\varphi 1}x_{\lambda\varphi 2} & x_{\lambda\varphi 1}x_{\lambda\varphi 3} & x_{\lambda\varphi 2}x_{\lambda\varphi 3} & x_{\lambda\varphi 1}^2 & x_{\lambda\varphi 2}^2 & x_{\lambda\varphi 3}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N_{\lambda\varphi}1} & x_{N_{\lambda\varphi}2} & x_{N_{\lambda\varphi}3} & x_{N_{\lambda\varphi}1}x_{N_{\lambda\varphi}2} & x_{N_{\lambda\varphi}1}x_{N_{\lambda\varphi}3} & x_{N_{\lambda\varphi}2}x_{N_{\lambda\varphi}3} & x_{N_{\lambda\varphi}1}^2 & x_{N_{\lambda\varphi}2}^2 & x_{N_{\lambda\varphi}3}^2 \\ & & & & \multicolumn{3}{c}{\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{bilinear}} & & & \end{array}\right) \tag{3.25}$$

at which $N_{\lambda\varphi}$ denotes the total number of grid points over land surface. It is important to note that the expansion from linear to quadratic introduces bilinear terms. Stolzenberger (2011) and Schult (2013) emphasize that this part contributes important information like the orientation in the climate phase space.

One point remains, the definition of the link function g. Due to the fact that the occurrence of one taxon is a binary variable (only presence $t_{\lambda\varphi} = 1$ or absence $t_{\lambda\varphi} = 0$, mentioned in chapter 3.1.1.), the corresponding probability distribution is the Bernoulli distribution or more general the Binomial distribution with $n = 1$ (see Tab. 3.1 for the formula). If the identity

$$g\left(p\right) = p, \tag{3.26}$$

is used as link function g, it is possible that the fitted values p_{opt} are not restricted to the interval $[0;1]$. For restricting to that interval the so called tolerance function $\mathit{tol}\left(\cdot\right)$ is used

$$p = \int_{-\infty}^{\boldsymbol{X}} \mathit{tol}\left(\boldsymbol{T}\right) d\boldsymbol{T}, \tag{3.27}$$

with the characteristics $\mathit{tol}\left(\boldsymbol{T}\right) \geq 0$ and $\int_{-\infty}^{\infty} \mathit{tol}\left(\boldsymbol{T}\right) = 1$, the requirements for a pdf. $\boldsymbol{T}$ is a matrix with the same properties as the design matrix $\boldsymbol{X}$. If we define

$$tol\left(\boldsymbol{T}\right) := \frac{\vec{\alpha}\exp\vec{\alpha}\boldsymbol{T}}{\left[1+\exp\vec{\alpha}\boldsymbol{T}\right]^{2}} \tag{3.28}$$

for the case considered here, integrate then equation 3.27 and solve the result for $\vec{\alpha}\boldsymbol{T}$, the equation for GLM (s. eq. 3.22) follows as

$$\underbrace{\ln\left(\frac{p}{1-p}\right)}_{=:g(p)} = \vec{\alpha}\boldsymbol{X}. \tag{3.29}$$

Here $g\left(p\right)$ is called the logistic link function. The transfer function as used in equation 3.14b is then interpreted as the aforementioned probability p estimated with the MLE $\vec{\alpha}_{opt}$

$$\mathbb{P}_{T_k|\vec{C}}\left(t_{k,0}|\,\vec{c}_0\right) := p = \frac{\exp\vec{\alpha}_{opt}\boldsymbol{X}}{1+\exp\vec{\alpha}_{opt}\boldsymbol{X}}. \tag{3.30}$$

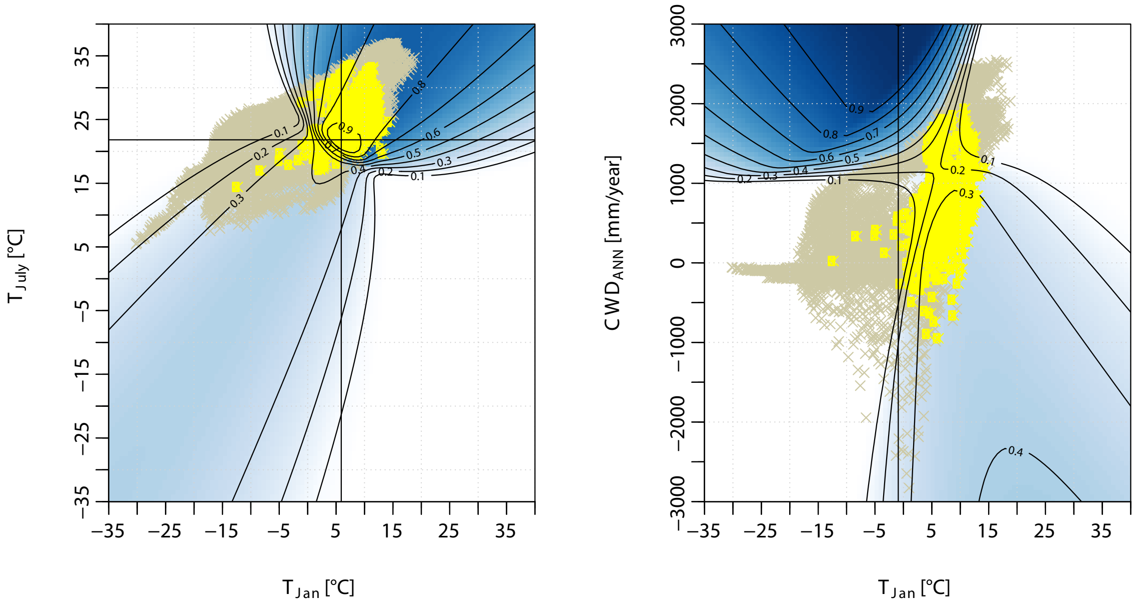
Figure 3.4 shows an example for the two-dimensional marginal distributions of the three-dimensional $\mathbb{P}_{T_k|\vec{C}}\left(t_{k,0}|\,\vec{c}_0\right)$ for *Olea*. Marginal here and later on denotes that an integration over the remaining dimensions of the climate state vector is applied, e. g. for Fig. 3.4a:

$$\mathbb{P}_{T_k|\vec{C}}\left(t_{k,0}|\,\vec{c}_0 = \left(T_{Jan}, T_{Jul}, CWD_{ANN}\right)^{T}\right) = \int\limits_{\mathcal{K}_{0,CWD_{ANN}}} \mathbb{P}_{T_k|\vec{C}}\left(t_{k,0}|\,\vec{c}_0\right) d(\vec{c}_{0,CWD_{ANN}}) \tag{3.31}$$

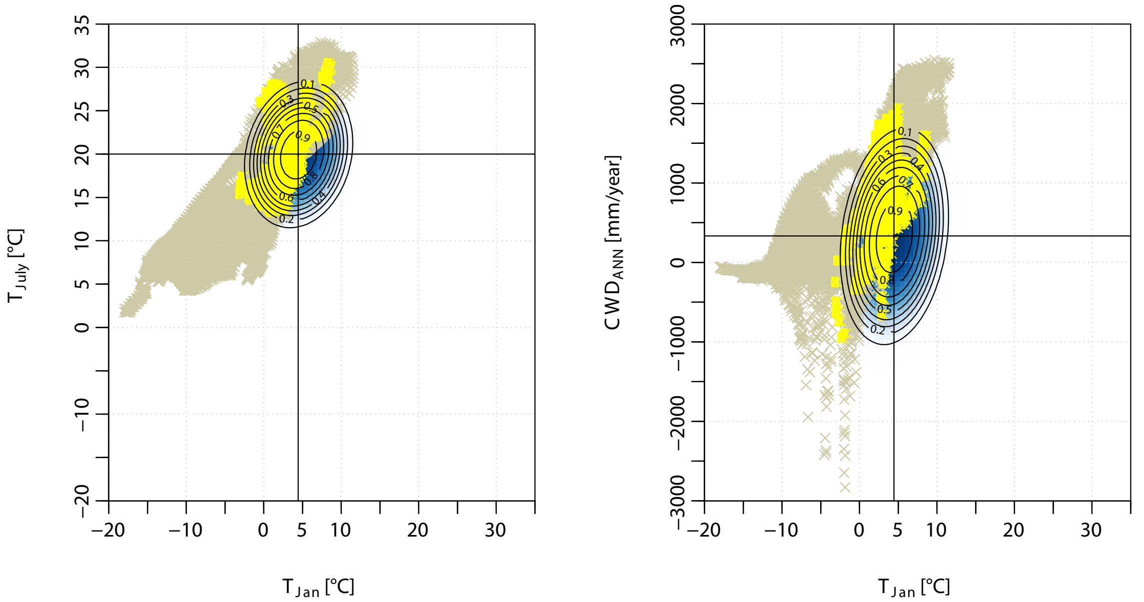
Figure 3.4a and 3.4b are estimated with the CRU TS 3.1 climate dataset and 3.4c and 3.4d with ERA-Interim for the 850 hPa temperatures. The missing two dimensional marginal distributions (T_{Jul}/CWD_{ANN} resp. $T_{Jul}^{850hpa}/CWD_{ANN}$) can be found in appendix A.2 (Fig. A.10 on p. 142) since Figure 3.4 only demonstrates how these transfer functions look like and what are the main differences between CRU TS 3.1 and ERA-Interim. In Figure 3.4 each cross represents one of the aforementioned realisations $\vec{c}_{\lambda\varphi}$ at all grid points over land surface. The colour of them marks if the taxon is present (yellow) or absent (grey). The contour lines represent the two dimensional marginal distributions of the estimated three dimensional GLM $\equiv$ the taxon specific palaeoclimate transfer function. The horizontal and vertical black line marks the maximum value of the estimated GLM.

The range of all $N_{\lambda\varphi} = 8438$ realisations $\vec{c}_{\lambda\varphi}$ in Figure 3.4 covers an interval of $[-30;20]$ °C for T_{Jan}, $[5;40]$ °C for T_{Jul}, $[-3000;2500]$ $\frac{\text{mm}}{\text{year}}$ for CWD_{ANN}, $[-20;15]$ °C for T_{Jan}^{850hpa} and $[0;35]$ °C for T_{Jul}^{850hpa}. This range does not change for the other taxa since it is the range of the considered datasets described in chapter 2.1. The range of the 731 realisations $\vec{c}_{\lambda\varphi}$ with *Olea* present ($t_k = 1$) is smaller: $[-15;15]$ °C for T_{Jan}, $[15;35]$ °C for T_{Jul}, $[-1000;2000]$ $\frac{\text{mm}}{\text{year}}$ for CWD_{ANN}, $[-5;10]$ °C for T_{Jan}^{850hpa} and $[10;30]$ °C for T_{Jul}^{850hpa}.

Summarized there is a difference in the input data for the GLM estimation in the temperatures,



(a) CRU TS 3.1 **(b)** CRU TS 3.1



(c) ERA-Interim-CRU TS 3.1 **(d)** ERA-Interim-CRU TS 3.1

Figure 3.4.: Palaeoclimatic transfer function for *Olea* estimated with a GLM based on the CRU TS 3.1 dataset ((a), (b)) resp. the ERA-Interim dataset for the 850 hPa temperatures ((c) and (d)). Each cross represents one realisation $\vec{c}_{\lambda\varphi}$ at all grid points over land surface. The colour of them marks if the taxon is present (yellow) or absent (grey). The contour lines and the colour scale represent the two dimensional marginal distributions of the estimated three dimensional GLM. Dark blue represent high values and light blue low values. The horizontal and vertical black line marks the maximum value of the estimated GLM.

which is in agreement with meteorological knowledge . More important is that the shape and orientation of the point cloud in Figure 3.4c and 3.4d is slightly rotated and more compact compared to Figure 3.4a, 3.4b. These differences lead to differences in the estimated GLM.

The estimated GLM for the surface climate variables has the shape of saddle. It matches for the T_{Jan}/T_{Jul} marginal distribution (Fig. 3.4a). For a detailed evaluation of the T_{Jan}/CWD_{ANN} marginal distribution (Fig. 3.4b) or how the input data in general is represented by the GLM the tools of verification as done by Stolzenberger (2011) have to be applied. This is not the aim of this study. Compared to the other presented transfer function the GLM in Figure 3.4c and 3.4d indicates a better agreement with the input dataset. This is in agreement with Stolzenberger (2011) who pointed out that the transfer functions estimated with a GLM and the temperatures of the middle troposphere (850 hPa) are a better representation of the input data.

3.2.3. $\pi_{\vec{C}}(\vec{c})$ Prior Distribution for the climate state vector

In BHM the general aim is "what is learnable from the data". In this study and specially in this chapter what is learnable from the pollen proxy data obtained from the sediment core of Lake Prespa in terms of climate. More general learnable means in other words (Schölzel 2006, p. 43): "*Every calculation of a posterior probability or probability density directly depends on the selection of the corresponding prior probability or probability density.*" Therefore the climate prior $\pi_{\vec{C}}\left(\vec{c}_0\right)$ (s. eq. 3.14) has to be defined.

Schölzel (2006) describes and summarises in detail the different types of climate priors for palaeoclimate reconstructions. The three categorised classes of priors are non-informative priors, maximum entropy priors and subjective priors. He also emphasises that "*there is vast information about the prior selection*". Due to the fact that the same type of prior is used in this work the selection and classification of priors is not further considered here. Only the influence on the posterior pdf is discussed in chapter 3.3. The prior type is an subjective prior determined by product of the marginal distributions of the multivariate prior distribution, i. e.

$$\pi_{\vec{C}}(\vec{c}_0 = (c_1, c_2, c_3)^T) := f_{\mathcal{N}\left(\mu^{prior}_{T_{Jan}}, \sigma^{prior}_{T_{Jan}}\right)}(c_1) \cdot f_{\mathcal{N}\left(\mu^{prior}_{T_{Jul}}, \sigma^{prior}_{T_{Jul}}\right)}(c_2) \cdot f_{\mathcal{G}\left(\nu^{prior}_{P_{ANN}}, \lambda^{prior}_{P_{ANN}}\right)}(c_3) \tag{3.32a}$$

$$\overset{or}{:=} f_{\mathcal{N}\left(\mu^{prior}_{T_{Jan}}, \sigma^{prior}_{T_{Jan}}\right)}(c_1) \cdot f_{\mathcal{N}\left(\mu^{prior}_{T_{Jul}}, \sigma^{prior}_{T_{Jul}}\right)}(c_2) \cdot f_{\mathcal{N}\left(\mu^{prior}_{CWD_{ANN}}, \sigma^{prior}_{CWD_{ANN}}\right)}(c_3) . \tag{3.32b}$$

The prior for the palaeoclimate reconstruction of the middle troposphere temperatures is of course estimated with the corresponding parameters.

For the estimation of this prior type the modern climate state on the location of Lake Prespa is used. In more detail histograms of the complete available time series of the CRU TS 3.1 and

ERA-Interim for the 850 hPa temperatures (presented in chap. 2.1.1 and chap. 2.1.3) are used for the definition of the required parameters. For a first guess the parameters $\mu_{\dots}^{prior}$ and $\sigma_{\dots}^{prior}$ of the univariate marginal distributions are estimated with MLE. Then the resulting $\mu_{\dots}^{prior}$ is rounded to the next integer. The variance $\sigma_{\dots}^{prior}$ is set to larger value as the estimated in order to allow to get enough information from the likelihood part of the BHM. Figure 3.5 and Table 3.2 show the result. Since the precipitation is not normally distributed the aforementioned inverse cumulative distribution function (CDF) method is applied (chap. 3.2.2) to obtain the parameters for the gamma distribution. Figure 3.5 demonstrates also that the defined climate prior is a good representation of the modern climate state.

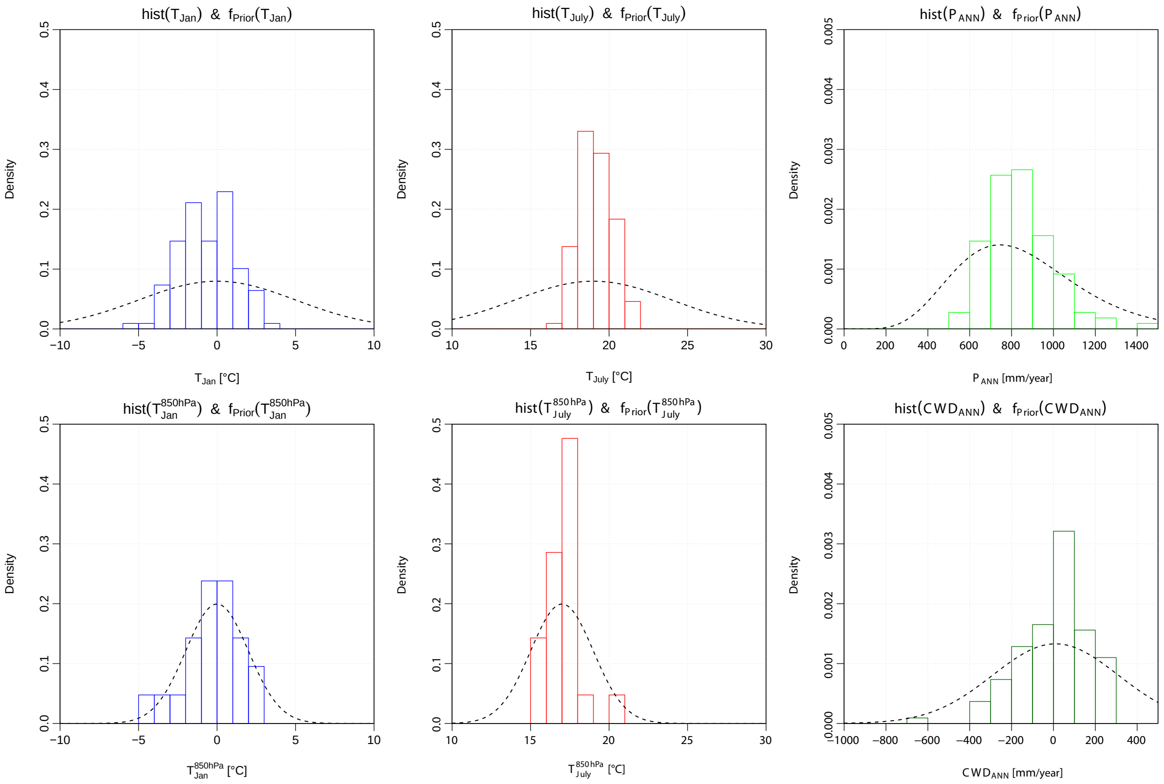


Figure 3.5.: Definition of the climate prior for Lake Prespa.
Each histogram bases on the complete available time series of the CRU TS Version 3.1 resp. ERA-Interim dataset on the location of Lake Prespa. The coloured curves are the estimated univariate climate prior pdfs. The corresponding parameters are listed in Table 3.2.

Table 3.2.: Parameters for the climate prior for Lake Prespa.

$\mu_{T_{Jan}}$	=	$0.0\,^{\circ}\mathrm{C}$	$\mu_{T_{Jul}}$	=	$19.0\,^{\circ}\mathrm{C}$	$\mu_{T_{Jan}^{850hpa}}$	=	$0.0\,^{\circ}\mathrm{C}$	$\mu_{T_{Jul}^{850hpa}}$	=	$17.0\,^{\circ}\mathrm{C}$
$\sigma_{T_{Jan}}$	=	$5.0\,^{\circ}\mathrm{C}$	$\sigma_{T_{Jul}}$	=	$5.0\,^{\circ}\mathrm{C}$	$\sigma_{T_{Jan}^{850hpa}}$	=	$2.0\,^{\circ}\mathrm{C}$	$\sigma_{T_{Jul}^{850hpa}}$	=	$2.0\,^{\circ}\mathrm{C}$
$\mu_{P_{ANN}}$	=	$850\,\frac{\mathrm{mm}}{\mathrm{year}}$	$\nu_{P_{ANN}}$	=	8.0278	$\mu_{CWD_{ANN}}$	=	$10.0\,\frac{\mathrm{mm}}{\mathrm{year}}$			
$\sigma_{P_{ANN}}$	=	$300\,\frac{\mathrm{mm}}{\mathrm{year}}$	$\lambda_{P_{ANN}}$	=	0.0094	$\sigma_{CWD_{ANN}}$	=	$300.0\,\frac{\mathrm{mm}}{\mathrm{year}}$			

3.3. Lake Prespa - Palaeoclimate Reconstruction Result

Figures 3.6 and 3.7 show the estimated result of the palaeoclimate reconstruction for Lake Prespa in terms of depth profile of the one-dimensional marginal pdfs. The term marginal is already defined in chapter 3.2.2. The integration required for the pdf is applied over the two remaining dimensions of the climate state vector. Thereby abscissa depth is replaced by age in y. cal BP due to the fact that an age-to-depth model established by Aufgebauer *et al.* (2012) is available (s. chap. 2.3.1). These pdfs are shown with a coloured scale, ranging from blue for low values to red for high values. The solid black line marks the mode, which is the highest value of each pdf in each sample layer. The dashed black lines mark the 10 %, 25 %, 75 % and the 90 % quantile[2]. The dotted line is the median, which is equal to the 50 % quantile. The red line is the weighted arithmetic mean for each layer, which is interpreted as the numerical calculated expectation value. The weights are the values of the marginal pdf.

The palaeoclimate reconstruction presented in Figure 3.6a, 3.6b and 3.6c base on taxa transfer functions estimated with CRU Time Series (CRU TS) 3.1 as climate dataset and $\vec{C} = (T_{Jan}, T_{Jul}, P_{ANN})^T$. Additionally Figure 3.6d shows the marginal pdf for CWD_{ANN} whose setup differs only in the third component of the climate state vector[3]. The corresponding figures for T_{Jan} and T_{Jul} are not shown since the result is equal.

Figure 3.7 presents the palaeoclimate reconstruction of the temperatures at the 850 hPa level. Surface temperatures based on the temperatures T_{Jan}^{850hpa} and T_{Jul}^{850hpa} could be determined with a simple height correction (s. e. g. for palaeoclimate reconstructions Stolzenberger 2011), based on the temperature gradient of the standard atmosphere ($6.5 \times 10^{-3}\frac{K}{m}$), the mean height of the 850 hPa level (1500 m) and the height above mean sea level (AMSL) for Lake Prespa (849 m AMSL, s. chap. 2.3.1): $\Delta T \approx 5\,^{\circ}\mathrm{C}$. In general the results of the palaeoclimate reconstruction of the surface temperatures (Fig. 3.6) are in accordance with the results of Panagiotopoulos *et al.* (2013). This is not the case for P_{ANN} and CWD_{ANN}, a fact that is discussed in the following.

In all figures a linear trend is estimated for several time slices for a better evaluation of

---

[2] The quantiles are estimated with the wtd.quantile function of the R-package Hmisc (Harrell Jr., with contributions from Charles Dupont and many others. 2013)

[3] in detail $\vec{C} = (T_{Jan}, T_{Jul}, CWD_{ANN})^T$

Table 3.3.: Results for the parameters of the estimated linear trends in Figures 3.6 and 3.7. The the acronym t. r. `X` stands for time range `X`.

	$c_i = a \cdot t + b$	$c_i = T_{Jan}$ Fig 3.6a	$c_i = T_{Jul}$ Fig 3.6b	$c_i = P_{ANN}$ Fig 3.6c	$c_i = CWD_{ANN}$ Fig 3.6d	$c_i = T^{850hpa}_{Jan}$ Fig 3.7a	$c_i = T^{850hpa}_{Jul}$ Fig 3.7b	$c_i = P_{ANN}$ Fig 3.7c	$c_i = CWD_{ANN}$ Fig 3.7d
t. r. 4	a	-0.0003	0.0000	-0.0022	-0.0141	-0.0002	-0.0001	-0.0180	0.0189
	$\pm\sigma$	0.0003	0.0002	0.0177	0.0282	0.0002	0.0003	0.0256	0.0342
	b	4.1426	21.6325	880.7740	50.5454	0.7013	16.0536	920.2494	-33.5076
	$\pm\sigma$	0.3471	0.2816	20.1403	32.0933	0.1810	0.3485	29.1066	38.9868
t. r. 3	a	-0.0001	0.0001	-0.0057	0.0085	-0.0001	0.0000	0.0018	0.0117
	$\pm\sigma$	0.0002	0.0002	0.0069	0.0161	0.0001	0.0001	0.0064	0.0190
	b	2.6475	20.7931	973.1271	-121.8717	0.0623	15.2152	923.0577	-190.9814
	$\pm\sigma$	1.7879	1.1211	50.3786	117.8900	0.6402	0.9527	46.8177	138.6456
t. r. 2	a	-0.0005	0.0003	-0.0012	0.0082	-0.0003	0.0001	-0.0027	-0.0201
	$\pm\sigma$	0.0006	0.0001	0.0124	0.0245	0.0003	0.0002	0.0278	0.0440
	b	5.7528	16.9299	1009.8411	-307.4034	2.3911	13.7953	1035.6623	33.9167
	$\pm\sigma$	7.1863	1.6707	149.5257	295.7989	3.8424	1.8240	336.3809	532.3978
t. r. 1	a	0.0014	0.0000	-0.0026	0.0123	0.0009	0.0002	-0.0049	-0.0184
	$\pm\sigma$	0.0015	0.0005	0.0311	0.0717	0.0010	0.0003	0.0586	0.1027
	b	-21.8232	20.3297	935.8191	-149.9581	-16.8841	11.9143	859.0741	372.3711
	$\pm\sigma$	22.8469	8.2093	480.1508	1106.8871	16.0768	4.6550	905.0205	1585.9224

climate changes. This trend is estimated on the numerical expectation value (red line) of the marginal pdfs of each climatolgical variable c_i. For the following time ranges trends are estimated :

time range 4 PAZ 1a to 1b, in detail pollen sample layer 1 to 11 which is equal to $0\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 2\,\mathrm{cm}$ and $2\,007\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 74\,\mathrm{cm}$ ($N_\iota = 11$)

time range 3 PAZ 1c to middle of 1e, in detail pollen sample layer 12 to 26 which is equal to $2\,713\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 82\,\mathrm{cm}$ and $9\,849\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 170\,\mathrm{cm}$ ($N_\iota = 15$)

time range 2 directly subsequent the next layer in PAZ 1e to second layer in PAZ 2b, in detail pollen sample layer 27 to 42 which is equal to $10\,239\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 178\,\mathrm{cm}$ and $13\,547\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 246\,\mathrm{cm}$ ($N_\iota = 16$)

time range 1 directly subsequent the next layer in PAZ 2b to PAZ 2c, in detail pollen sample layer 43 to 60 which is equal to $13\,790\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 251\,\mathrm{cm}$ and $17\,022\,\mathrm{y.cal\,BP} \overset{\approx}{\Leftrightarrow} 319\,\mathrm{cm}$ ($N_\iota = 18$)

These four time ranges, numbers from past to present and shortened from now on with t. r. X, are identifiable in all palaeoclimate reconstructions since the marginal pdfs profile before and after differs sometimes more sometimes less. In detail the definition of the time ranges bases in principle on a visual analysis of the T_{Jan} palaeoclimate reconstruction (Fig. 3.6a): What are the distinguishable zones? Due to the large variability in PAZ 2b and PAZ 2c the lower part of the profile is divided in two parts (t. r. 1 and 2). A fact that is supported by the visual analysis of the CWD_{ANN} marginal pdf profile (Fig. 3.6d): The profile in PAZ 2a differs from PAZ 2b and PAZ 2c.

The result of all estimated linear trends are summarized in Tab. 3.3. All slopes a of the estimated linear trends in Figure 3.6 and 3.7 ($c_i = a \cdot t + b$) are small with a maximum value of $|a| = 0.0201$ for the t. r. 2 and CWD_{ANN} with a large error ($\sigma = \pm 0.0440 \equiv 219\,\%$). If all errors are converted to relative errors (not shown) these are large for the most cases, except t. r. 2 and T_{Jul} ($\sigma = 0.0001 \equiv 33\,\%$). Thereby the largest relative errors for a arise for P_{ANN}. The slopes for the temperatures are in general one order of magnitude lower than those of CWD_{ANN} and P_{ANN} with a maximum value of $|a| = 0.0014$. The same holds for the most of the associated errors except that they differ not one order of magnitude. If the time ranges are compared among themselves it becomes obvious that time ranges 3 and 4 have lower values for a and also smaller relative errors for a and b.

The conclusion is that it is an acceptable approximation to consider only the intercept b as a representation of the expectation value in each time range or a simple mean of the expectation values in each layer (red line) for a simple hypothesis testing. This hypothesis testing is carried out with a Student's t-test for the means in the respective time ranges with null hypothesis that the means are equal. For this aim the R-function[4] *t.test()* is used. A significance level of 0.05 is applied for all tests.

It is indicated in all four subfigures of Figure 3.6 that there is a climate change which is diverse in its strength/ intensity in the considered climate variables. The clearest changes are indicated in the palaeoclimate reconstruction of T_{Jan} (Figure 3.6a) which is discussed first: The difference between t. r. 1 and t. r. 2 is insignificant, the Student's t-test results in a p-value of 0.85 for the two group means $-0.8\,°C$ and $-0.5\,°C$, so that the null hypothesis (means are equal) is accepted on a 5 % significance level. The deviation between the values of group means and the respective slopes is large, but the uncertainty of the slope a and the intercept b is also large (s. Tab. 3.3). This is in agreement with the aforementioned conclusion for applying the Student's t-test. The large variability in a short time in the considered time ranges is the reason for that uncertainty. The greatest change between two time ranges is identifiable between t. r. 2 and 3: The slopes b difference is $\approx 3\,°C$ (s. Tab. 3.3). The hypothesis testing also supports this with the lowest p-value for T_{Jan}: 4.5×10^{-4} (group means $2.1\,°C$ and $-0.8\,°C$) which leads to a rejection of the null hypothesis. The comparison of t. r. 3 and t. r. 4 also shows a significant difference (p-value 0.0032 with mean values of $3.9\,°C$ and $2.1\,°C$) .

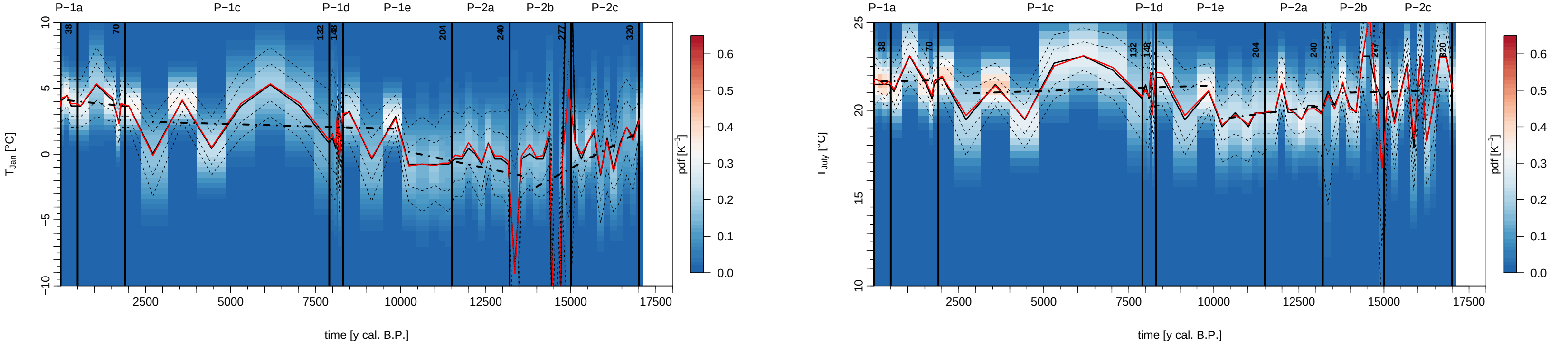
The visual analysis of the marginal pdf profile supports this: There are more and larger jumps in t. r. 3 than in t. r. 4 in relation to the estimated trend lines. This also holds for t. r. 3 versus t. r. 2 except that there is one extremely large jump ($\approx -8\,°C$). An extrapolation of the estimated linear trends from t. r. 2 into t. r. 3 resp. t. r. 3 into t. r. 4 and a consideration of the quantile curves indicates that the estimated lines are outside of the 25 % / 75 % quantile curves. In some cases they are actually outside of the 10 % resp. 90 % quantile curves. This point of view supports the just mentioned climate changes between the time ranges and a slight warming since the oldest time slice compared to the modern time slice.

The palaeoclimate reconstruction result for T_{Jan} PAZ 2b is in agreement with Panagiotopoulos *et al.* (2013): This analysis assumes a cold climate with oscillations in PAZ 2b. Also the rising winter temperatures in PAZ 1 can be confirmed as well as the so called 8.2-event $\equiv$ PAZ 1d: This event is associated with temperature oscillations in a very short time range which are identifiable in Figure 3.6a. Also proved are rising temperatures in the Middle and Late Holocene. The large winter temperature decline in PAZ 2a is not confirmable.

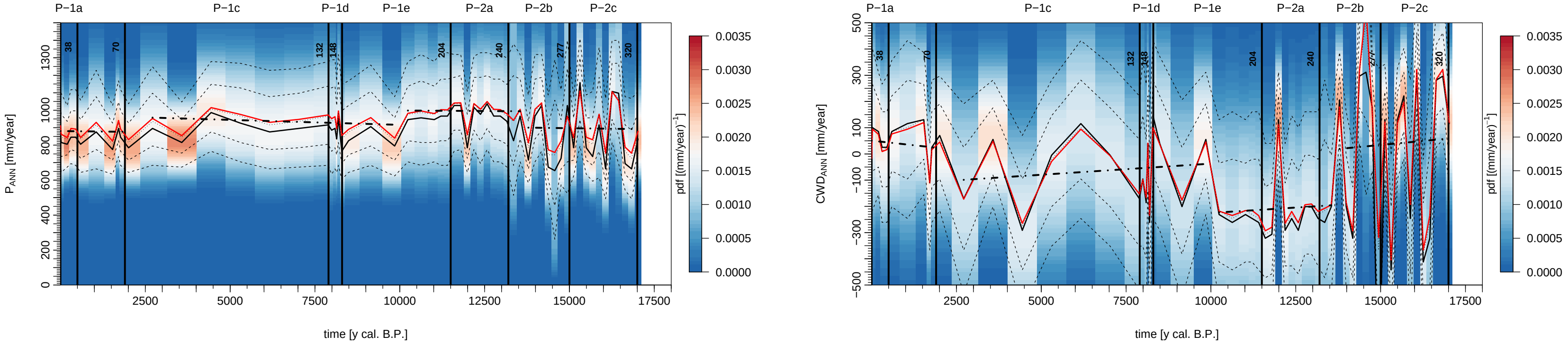
The palaeoclimate reconstruction of T_{Jul} (Figure 3.6b) shows four clearly discriminable time ranges but the climate changes are not as large as for T_{Jan}. With a cooling average of $\approx 1\,°C$ t. r. 1 is visual more distinguishable from t. r. 2 as it is the case for T_{Jan}. This is due to fact that the variability in t. r. 1 is larger for T_{Jul} if the large jump at the border of PAZ 2c is left out in this consideration. This is illustrated using the expectation value (red line): oscillations of $T_{Jan} \approx 3.5\,°C$ compared to $T_{Jul} \approx 7\,°C$. The Student's t-test results in a p-value of 0.048 (group means $21.1\,°C$ and $20.0\,°C$) which leads to a rejection of the null hypothesis on the given

---

[4] with its default setup (two-sided, no equal variances) of the base package *stats* of R version 3.0.1, s. Team (2013) for more details



(a) $c_0 = T_{Jan}$ **(b)** $c_0 = T_{Jul}$



(c) $c_0 = P_{ANN}$ **(d)** $c_0 = CWD_{ANN}$

Figure 3.6.: Marginal posterior pdfs $f_{C_0|\vec{T}_0}\left(c_0|\vec{t}_0\right)$ for T_{Jan}, T_{Jul}, P_{ANN} and CWD_{ANN} for the Lake Prespa palaeoclimate reconstruction based on equation 3.14b. In detail all base on the filtered Indicator Taxa Matrix for Lake Prespa $\mathcal{I}^{N_t \times N_{k_0}}_{filtered}$ (Fig. 3.3c). (a), (b) and (c) base on $\vec{C} = (T_{Jan}, T_{Jul}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset. (d) differs only in the definition the climate state vector ($\vec{C} = (T_{Jan}, T_{Jul}, CWD_{ANN})^T$). The pdfs are shown with a coloured scale, ranging from blue for low values to red for high values. The solid black line mark the mode, which is the highest value of each pdf in each sample layer. The dashed black lines mark the 10 %, 25 %, 75 % and the 90 % quantile. The dotted line is the median, which is equal to the 50 % quantile. The red line is the weighted arithmetic mean for each layer, which is interpreted as the numerical calculated expectation value. The weights are the values of the marginal pdf. The PAZ as defined in Panagiotopoulos *et al.* (2013) are shown in black and are labelled with the according depth. The age-to-depth model is established by Aufgebauer *et al.* (2012) (mentioned in chapter 2.3.1).

significance level. If the complete marginal pdf is considered and not the expectation value alone the climate change has to be classified as a change of the variability and not as shift to a colder time range. This is a fact which is also supported by the order of magnitude of the p-value.

The comparison of t. r. 2 with t. r. 3 with the hypothesis testing results in a p-value of 0.0013 for the group means 20.0 °C and 21.2 °C (null hypothesis rejected). The analysis of the marginal pdf in t. r. 3 indicates variability on the scale of t. r. 1 but on a larger time scale and with a longer period with higher T_{Jul} values (around 6 000 y. cal BP) as the estimated trend. A comparison of t. r. 3 with t. r. 4, the time range which includes the modern climate, shows that the difference between these time ranges on average is comparably small: The p-value is 0.22 with mean values of 21.2 °C and 21.7 °C. This leads to the conclusion that the only time range with a significant difference to the modern T_{Jul} climate for a longer time scale is t. r. 2. This cooling in T_{Jul} is supported by the above presented extrapolation of the linear trend of t. r. 2 into t. r. 3 and t. r. 4 and a consideration of the quantile curves: the extrapolation is outside of the 10 % quantile except for the two time slices in PAZ 1c (around 2 500 y. cal BP and 3 500 y. cal BP), the also recognizable 8.2-event and time slice 9 000 y. cal BP. These results are in accordance with Panagiotopoulos *et al.* (2013): "*a negligible decline in summer temperatures in PAZ 2a*" and "*climate oscillations in PAZ 2b*".

The marginal pdf for the climatolgical variable $c_i = P_{ANN}$ presented in Figure 3.6c shows almost no climate change: The extrapolation of each estimated linear trend into the other ones are within the 25 % / 75 % quantile curves. The only time range that differs from the other ones is t. r. 1: The variability in this time range is larger than in the other time ranges. A comparison of t. r. 1 with e. g. t. r. 4 by applying the Student's t-test results in a p-value of $p = 0.59$ for $879 \, \frac{\text{mm}}{\text{year}}$ and $896 \, \frac{\text{mm}}{\text{year}}$ which leads to an acception of the null hypothesis that the mean values for these time ranges are equal. This fact does not change if the complete marginal pdf profile is considered: The 10 % / 90 % quantile curves vary in the complete profile but only in the lowest/oldest time slice the range of them shifts to lower annual precipitation values. Two results of Panagiotopoulos *et al.* (2013) can be confirmed: the already mentioned climate oscillations and the short time range variation within 8.2-event $\equiv$ PAZ 1d.

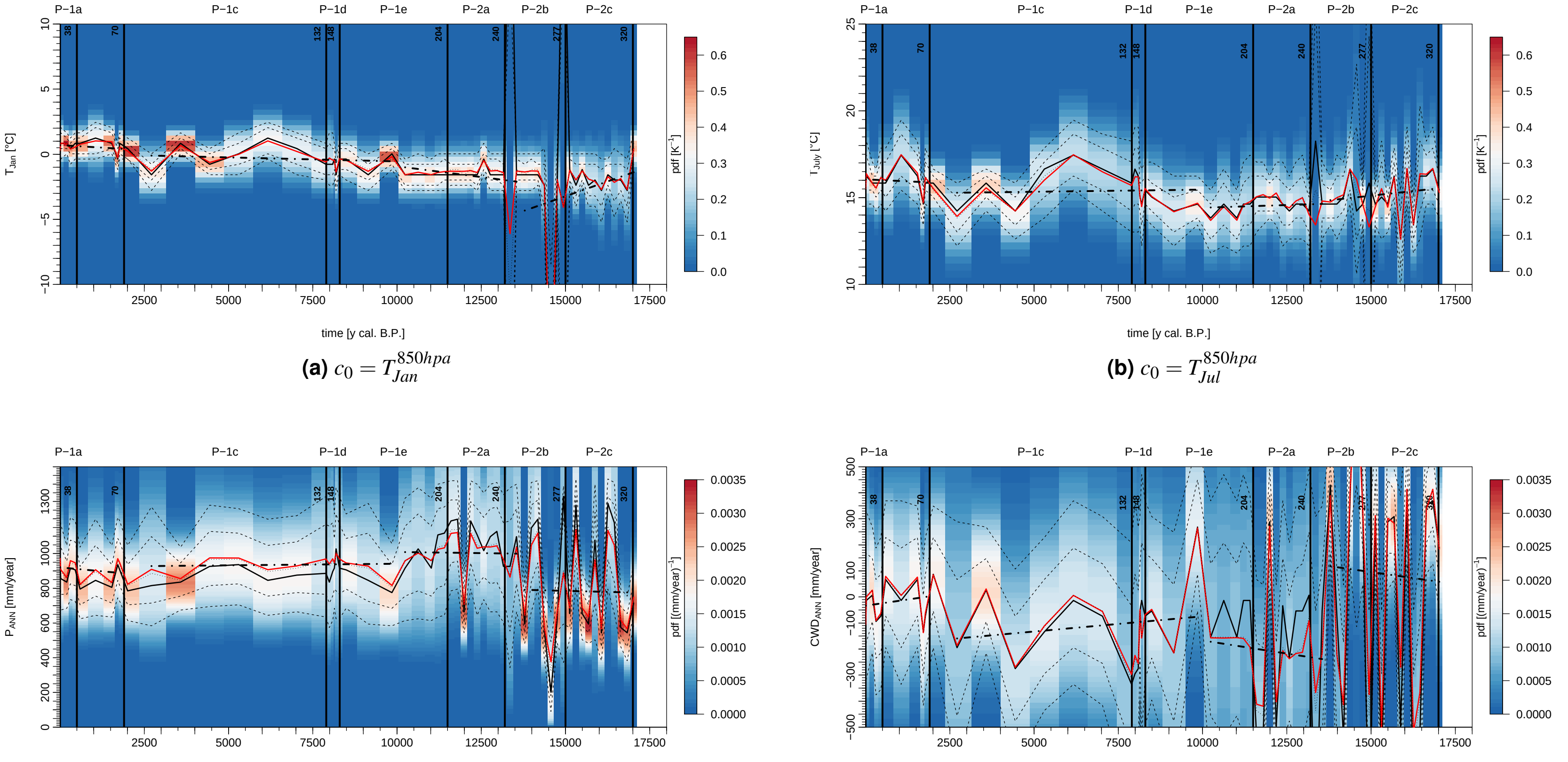
In contrast to the P_{ANN} palaeoclimate reconstruction CWD_{ANN} (Figure 3.6d) indicates four clearly discriminable time ranges. Just to remind: $CWD_{ANN} > 0$ is equivalent to dry and $CWD_{ANN} < 0$ to wet environmental condition. On average the climate in t. r. 1 is dry with a large variability. The change compared to t. r. 2 is clearly recognizable and significant as the Student's t-test shows a p-value of 0.0025 for the group means $39 \, \frac{\text{mm}}{\text{year}}$ and $-208 \, \frac{\text{mm}}{\text{year}}$. This is also the case if the marginal pdf, represented by the quantile curves, is considered: The aforementioned trend extrapolated into t. r. 2 lies outside of the 75 % quantile curve. But a comparison of t. r. 1 with t. r. 4, the quasi modern climate, indicates that on average the climate in t. r. 1 is not significantly different (p-value of 0.99 for $39 \, \frac{\text{mm}}{\text{year}}$ and $38 \, \frac{\text{mm}}{\text{year}}$).

On average the difference of t. r. 2 to t. r. 3 is also significant as the result of the Student's t-test shows: a p-value of 0.0010 for $-208\,\frac{\text{mm}}{\text{year}}$ and $-62\,\frac{\text{mm}}{\text{year}}$. Although in t. r. 2 almost all time slices show a marginal pdf where 75% of the pdf-range indicates a wet climate: the 75 % quantile curve lies below $CWD_{ANN} \equiv 0$ except for time slice 11 990 y. cal BP ($\equiv$ 214 cm). On average the climate in t. r. 3 is also wet (group mean $-62\,\frac{\text{mm}}{\text{year}}$) but the variability is larger as in t. r. 2. This variability is in the order of magnitude of t. r. 1 ($\approx 200\,\frac{\text{mm}}{\text{year}}$) but not on the time scale: there is e. g. a longer period with values of $CWD_{ANN} > 0$ around 6 000 y. cal BP. The climate change (on average) from t. r. 3 to t. r. 4 is significant on the 5 % significance level (p-value of $p = 0.01$ for $-62\,\frac{\text{mm}}{\text{year}}$ and $38\,\frac{\text{mm}}{\text{year}}$). In summary the time ranges differ significantly among themselves but only time range 2 with a significantly different climate to the modern time slice.

There is only partial accordance with the results of Panagiotopoulos *et al.* (2013): Confirmed are a dry climate on average as well as climate oscillations in PAZ 2c. These oscillations continue in the next zone, PAZ 2b. Also confirmed is the 8.2-event $\equiv$ PAZ 1d but not as described by Panagiotopoulos *et al.* (2013) with a short return to return to dry but wet conditions. Rather disproved than confirmed is the moisture deficit in PAZ 2a since $CWD_{ANN} < 0$ in Figure 3.6d. Additionally, the described increasing moisture availability in PAZ 1e which should continue in PAZ 1c, 1b, 1a is disproved. Figure 3.6d indicates a decreasing moisture availability. Since two components of the middle troposphere temperature reconstruction are also a P_{ANN} resp. CWD_{ANN} reconstruction, the final conclusion for these variables is drawn after the discussion of Figure 3.7c and 3.7d.

The result for the palaeoclimate reconstructions of the middle troposphere temperatures in Figure 3.7 is different for some of the marginal pdfs. The ordinate scale in Figure 3.7 is the same as in 3.6 so that one main difference between the 850 hPa temperature reconstruction and those of surface values becomes obvious: A slightly narrower marginal pdf profile for T_{Jan}^{850hpa} quantified by the distance between e. g. the 10 % and 90 % quantile curves: $T_{Jan}^{850hpa} \approx \max 4-5\,^{\circ}\text{C}$ (with 5 °C in the lower part of the sediment core) and $T_{Jan} \approx \max 6\,^{\circ}\text{C}$. This fact holds true exept for two layers in PAZ-2b with $T_{Jan}^{850hpa} \approx \max >> 20\,^{\circ}\text{C}$ and $T_{Jan} \approx \max 15\,^{\circ}\text{C}$. The distance for $T_{Jul}^{850hpa} \approx T_{Jul} \approx \max 6\,^{\circ}\text{C}$, $P_{ANN} \approx \max 1000\,\frac{\text{mm}}{\text{year}}$ and $CWD_{ANN} \approx \max 1200\,\frac{\text{mm}}{\text{year}}$ is the same. The discussed differences between the time ranges in the surface temperature palaeoclimate reconstruction (Fig. 3.6) are generally persistent respective to the significance and the direction of the change. The exceptions are presented in the following.

The just mentioned fact of a narrower marginal pdf for T_{Jan}^{850hpa} is also indicated with a reduced variability in PAZ-2c. This is e. g. also visible in the jumps of the expectation value (red line): $\Delta T_{Jan} \approx 3\,^{\circ}\text{C}$ and $\Delta T_{Jan}^{850hpa} \approx 1\,^{\circ}\text{C}$. Also the time ranges are more similar so that the aforementioned warming since the oldest time slice is weaker. The most clear quantifiable change is that T_{Jan}^{850hpa} is on average in t. r. 2 more similar to t. r. 3 as it is the case for T_{Jan}: The p-value is 0.0018 for the group means$-1.60\,^{\circ}\text{C}$ $-0.37\,^{\circ}\text{C}$, so that the null hypothesis is rejected.



(a) $c_0 = T_{Jan}^{850hpa}$

(b) $c_0 = T_{Jul}^{850hpa}$

(c) $c_0 = P_{ANN}$

(d) $c_0 = CWD_{ANN}$

Figure 3.7.: Marginal posterior pdfs $f_{C_0|\vec{t}_0}\left(c_0|\vec{t}_0\right)$ for T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, P_{ANN} and CWD_{ANN} for the Lake Prespa palaeoclimate reconstruction based on equation 3.14b. In detail all base on the filtered Indicator Taxa Matrix for Lake Prespa $\mathcal{I}_{filtered}^{N_t \times N_{k_0}}$ (Fig. 3.3c). (a), (b) and (c) base on $\vec{C} = \left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, P_{ANN}\right)^T$ and ERA-Interim for the 850 hPa temperatures as estimation dataset. (d) differs only in the definition the climate state vector ($\vec{C} = \left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, CWD_{ANN}\right)^T$). The mapping of the figure is identical to Figure 3.6.

The differences for the palaeoclimate reconstructions of T_{Jul} and T_{Jul}^{850hpa} are not as easily identifiable as in the last case. The most important result is that the main difference between them is almost only a shift of about 5 °C. The marginal pdf for T_{Jul}^{850hpa} is also a little narrower but it is not easy to recognize and quantify this in the younger/upper part of the palaeoclimate reconstruction. In the older/lower part (≈ t. r. 1 and t. r. 2) narrow is here equivalent to a reduced variability of 6 °C compared to 4 °C for e. g. the expectation value. So one main conclusion is that the aforementioned height correction for the middle troposphere temperatures is generally confirmed, whereby the better representation for this fact is $T_{Jul} \Leftrightarrow T_{Jul}^{850hpa}$.

One important result for the palaeoclimate reconstructions of P_{ANN} and CWD_{ANN} as component of different setups of the climate state vector is that they are in general in accordance and not inconsistent. In detail this means that on average (comparison of the group means) the sign of the differences between the different time ranges is the same. As aforementioned the interquantile distance is also the same.

The group means are on average different whereby the greatest differences appear for P_{ANN} in t. r. 1 with $|\Delta| \approx 103\,\frac{\mathrm{mm}}{\mathrm{year}}$ and t. r. 4 with $|\Delta| \approx 25\,\frac{\mathrm{mm}}{\mathrm{year}}$. A direct consequence of the last mentioned result is that the oldest time range is more discriminable in Figure 3.7c than in 3.6c from the rest of the marginal pdf profile. Thereby this is not statistically significant: The Student's t-test result is a p-value of 0.05 which is equal to the defined significance level. Additionally the trend extrapolation from t. r. 1 in t. r. 2 in Figure 3.7c supports this: It lies almost lower as the 25 % quantile curve instead of almost above. The only fact of the P_{ANN} palaeoclimate which is reliable are the climatic oscillations in t. r. 1.

The palaeoclimate reconstruction for CWD_{ANN} presented in Figure 3.7d differs from Figure 3.6d on average in each time range except in t. r. 2: t. r. 1 with $|\Delta| \approx 49\,\frac{\mathrm{mm}}{\mathrm{year}}$, t. r. 3 with $|\Delta| \approx 47\,\frac{\mathrm{mm}}{\mathrm{year}}$ and t. r. 4 with $|\Delta| \approx 55\,\frac{\mathrm{mm}}{\mathrm{year}}$. Thereby the change in t. r. 4 is from dry (Fig. 3.6d) to wet (Fig. 3.7d) environmental conditions. On average t. r. 1 becomes more wet and t. r. 3 and t. r. 4 more wet as the comparison of Figure 3.6d and 3.7d indicates. As a direct consequence the result from Figure 3.6d that t. r. 2 is the only time range which differs significantly from the modern time range is weakened: Figure 3.6d with $p = 1.5 \times 10^{-8}$ compared with Figure 3.7d with $p = 0.0005$, so the null hypothesis is rejected and the difference of the mean values is significant.

The aforementioned conclusions of the palaeoclimate reconstruction results of the water variables respective Panagiotopoulos *et al.* (2013) are also consistent with the results discussed in the last two paragraphs. Over all it is important to note that the palaeoclimate reconstruction of P_{ANN} and CWD_{ANN} of Lake Prespa are in general sensitive to the setup of the climate state vector $\vec{c}$, more precisely: What are the other components (middle troposphere temperatures or surface temperatures)?

In this last paragraph the palaeoclimate reconstruction results for Lake Prespa are briefly summarized: There are four discriminable time ranges identifiable in the surface temperature palaeoclimate reconstructions for T_{Jan} and T_{Jul} with greater climate changes for T_{Jan} than for T_{Jul}. For T_{Jan} the greatest change between two time ranges is identifiable between t. r. 2 and 3. Time range 3 also differs significantly from the modern climate. For T_{Jul} only one time range is identifiable with a significant difference to the modern T_{Jul} climate for a longer time scale: t. r. 2. The palaeoclimate reconstruction of the middle troposphere temperatures T_{Jan}^{850hpa} and T_{Jul} (Fig. 3.7a resp. 3.7b) support this results: In general the differences between the time ranges are persistent respective to the significance and the direction of the change compared to the surface temperature results presented in Figure 3.6a resp. 3.6b. The most clear quantifiable change is that T_{Jan}^{850hpa} is on average in t. r. 2 more similar to t. r. 3 as it is the case for T_{Jan}. Also the marginal pdfs for T_{Jan}^{850hpa} and T_{Jul}^{850hpa} are slightly narrower than those of T_{Jan} and T_{Jul}. As a consequence the simple height correction mentioned in the beginning is also confirmed. The result for the palaeoclimate reconstructions of the water variables P_{ANN} and CWD_{ANN} is different from that: One important result is that the different setup of the climate state vector yields to results which are in general not inconsistent. This means that the sign of the differences between the group means of the time ranges is the same. For both P_{ANN} reconstruction results, Figure 3.6c and 3.7c, only the climatic oscillations in t. r. 1 are reliable. Compared to the P_{ANN} palaeoclimate reconstruction both CWD_{ANN} results (Figure 3.6d and 3.7d) indicate four clearly discriminable time ranges. These differ significantly among themselves with only one time range with a really visible and significantly different climate to the modern time slice: time range 2.

4

Lake Kinneret and Ein Gedi

This chapter presents the result and the mathematical tools that are used for the palaeoclimate reconstruction that bases on the sediment core data of Lake Kinneret and Ein Gedi with a focus on Lake Kinneret. The sediment cores and the locations are described in chapter 2.3.2. The basic mathematical principles for this purpose are introduced in chapter 1.2 and are extended in the first section of this chapter to the Bayesian Biome Model (BBM) which is applied for both sediment cores. The following section describes the estimation of each component of the BBM with the corresponding results for Lake Kinneret due to the fact that the sediment core of Ein Gedi is already analysed in Litt *et al.* (2012). The next part shows an experiment with the pollen-ratio model (PRM) for both sediment cores. Directly follows the palaeoclimate reconstruction for Ein Gedi whereby the focus are the adaptations of Litt *et al.* (2012) on the requirements of the Jordan Valley reconstruction presented in chapter 6. In the last section some general conclusions for a palaeoclimate reconstruction with the BBM are presented.

4.1. The Bayesian Biome Model - Introduction

The BBM is developed by Schölzel (2006) and after some improvements successfully applied to the Ein-Gedi Sediment core (Litt *et al.* 2012). It bases on the concept of biomisation, which is an "*empirical classification scheme of plant taxa into well defined groups*" (Schölzel 2006). For this purpose each plant taxon which could be identified in the palaeo record is assigned to one biome (s. chap. 4.2.2). In a nutshell the BBM is a reduction of complexity of the vegetation (Schölzel (2006), 5.1).

BBM defines a biome as a vegetation area. It can be characterized by:

- **the botanical composition**, defined as a list of the taxa assigned to the biome type;
- **the geographical distribution**, here synonymous to the vegetation areas;
- **the climatic range**, defined by a list of typical ranges for climate variables

The basic hypothesis of the biome model is (see Schölzel (2006), p. 56):

> *"Climate changes during the Holocene did not change the number of biomes or their botanical composition, but caused a relocation of the three territories (two in the case of Kinneret). The influence of each biome on the Dead Sea (or Lake Kinneret), as had been recorded by the fossil pollen spectra of the sediment core, changed over time."*

The Bayesian Indicator taxa method, used for the palaeoclimate reconstructions described in chapter 3, defines the "*transfer functions as climate given taxa* " (Schölzel 2006, chap. 5.1, p. 56). This is no reasonable interpretation in the case of biomes, e.g. "*climate given a 75 % probability of biome*". Litt *et al.* (2012) describe the correct view: "*we always observe a mixture signal of all biomes*".

$$\vec{\mathcal{B}} := \left(\mathscr{b}_1, \ldots, \mathscr{b}_{N_l}\right)^T, \vec{\mathscr{b}} \in \{1, 2, \ldots, l, \ldots, N_l\}^{N_l} \tag{4.1}$$

defines the random variable for biomes needed for the following definitions. Since the purpose is to estimate the probability or pdf of climate data given proxy data the starting point is

$$\mathbb{P}_{\vec{C},\vec{P}}\left(\vec{c}, \vec{p}\right) = \sum_{l=1}^{N_l} \mathbb{P}_{\vec{C},\vec{P},\mathcal{B}_l}\left(\vec{c}, \vec{p}, \mathscr{b}_l\right) \tag{4.2}$$

The right hand side reflects the above described mixture signal of all biomes and is the summation over all different biome types N_l. The Application of the Bayes theorem on both sides of the equation leads to

$$\mathbb{P}_{\vec{C}|\vec{P}}\left(\vec{c} \,\middle|\, \vec{p}\right) = \sum_{l=1}^{N_l} \mathbb{P}_{\mathcal{B}_l|\vec{P}}\left(\mathscr{b}_l | \vec{p}\right) \mathbb{P}_{\vec{C}|\vec{P},\mathcal{B}_l}\left(\vec{c} \,\middle|\, \vec{p}, \mathscr{b}_l\right) \tag{4.3}$$

The equation for the BBM is obtained after the realisation of three further steps:

1.) Assume that the connection between climate and biomes is conditional independent given the proxy data (established by Litt *et al.* (2012)).

2.) The application of the basic assumption described in chapter 1.2 (eq. 1.5 on p. 5).

3.) Another application of the Bayes theorem to get the required definition for the biome transfer function:

$$\mathbb{P}_{\vec{C}_0|\vec{P}_0}\left(\vec{c}_0 | \vec{p}_0\right) = \sum_{l=1}^{N_l} \mathbb{P}_{\mathcal{B}_l|\vec{P}}\left(\mathscr{b}_{l_0} | \vec{p}_0\right) \frac{\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0} | \vec{c}_0\right) \pi_{\vec{C}}\left(\vec{c}_0\right)}{\mathbb{P}_{\mathcal{B}_l}\left(\mathscr{b}_{l_0}\right)} \tag{4.4}$$

This is in principle the equation for the palaeoclimate reconstructions based on the BBM. The reconstructions presented in the following are pdfs. As discussed in the last chapter 3 it is possible to switch to a posterior probability density function w. l. o. g. for every pollen sample layer ι found in depth d_ι

$$f_{\vec{C}_0|\vec{P}_0}\left(\vec{c}_0|\vec{p}_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right)\cdot\sum_{l=1}^{N_l} W_{B_{l_0}} \frac{\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\vec{c}_0\right)}{m_{\mathcal{B}_l}\left(b_{l_0}\right)}. \tag{4.5}$$

In equation 4.5 $\mathbb{P}_{\mathcal{B}_l|\vec{P}}\left(b_{l_0}|\vec{p}_0\right)$ is replaced by the biome probability $W_{B_{l_0}}$ (s. chap. 4.2.2 for details) and denotes the contribution of the pollen spectrum. Also $\mathbb{P}_{\mathcal{B}_l}\left(b_{l_0}\right)$ is replaced by $m_{\mathcal{B}_l}\left(b_{l_0}\right)$ since it fulfils the definition for a marginal probability. $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\vec{c}_0\right)$ is the biome transfer function or biome likelihood described in chapter 4.2.3.

4.2. The BBM in Detail for Lake Kinneret

The BBM is used for the reconstruction for Lake Kinneret because the circumstances are similar to those of Ein Gedi: The vegetation zones changes within a few kilometres in this area and the collection area for the pollen, present in the core, covers a wide region. This section explains each component of the BBM and presents the corresponding results for Lake Kinneret.

4.2.1. $\pi_{\vec{C}}(\vec{c}_0)$ Prior Distribution for the climate state vector

Chapter 6 presents an approach which could be classified as climate field reconstruction (CFR): an interpolation of the local palaeoclimate reconstructions in the Jordan Valley. As mentioned in chapter 3.2.3 in BHM the general aim is "what is learnable from the data". Because in chapter 6 the aim is to learn from the data something about the climate changes in the Jordan Valley a climate prior $\pi_{\vec{C}}\left(\vec{c}_0\right)$ for the whole Valley is defined. The estimation process is not changed compared to the Lake Prespa palaeoclimate reconstruction. The climate prior definition in equation 3.32 is not changed for the three-dimensional climate state vector here. For two dimensions the definition changes only in that way, that it is only a product of the required two marginal distributions which are connected to the corresponding climate variable. The result of the estimation for the required parameters in the definition of the climate prior are presented in Table 4.1.

Table 4.1.: Definition of the climate prior for the Jordan Valley in Israel

$\mu_{T_{DJF}}$	=	12.0°C	$\mu_{T_{JJA}}$	=	25.0°C	$\mu_{T^{850hpa}_{DJF}}$	=	5.0°C	$\mu_{T^{850hpa}_{JJA}}$	=	21.0°C
$\sigma_{T_{DJF}}$	=	4.0°C	$\sigma_{T_{JJA}}$	=	4.0°C	$\sigma_{T^{850hpa}_{DJF}}$	=	2.0°C	$\sigma_{T^{850hpa}_{JJA}}$	=	2.0°C
$\mu_{P_{ANN}}$	=	$410\,\frac{\text{mm}}{\text{year}}$	$\nu_{P_{ANN}}$	=	4.912	$\mu_{CWD_{ANN}}$	=	$1000\,\frac{\text{mm}}{\text{year}}$			
$\sigma_{P_{ANN}}$	=	$185\,\frac{\text{mm}}{\text{year}}$	$\lambda_{P_{ANN}}$	=	0.012	$\sigma_{CWD_{ANN}}$	=	$250\,\frac{\text{mm}}{\text{year}}$			

4.2.2. $W_{B_{l_0}}$ Biome Probability

$W_{B_{l_0}}$ is the probability that a biome contributes to a climate state, that is going to be reconstructed (Schölzel 2006). Its determination bases on the calculation of affinity scores. For that reason it replaces the probability $\mathbb{P}_{\mathcal{B}_l|\vec{P}}\left(b_{l_0}|\,\vec{p}_0\right)$ in equation 4.5. If all relative abundances $\widetilde{\omega}_{\iota\,k_0}$ are summarized in the same way as in chapter 3.2.1 (s. eq. 3.11, p. 29) in one matrix $\boldsymbol{\Omega}^{N_\iota \times N_{k_0}}$ the affinity score matrix is defined as

$$\boldsymbol{\mathcal{A}}^{N_\iota \times N_{l_0}} := \left(\frac{\max\left(\boldsymbol{\Omega}^{N_\iota \times N_{k_0}} - \boldsymbol{\vartheta}^{N_\iota \times N_{k_0}}, 0\right)}{\boldsymbol{\Omega}_{max(k_0)}^{N_\iota \times N_{k_0}} - \boldsymbol{\vartheta}^{N_\iota \times N_{k_0}}}\right) \cdot \underbrace{\begin{pmatrix} m_{1\,1} & \cdots & m_{1\,l_0} & \cdots & m_{1\,N_{l_0}} \\ \vdots & \ddots & & & \vdots \\ m_{k_0\,1} & & m_{k_0\,l_0} & & m_{k_0\,N_{l_0}} \\ \vdots & & & \ddots & \vdots \\ m_{N_{k_0}\,1} & \cdots & m_{N_{k_0}\,l_0} & \cdots & m_{N_{k_0}\,N_{l_0}} \end{pmatrix}}_{=:\boldsymbol{\mathcal{M}}^{N_{k_0} \times N_{l_0}}} . \qquad (4.6)$$

This matrix summarizes the affinity scores for all $N_\iota = 73$ sample layers and $N_{l_0} = 2$ numbers of biomes for Lake Kinneret . The assignment of each taxon to one biome is described by the biome assignment matrix $\boldsymbol{\mathcal{M}}^{N_{k_0} \times N_{l_0}}$. Each column vector $\vec{m}_{l_0}$ (in eq. 4.6 marked in green) describes if the regarded taxon k_0 belongs to this biome l_0 ($m_{k_0 l_0} = 1$) or not ($m_{k_0 l_0} = 0$). The corresponding values for Lake Kinneret are shown in the Appendix in Table B.1 (p. 145) and in Table C.1 for Ein Gedi (p. 201).

The $\boldsymbol{\Omega}_{max(k_0)}^{N_\iota \times N_{k_0}}$ denotes the maximum value in each column vector $\vec{\widetilde{\omega}}_{k_0}$. Together with the maximum function in the numerator this transfers the relative abundances $\widetilde{\omega}_{\iota\,k_0}$ to deviation from the maximum value of each taxon.

$\boldsymbol{\vartheta}^{N_\iota \times N_{k_0}}$ is the taxa threshold matrix which summarizes all taxon specific threshold values ϑ_{k_0}. Due to the fact that the threshold is defined for each taxon it is a matrix with column-wise constant values. For the Lake Kinneret pollen data different settings were tested for evaluating the influence of this factor: $\forall k_0 \vartheta_{k_0} = 0$, $\vartheta_{k_0} = Q_{k_0}\left(\vec{\widetilde{\omega}}_{k_0}\right) = 33\,\%$ the 33 % quantile of $\vec{\widetilde{\omega}}_{k_0}$ and a taxon specific individual definition. The last mentioned ϑ_{k_0} estimation approach is the same as described in chapter 3.2.1. The approach has still the same advantages and disadvantages but it also allows the interpretation why it is introduced in equation 4.6: The subtraction is done for decreasing the influence of clusters in each taxon profile which are regarded as background noise (Litt *et al.* 2012).

The required figures for the taxon specific individual threshold definitions, depth profile of $\vec{\widetilde{\omega}}_{k_0}$ and ECDF, can be found in appendix B.1.2 on page 147. The corresponding results for ϑ_{k_0} are summarized in Table B.1 (p. 145).

The biome probability $W_{B_{l_0}}$ for the pollen sample found in depth d_ι is now given, after a normalisation for each layer

$$W_{B_{l_0}} := \frac{a_{\iota\, l_0}}{\sum\limits_{l_0=1}^{N_{l_0}} a_{\iota\, l_0}}. \tag{4.7}$$

This normalisation ensures that the requirement for a probability is fulfilled. Together with equation 4.6 this definition of the biome probability $W_{B_{l_0}}$ is an alternative to the equation used in Litt *et al.* (2012).

The palaeoclimate reconstruction for the Lake Kinneret sediment core includes only two biome types: the Mediterranean l or resp. $l_0 = 1$ and Irano-Turanian l or resp. $l_0 = 2$. The reason is that only one taxon is included in the palaeoclimate reconstruction which is assigned to the Saharo-Arabian biome (l or resp. $l_0 = 3$) at the Ein-Gedi reconstruction. It could also be assigned to the Irano-Turanian biome. Additionally the Lake Kinneret is not that focus point of the three Biomes like the dead sea.

Figure 4.1 shows the biome probability $W_{B_{l_0}}$ for these two biome types and summarizes the results for different threshold values ϑ_{k_0}. Additionally the result for the ratio trees & shrubs/grasses & herbs pollen and the local pollen assemblage zones (LPAZ) defined by Schiebel (2013) are shown. The area below the curves is the biome probability $W_{B_{l_0=1}}$ for the Mediterranean biome. Consequently the probability $W_{B_{l_0=2}}$ for the Irano-Turanian biome is the area between one and the curves. The influence of the different threshold compared $\vartheta_{k_0} = 0$ is an increasing of the probability for the Mediterranean biome in the upper part of the record

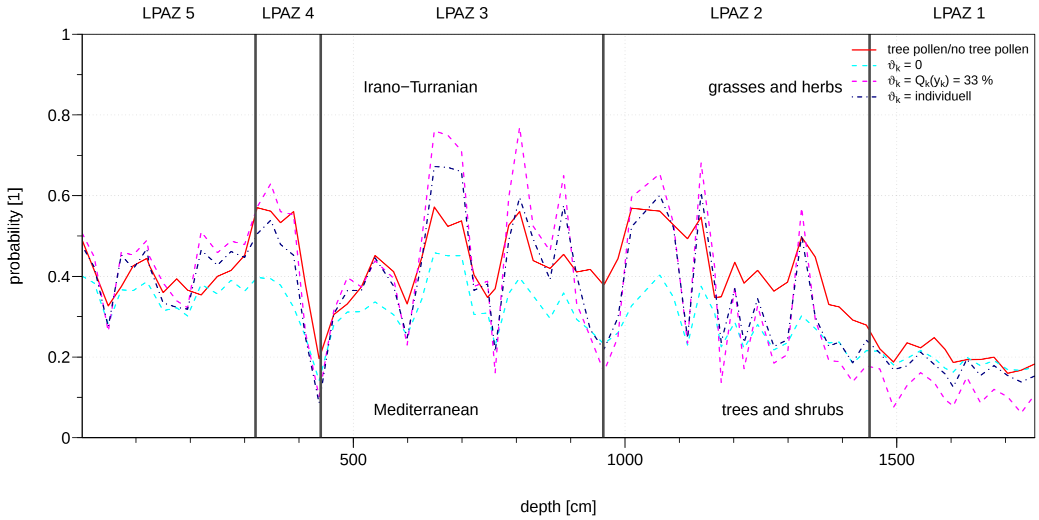


Figure 4.1.: Depth profile of the biome probability $W_{B_{l_0}}$ for Lake Kinneret estimated with different threshold values ϑ_{k_0}. The area under the curves represents the Mediterranean biome probability $W_{B_{l_0=1}}$ and above the curves to one represents the Irano-Turanian biome probability $W_{B_{l_0=2}}$. The vertical lines separate the local pollen assemblage zones (LPAZ) of Lake Kinneret defined in Schiebel (2013, chap. 5.1.3, Tab. 5.1 and Fig. 5.1). The red curve is the result from Schiebel (2013) for the ratio of trees & shrubs/grasses & herbs pollen.

(LPAZ 5 till the end of LPAZ 3 at about 850 cm) and a decrease in lower part (from the end of LPAZ 2 till the end of LPAZ 1). In this regard LPAZ 2 is the transition zone. It is clearly deducible that the trend of the biome probability varies in the same way as the ratio trees & shrubs/grasses & herbs. This fits very well with the fact that the Irano-Turanian biome is a steppe biome type dominated by grasses & herbs (Schiebel 2013). Also identifiable is that for almost all LPAZ (LPAZ 5 till the end of LPAZ 2 at about 1250 cm) there is no preferred biome type, since the biome probability always varies around 0.5. Only in the lower part of the sediment core the Irano-Turanian biome is preferred.

4.2.3. $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}\middle|\vec{c}_0\right)$ Biome Likelihood - Biome Transfer Function

The biome likelihood or biome transfer function $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\,\vec{c}_0\right)$ connects the modern climate data and the modern biome distribution maps (s. Fig. 2.11, p. 15). The transfer functions presented in chapter 3.2.2 describe the same for a single taxon but in contrast to taxa there is no information available about the absence of a biome. Therefore Litt *et al.* (2012) use a quadratic discriminant analysis (QDA) for estimating $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\,\vec{c}_0\right)$ in the BBM for the palaeoclimate reconstruction based on the Ein Gedi sediment core.

A detailed description of QDAs can be found in Venables, Ripley (2002) and is here only shortly described. The general purpose of a discriminant analysis is to assign objects into groups, determining the boundary between them. The considered boundary here is the boundary between biomes in the climate phase space. For this aim the Bayes theorem is applied on the biome likelihood:

$$\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_l = l|\,\vec{c}\right) := \frac{f_{\vec{C}|\mathcal{B}_l}\left(\vec{c}\,\middle|\, b_l = l\right)\cdot \pi_{\mathcal{B}_l}\left(b_l = l\right)}{m_{\vec{C}}\left(\vec{c}\right)}. \tag{4.8}$$

Thereby $b_l = 1$ reflects the aforementioned fact that only the presence information of each biome is available. The prior knowledge is defined as

$$\pi_{\mathcal{B}_l}\left(b_l = l\right) := \frac{N_{B_l\,obs}}{N_{obs}} = \frac{N_{B_l\,obs}}{\sum_{l=1}^{N_l} N_{B_l\,obs}} =: RF_{B_l}. \tag{4.9}$$

$N_{B_l\,obs}$ is the number of all grid points $i := \lambda\varphi$ in the biome area maps at which presence information of biome type B_l is available. N_{obs} is the number of grid points i with presence information of any biome type B_l. In other words the prior knowledge is the relative frequency RF_{B_l} of biome type B_l referred to the number of biome area grid points N_{obs}.

As last step it is necessary to define $f_{\vec{C}|\mathcal{B}_l}\left(\vec{c}\,\middle|\, b_l = 1\right)$, the so called group pdf. In the present case a multivariate normal distribution $\mathcal{N}_{MV}\left(\vec{c}, \vec{\mu}_l, \mathbf{\Sigma}_l\right)$ is assumed . The exact values of the covariance matrix $\mathbf{\Sigma}_l$ and the mean $\vec{\mu}_l$ of each class l are not known.

If $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_l = l|\vec{c}\right)$ is maximised the best or most likely boundary between the groups is estimated. In other words equation 4.8 defines the maximum likelihood function for the classification problem. As mentioned at the beginning of chapter 3.2.2 it is also possible to apply the log-likelihood function. In the context of classification problems with a multivariate normal distribution as class pdf this function is called the discriminant function:

$$d_{\mathcal{B}_l}\left(\vec{c}\right) = \left(\vec{c} - \vec{\mu}_l\right)^T \mathbf{\Sigma}_l^{-1}\left(\vec{c} - \vec{\mu}_l\right) + \ln|\mathbf{\Sigma}_l| - 2\ln\left(\pi_{\mathcal{B}_l}\left(b_l = l\right)\right). \tag{4.10}$$

It is derived from equation 4.8 by taking the natural logarithm, inserting $\mathcal{N}_{MV}\left(\vec{c}, \vec{\mu}_l, \mathbf{\Sigma}_l\right)$ and omitting all terms not depending on the parameters related with biome type l. The estimation of $\vec{\mu}_l$, $\mathbf{\Sigma}_l$ and the associated coefficients $\vec{\alpha}_{opt}$ which describe the boundary, is performed with the statistical software R[1]. For the input data of the estimation process different configurations, which are described in the following, are tested and verificated.

On the other hand the definition of the climate state vector $\vec{C}$ resp. the realisation $\vec{c}$ compared to Lake Prespa (s. eq. 3.1, p. 25) is changed here. First of all different dimensions are considered, in detail the climate state vector is defined as

$$\vec{C} := (C_1, C_2, C_3)^T \quad \text{with the realisations} \quad \vec{c} \in \mathbb{R}^3 \text{ or} \tag{4.11a}$$

$$\vec{C} := (C_1, C_2)^T \quad \text{with the realisations} \quad \vec{c} \in \mathbb{R}^2. \tag{4.11b}$$

Secondly mean winter temperature (T_{DJF}) and mean summer temperature (T_{JJA}) are used (s. chap. 2.1.1, p. 7). Thirdly different coordinate combinations are used. In equation 4.11 this fact is already regarded by the coordinate definition with C_i. In detail ten two-dimensional and four three-dimensional combinations out of the near surface climate variables T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} and the climate variables at 850 hPa, the middle troposphere, T_{DJF}^{850hpa} and T_{JJA}^{850hpa} are tested (s. apx. B.3 Tab. B.4 for detailed definition of the combinations). The associated climate datasets are described in chapter 2.1. Fourthly three different databases for the near surface climate variables are tested: CRU TS 3.1, E-OBS and a mixture of CRU TS 3.1 and E-OBS for the Jordan Valley palaeoclimate reconstruction.

On the other hand, as mentioned in chapter 2.2.1, four different configurations for the biome areas are used. The first of these so named "*biome setups*" are the biome areas B_1, B_2 and B_3 as defined in Litt *et al.* (2012) and presented in Figure 2.11a (biome setup 1). The three remaining base on Figure 2.11b, at which the first one does not change the definition of B_1, B_2 and B_3 (biome setup 2). For biome setup 3 the Mediterranean biome area is extended: $B_1 := B_1 + B_{1-a} + B_{1-b}$ (s. Fig. B.15). In addition for last biome setup 4 the Irano-Turanian biome is extended: $B_2 := B_2 + B_{2-a}$ (s. Fig. B.16).

Together with the aforementioned different climate configurations each possible setup is

---

[1] in more detail with the package *MASS* for R version 3.0.1, s. Venables, Ripley (2002) for more details

abbreviated with one setup number (s.no.). Table B.2 in appendix B.3 summarizes all defined setup numbers. In order to evaluate all setups a tool known in probabilistic forecast verification is applied: the Brier skill score defined for each biome B_l. First the related Brier score has to be defined (after Brier 1950):

$$BS_{B_l} = \frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} \left[\mathbb{P}_{\mathcal{B}_l|\vec{C}} \left(\mathscr{b}_l = l | \vec{c}_i \right) - E_i\left(l\right) \right]^2 . \tag{4.12}$$

The biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}} \left(\mathscr{b}_l = l | \vec{c}_i \right)$ for biome B_l is then evaluated at all grid points i, defined above with the given climate state realisation $\vec{c}_i$. The detailed configuration of $\vec{c}_i$ for the dimensions, the combinations of climate variables and the climate dataset for the estimation is predetermined by the just mentioned setup number. $E_i\left(l\right)$ is a function that returns one if biome B_l is present at grid point i and zero otherwise. It is important to emphasise that the Brier score is zero if the result of the biome likelihood matches exactly with the biome area.

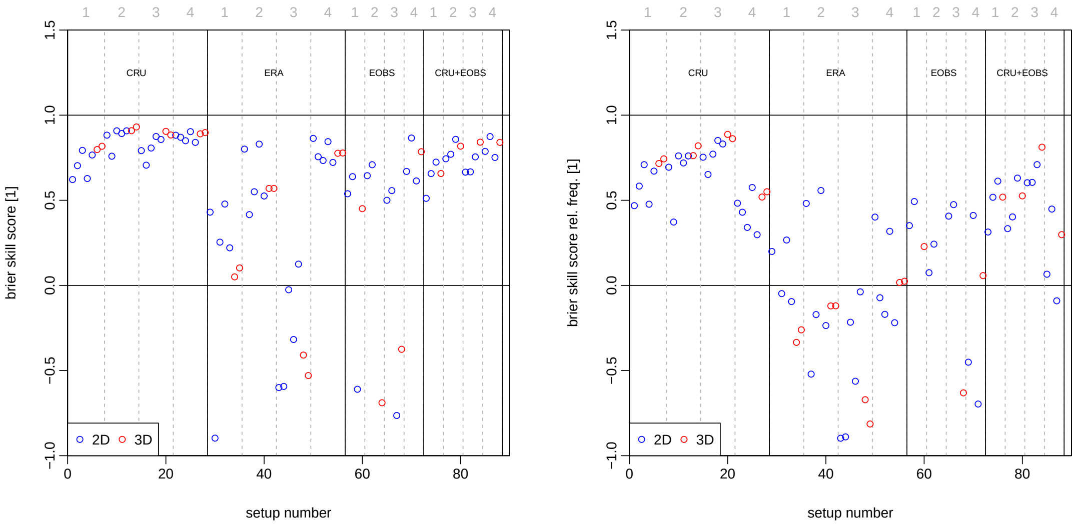
For the calculation of the Brier skill score BSS a reference score BS_{ref} is required. Here in this thesis two different reference scores are defined by replacing the Biome likelihood in equation 4.12 with the above defined relative frequency RF_{B_l} and a reference probability of $ZV = \frac{1}{N_l}$, so that the two Brier skill scores are defined by:

$$BSS_{B_l}^{RF} = 1 - \frac{BS_{B_l}}{BS_{ref}^{RF_{B_l}}}, \tag{4.13a}$$

$$BSS_{B_l}^{ZV} = 1 - \frac{BS_{B_l}}{BS_{ref}^{ZV}}. \tag{4.13b}$$

Both scores converge to one if BS_{B_l} is similar to the reference score (Stefanova, Krishnamurti 2002; Stolzenberger 2011). For the palaeoclimate reconstruction of Lake Kinneret and Birkat-Ram (s. chap. 5) only two biome types are available. As a result the Brier skill scores for B_1 and B_2 are identical, since the biome likelihood $\mathbb{P}_{\mathcal{B}_1|\vec{C}} \left(\mathscr{b}_1 = 1 | \vec{c} \right)$ is substitutable with $\mathbb{P}_{\mathcal{B}_2|\vec{C}} \left(\mathscr{b}_2 = 2 | \vec{c} \right)$ due to the fact that both together sum up to one.

The aim is to compare different setups. Therefore all required climate variables for the verification are calculated with the same dataset. Figure 4.2 shows the result for the CRU TS 3.1 dataset and Figure 4.3 for the ERA-Interim dataset, both for a total of two biomes. The respective Figures B.17 and B.18 for a total of three biomes are shown in appendix B.2. For the verification with the CRU TS 3.1 dataset the temperatures T_{DJF}^{850hpa} and T_{JJA}^{850hpa} are determined with a simple height correction, based on the temperature gradient of the standard atmosphere ($6.5 \times 10^{-3} \frac{K}{m}$), the mean height of the 850 hPa level (1500 m) and the CRU TS 3.1 orography as already applied in Stolzenberger (2011), Simonis (2009) or Gebhardt (2003). For the ERA-Interim verification this is done the other way round, for the near surface temperatures. It is important to emphasise that every setup number in Table B.2, B.3, B.4 and B.5 in appendix B.3



(a) $BSS_{B_1}^{ZV}$ for a total of two biomes

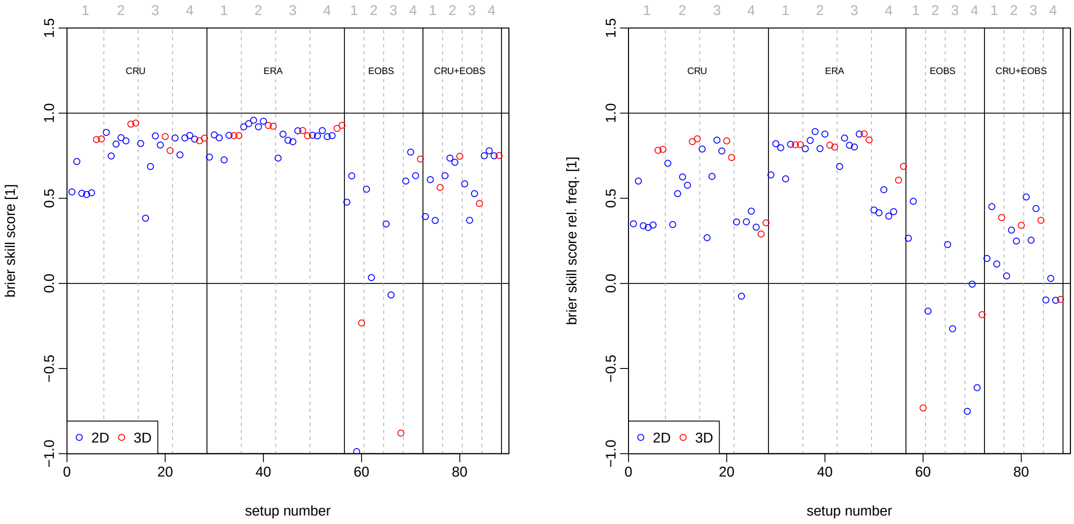
(b) $BSS_{B_1}^{RF}$ for a total of two biomes

Figure 4.2.: Verification of the biome likelihood $\mathbb{P}_{\mathcal{B}_1|\vec{C}}\left(\mathscr{b}_1 = 1|\vec{c}\right)$ with Brier skill score $BSS_{B_1}^{ZV}$ and $BSS_{B_1}^{RF}$ for a total of two biomes. The parameters which determine the biome likelihood are estimated as given by the setup number shown on the x-axis. The setup numbers are defined in Table B.2. The climate input data for the verification is always the CRU TS 3.1 dataset. The dotted grey vertical lines separate the different "*biome setups*" as specified on top of the Figures (grey numbers). The black vertical lines separate the different climate databases as input data in the estimation process.

defines the same setup aside from the total number of biomes.

The distribution of the data points in the figures for the $BSS_{B_l}^{RF}$ is similar to those of the $BSS_{B_l}^{ZV}$. This follows from the similarity in the definition of the reference scores, since the only difference in the definition is that BS_{ref}^{ZV} is always greater than $BS_{ref}^{RF_{B_l}}$. As a result the clusters for $BSS_{B_l}^{ZV}$ are more compact an shifted to slightly higher values. Consequently $BSS_{B_l}^{ZV}$ is more convenient for a general overview and $BSS_{B_l}^{RF}$ for a more detailed consideration.

Summarized there are some general results for a total of two and three biomes: One is that the highest values for the Brier skill scores are obtained with the dataset that is applied in the estimation process. This is an important result which maintains that there is no error in the estimation of the scores since by construction of the Brier skill scores this is the expected result. Another result is, that if all calculated $BSS_{B_l}^{\ddot{\cdot}}$ for $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l = l|\vec{c}\right)$ estimated with surface climate values (all s.no. except 29 - 56) are compared, the biome likelihood estimated with the E-OBS-dataset alone (s.no. 57 - 72) has the highest variance. This is probably a direct consequence of the considered geographical region, since it is situated in the edge of the E-OBS dataset area and therefore boundary effects in the result cannot be excluded completely. In contrast the mixture dataset (s.no. 73 - 88) varies in the same order as the CRU TS 3.1-dataset (s.no. 1 - 28). The variance of $BSS_{B_l}^{\ddot{\cdot}}$ values for $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l = l|\vec{c}\right)$ estimated with ERA-Interim-dataset and



(a) $BSS_{B_1}^{ZV}$ for a total of two biomes

(b) $BSS_{B_1}^{RF}$ for a total of two biomes

Figure 4.3.: Verification of the biome likelihood $\mathbb{P}_{\mathcal{B}_1|\vec{C}}\left(b_1 = 1|\vec{c}\right)$ with Brier skill score $BSS_{B_1}^{ZV}$ and $BSS_{B_1}^{RF}$ for a total of two biomes. The parameters which determine the biome likelihood are estimated as given by the setup number shown on the x-axis. The setup numbers are defined in Table B.3. The climate input data for the verification is always the ERA-Interim dataset. The dotted grey vertical lines separate the different "*biome setups*" as specified on top of the Figures (grey numbers). The black vertical lines separate the different climate databases as input data in the estimation process.

verificated on CRU TS 3.1-dataset is also high, but this is not the case the other way round. It is also identifiable that the $BSS_{\ddot{B}_l}$ for setups based on a three-dimensional climate state vector (red dots in the figures)) range at the top for the most biome setup. The exceptions are the verification of the E-OBS-dataset (s.no. 57 - 72) and the s.no. 29 - 56 for a total of three biomes verificated on the CRU TS 3.1-dataset (s. Fig. B.17).

The biome likelihood estimated only with climate data from the E-OBS-dataset is not applied in any palaeoclimate reconstruction presented in this work. Therefore the s.no. 57 - 72 are not discussed in detail. If a biome prediction is required alone the results presented in Figure 4.3 and B.18 show an interesting fact: It is possible to apply the ERA-Interim-dataset as input even if the likelihood fit is estimated with the CRU TS 3.1-dataset (s.no. 1 - 28) and the result is not much worse than the fit based on the ERA-Interim-dataset s.no. 29 - 56. Also not clearly detectable is the "*best*" biome setup for a total of two biomes. For biome setup two or three show in general higher $BSS_{\ddot{B}_l}$ values than one or four in Figure 4.2 and 4.3. However slightly higher values can be found in biome setup number three, which is supposed to be a result of the approximately same order of magnitude for the number of biome points ($N_{B_1 obs} = 61$ and $N_{B_2 obs} = 141$).

For a total of three biomes, as required for the Ein Gedi palaeoclimate reconstruction, it is necessary and possible to regard the three different $BSS_{\ddot{B}_l}$ in the Figures B.17 and B.18. The majority (90 %) of the data points in the plots for biome type B_1 show higher $BSS_{\ddot{B}_l}$ values than those for the other two biome types. The Brier skill scores for biome type B_2 calculated with the ERA-Interim-dataset show no substantial lower values than those calculated with the CRU TS 3.1-dataset. Also the pattern of the data points for every biome type is similar.

The conclusion is that the new digitalisation of the biome maps has a slightly positive effect and consequently one of the new biome setups (two till four) could and should be applied. Also a three-dimensional climate state vector should be applied in the reconstruction. It should be considered that the aim in palaeoclimate reconstructions is not to predict the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l = l|\vec{c}\right)$ alone. Therefore the effect of the biome setup should be also examined in connection with palaeoclimate reconstructions.

4.2.4. $m_{\mathcal{B}_l}\left(\mathscr{b}_{l_0}\right)$ Marginal Probability

The estimation of the marginal probability $m_{\mathcal{B}_l}\left(\mathscr{b}_{l_0}\right)$ is the same as in Litt *et al.* (2012). It is an integration over the hole climate space $\mathcal{K}_0$ over the product of the biome likelihood and the climate prior:

$$m_{\mathcal{B}_l}\left(\mathscr{b}_{l_0}\right) := \int\limits_{\mathcal{K}_0} \mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}|\vec{c}_0\right)\pi_{\vec{C}}\left(\vec{c}_0\right)d\vec{c}_0. \tag{4.14}$$

4.3. Lake Kinneret - Palaeoclimate Reconstruction Result

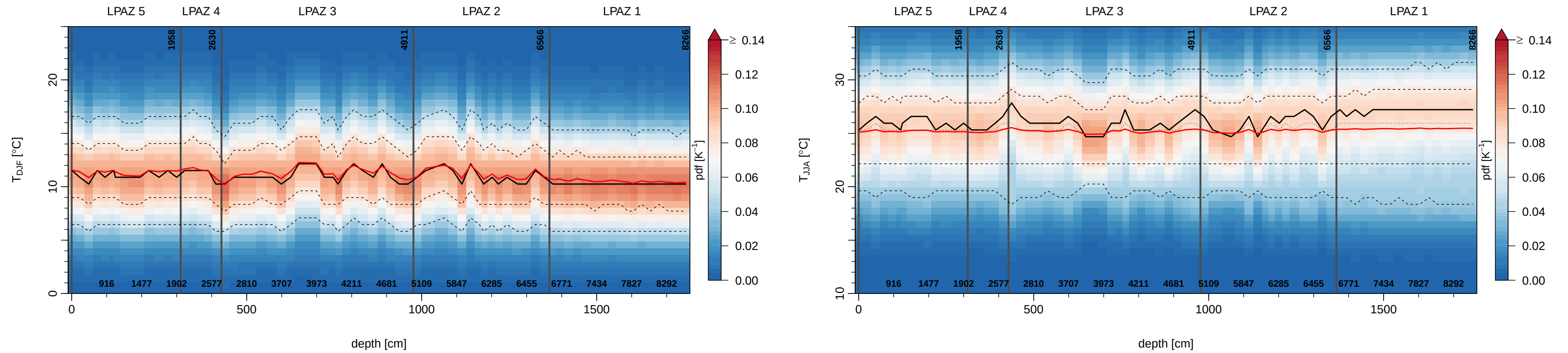
All figures (Fig. 4.4, 4.6, 4.7 and 4.8) in this chapter show the estimated result of the palaeoclimate reconstruction for Lake Kinneret in terms of a depth profile of the one-dimensional marginal pdfs. The term marginal is already defined in chapter 3.2.2. The design of the figures for the posterior pdf for the Lake Kinneret palaeoclimate reconstruction is the same as that of the Lake Prespa figures, which is already described at the beginning of chapter 3.3 (s. p. 40) and repeated in the caption of Figure 4.4. The palaeoclimate reconstruction presented in Figure 4.4a, 4.4b and 4.4c bases on the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l = l|\vec{c}\right)$ whose setup for the estimation is defined by s.no. 21[2] and the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$. Additionally Figure 4.4d shows the marginal pdf for CWD_{ANN} whose setup[3] differs only in the third component of the climate state vector. The corresponding figures for T_{DJF} and T_{JJA} are not shown since the result is equal.

All pdfs in Figure 4.4 show no significant signal for a period with climatic change. The 10 % and 90 % quantile show only a maximum variation of less than 2 K for the temperatures and resp. 120 $\frac{\text{mm}}{\text{year}}$ for P_{ANN} or CWD_{ANN}. The conclusion of no climatic changes matches with the

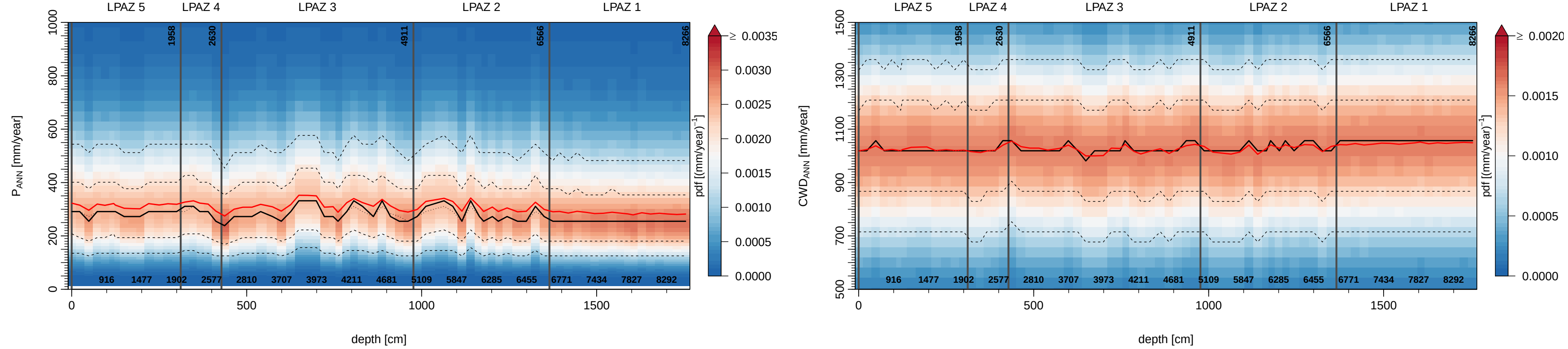
---

[2] s.no. 21 is equal to: biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset

[3] in detail $\vartheta_{k_0} = individual$ and s.no. 20 which equal to biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, CWD_{ANN})^T$ and CRU TS 3.1 as estimation dataset



(a) s.no. 21 **(b)** s.no. 21



(c) s.no. 21 **(d)** s.no. 20

Figure 4.4.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} for the Lake Kinneret palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$. (a), (b) and (c) base on s.no. 21 (biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset) and (d) on s.no. 20 which is equal to s.no. 21 and differs only in the definition the climate state vector ($\vec{C} = (T_{DJF}, T_{JJA}, CWD_{ANN})^T$). The pdfs are shown with a coloured scale, ranging from blue for low values to red for high values. The solid black line marks the mode, which is the highest value of each pdf in each sample layer. The dashed black lines mark the 10 %, 25 %, 75 % and the 90 % quantile. The dotted line is the median, which is equal to the 50 % quantile. The red line is the weighted arithmetic mean for each layer, which is interpreted as the numerical calculated expectation value. The weights are the values of the marginal pdf. The bold black numbers transfer the marked ticks of the depth axis in y. cal BP (s. chap. 2.3, p. 17) according to the age-to-depth model established by Schiebel (2013) and mentioned in chapter 2.3.2.1. The LPAZ are also shown and labelled with the according age.

result for the biome probability $W_{B_{l_0}}$ (s. chap. 4.2.2), since no biome type is clearly preferred up to a depth of approximate $1250\,\mathrm{cm} \equiv 6271 \pm^{84}_{99}\,\mathrm{y.\,cal\,BP}$. On the contrary the marginal posterior pdfs show no evidence for a climate change in the lower part ($\gtrapprox 1250\,\mathrm{cm}$), only less variation is detectable. If the definition of the BBM is considered (s. eq. 4.5) it becomes obvious that the weighted biome likelihoods $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0} = l|\vec{c}\right)$ multiplied with the climate prior $\pi_{\vec{C}}\left(\vec{c}_0\right)$ play an important role and have to be regarded in detail:

$$\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathscr{b}_{l_0}, \vec{c}_0\right) := \mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0} = l|\vec{c}\right)\pi_{\vec{C}}\left(\vec{c}_0\right). \tag{4.15}$$

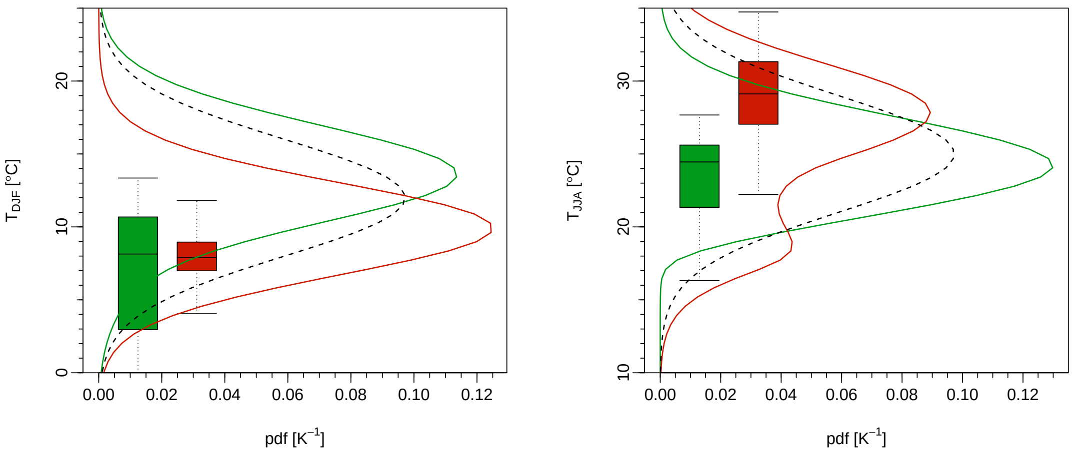
The marginal distributions of this probability are presented in Figure 4.5 dedicated to the reconstruction presented in Figure 4.4. The marginal prior pdf is shown with a dashed line, the coloured green line represents the Mediterranean biome B_1 and the red one the Irano-Turanian biome B_2. The box plots are the visualisation of the input dataset for the estimation of $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l = l|\vec{c}\right)$ with the same mapping of the colours. They are estimated with the standard function of the statistical software R in which the five vertical lines are estimated as the quantiles which are very close to the $0, 25, 50, 75, 100\,\%$ quantiles[4].

The comparison of the box plots, i. e. of the input data, shows that there is no distinction possible between B_1 and B_2 based alone on T_{DJF} values. This is not the case for T_{JJA}, P_{ANN} and CWD_{ANN}: at least $50\,\%$ of B_1 values illustrate a clear difference to those of B_2. This changes for the marginal distributions of the prior-likelihood product, since the multivariate distributions also include covariance structures. All four figures show a great influence of the prior on the resulting mixture product. This is no unexpected result due to the fact that a BHM is applied and, as aforementioned, the result has to be regarded with respect to the prior. Another possible explantation for this high sensitivity towards the prior is that the likelihoods of the observations (here the climate input data) do not contain enough information to estimate roubst values for $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l = l|\vec{c}\right)$. For T_{DJF} and P_{ANN} the mode of the prior is exactly the intersection point of $\mathbb{P}_{\mathcal{B}_1,\vec{C}}\left(\mathscr{b}_{1_0}, \vec{c}_0\right)$ and $\mathbb{P}_{\mathcal{B}_2,\vec{C}}\left(\mathscr{b}_{2_0}, \vec{c}_0\right)$ (of B_1 and B_2). This is not true for T_{JJA} and CWD_{ANN}: the mode of the prior is almost identical with the mode of $\mathbb{P}_{\mathcal{B}_1,\vec{C}}\left(\mathscr{b}_{1_0}, \vec{c}_0\right)$. Additionally almost the complete area under $\pi_{\vec{C}}\left(\vec{c}_0\right)$ is identical to B_1.

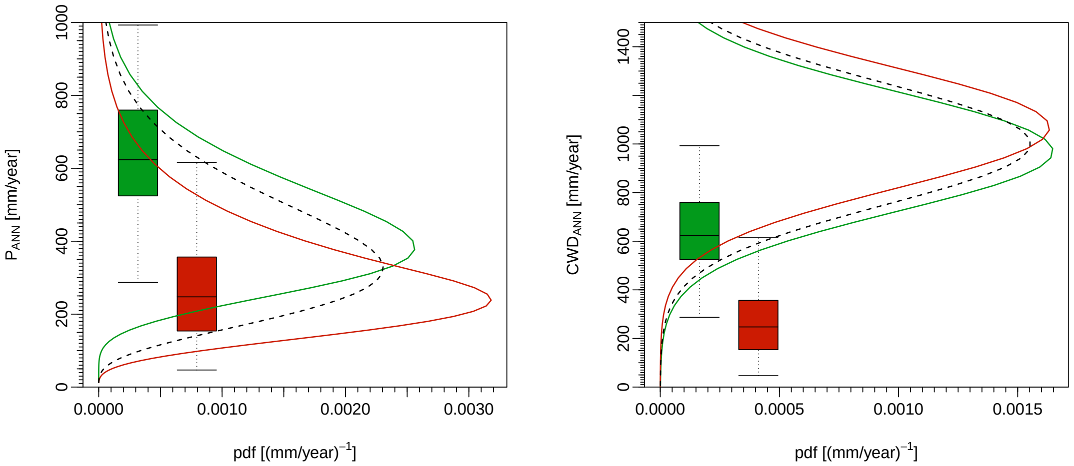
Similar difficulties could be identified in the palaeoclimate reconstruction for 850 hPa temperatures. Figure 4.6 presents the marginal pdf for T_{DJF}^{850hpa} (4.6a) and P_{ANN} (4.6b) with $\vartheta_{k_0} = \mathit{individual}$ and biome setup 4. This biome setup is selected by the reason that it reaches the highest Brier skill scores values for the verification based on CRU TS 3.1 data (Figure 4.2). The comparison of this palaeoclimate reconstruction with the one discussed in the last paragraphs is only reasonable if the aim is to reconstruct surface temperatures with transfer functions estimated with 850 hPa temperatures. This makes the selection of biome setup 4

---

[4] for more detail s. R-help for *boxplot.stats*



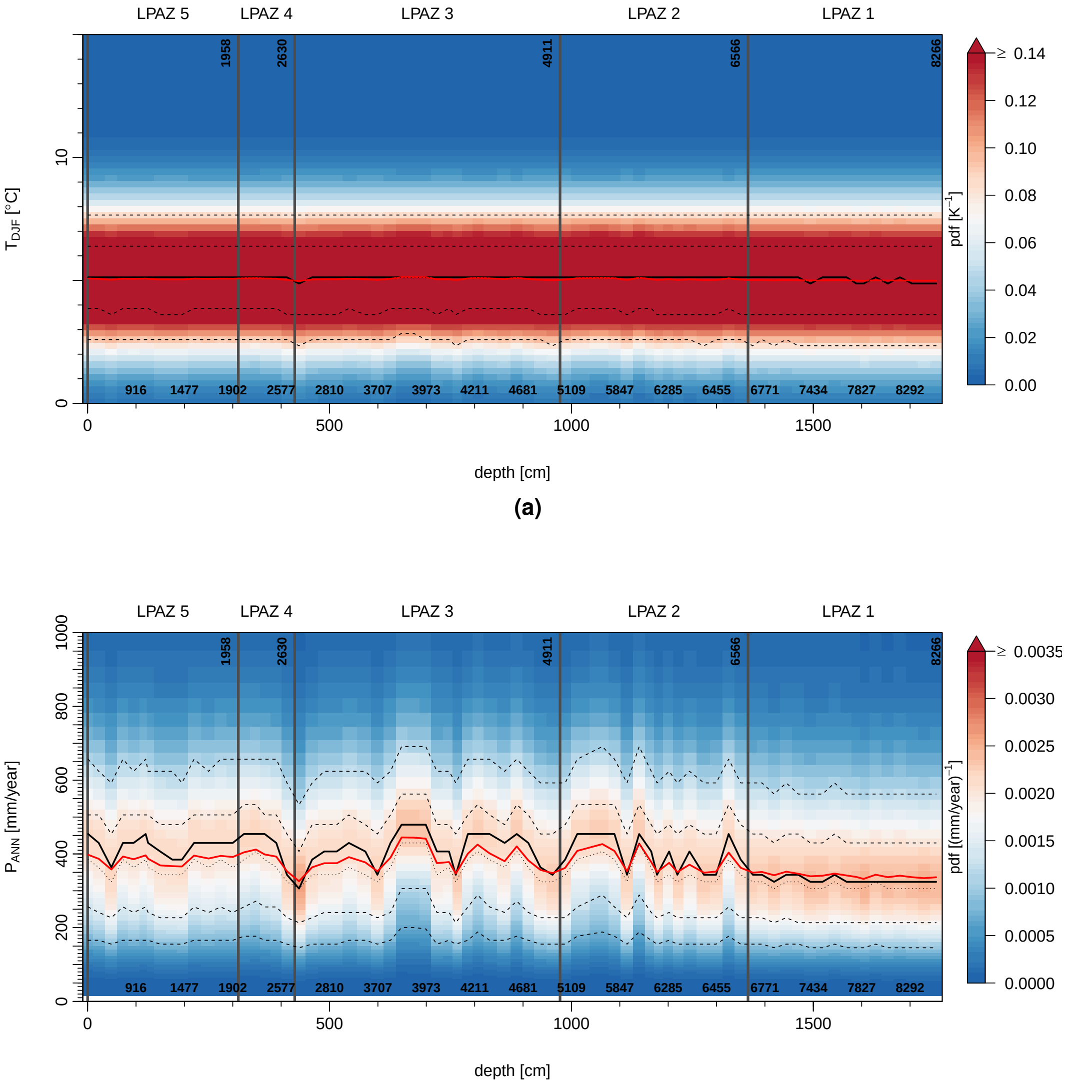
(a) dedicated to Fig. 4.4a **(b)** dedicated to Fig. 4.4b



(c) dedicated to Fig. 4.4c **(d)** dedicated to Fig. 4.4d

Figure 4.5.: Marginal distributions of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathcal{b}_{l_0},\vec{c}_0\right)$ as defined in eq. 4.15 (solid coloured lines) and the climate prior $\pi_{\vec{C}}\left(\vec{c}_0\right)$ (dashed black line) for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} dedicated to the palaeoclimate reconstruction presented in Figure 4.4. The box plots are the visualisation of the input dataset for the estimation of the biome likelihoods (biome setup 3). Green colours represent the Mediterranean biome B_1 and red colours the Irano-Turanian biome B_2.

reasonable. The surface temperatures are obtained by the application of the temperature correction described in chapter 4.2.3. If the aim is to reconstruct 850 hPa temperatures it is better to select biome setup 3 (Fig. 4.3). In Figure 4.6a there is almost no variability detectable. This is also the case for the marginal pdf for T_{JJA} (only shown in Figure B.19) but not so



(a)

(b)

Figure 4.6.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}^{850hpa} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$ and s.no. 56 ((biome setup 4, $\vec{C} = \left(T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}\right)^T$ and ERA-Interim as estimation dataset for the 850 hPa temperatures). The mapping of the figure is identical to Figure 4.4.

eye-catching like here. The absence of variability is not a result of the narrow prior pdf as Figure 4.7 proves for T_{DJF}^{850hpa} and $\sigma_{T_{DJF}^{850hpa}} = \sigma_{T_{JJA}^{850hpa}} = 4.0\,°\mathrm{C}$. The reconstructed marginal pdf for P_{ANN} in Figure 4.6b shows more variability compared to Figure 4.4c and a general shift of the complete marginal pdf-depth profile to higher values of P_{ANN}, which is a result of the connection in the transfer function to the lower temperature values at 850 hPa.

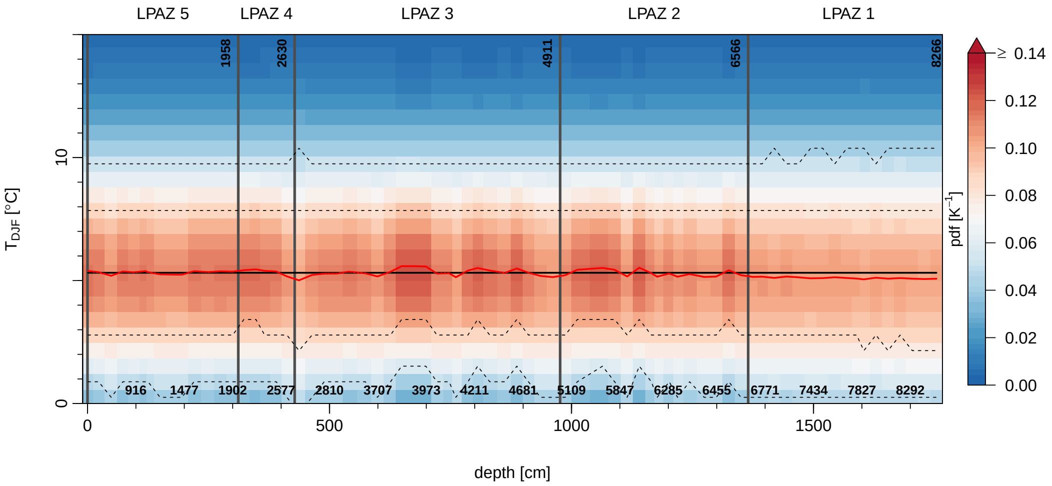
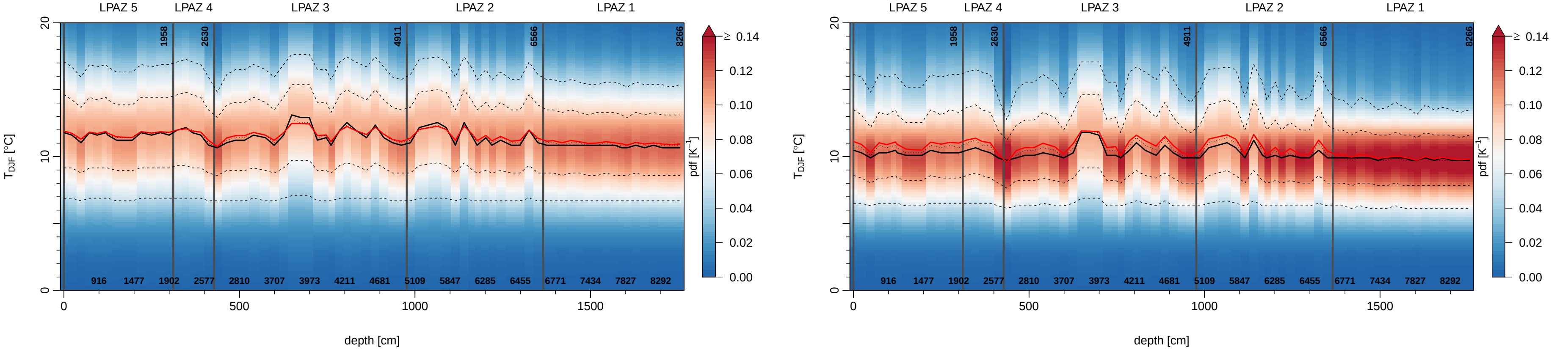


Figure 4.7.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}^{850hpa} for the Lake Kinneret palaeoclimate reconstruction with an alternative value of the standard deviation of the climate prior. In detail the same setup as in Figure 4.6 is applied. Only the standard deviation of $\pi_{\vec{C}}\left(\vec{c}_0\right)$ is modified to $\sigma_{T_{DJF}^{850hpa}} = \sigma_{T_{JJA}^{850hpa}} = 4.0\,°C$. The mapping of the figure is identical to Figure 4.4.

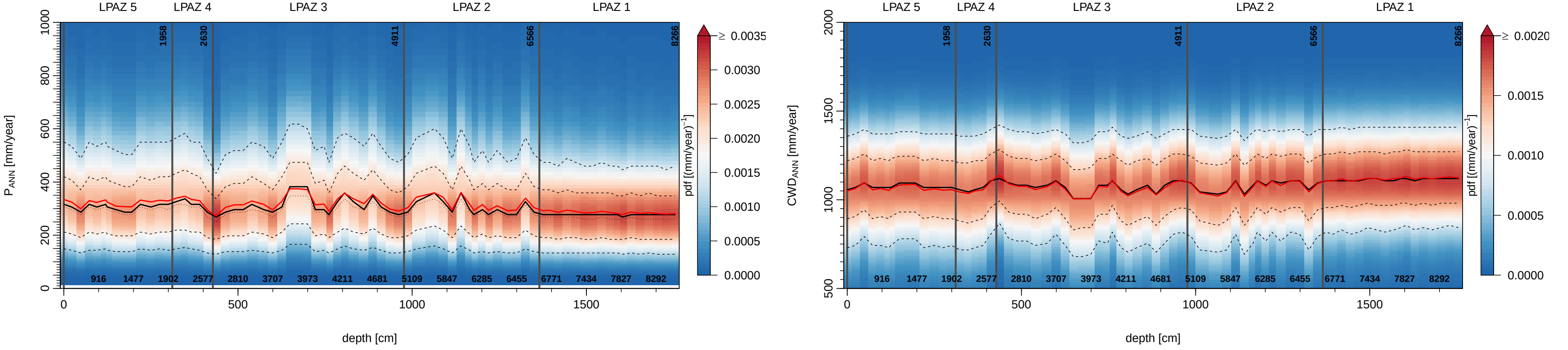
The analysis of the reconstructions estimated with different biome setups (Figures only shown in apx. B.4) reveals that the adjustment of this parameter has only little effects. Changing only the biome setup compared to the palaeoclimate reconstruction presented in Figure 4.4a - 4.4c shows that the greatest change appears compared to biome setup 1 (Fig. B.20): a shift of about 1 K of the marginal pdfs for the temperatures and an increase of variability in the 10 % and 25 % quantiles of the T_{JJA} profile. The comparison with biome setup 2 (Fig. B.21) shows only a slight increase of variability in the aforementioned quantiles of the T_{JJA} profile. In the last remaining biome setup, biome setup 4 (Fig. B.22), the complete marginal pdf depth profile for T_{DJF} shows a decrease of variability and the associated P_{ANN} pdf depth profile an increase.

As aforementioned the biome probability $W_{B_{l_0}}$ is estimated with different configurations. The influence on the posterior pdf is evaluated referred to the same palaeoclimate reconstruction as in the last paragraph (Figure 4.4a - 4.4c). The definition of $\vartheta_{k_0} = 0$ for all used taxa reduces the variability in all marginal pdfs in a rigorous manner, almost all variability is vanished (Fig. B.23). This is especially apparent for the marginal pdf for *Pann* due to the fact that all changes in the setup discussed so far do not have an effect in the same order of magnitude on the variability on the whole pdf. The third definition tested for the taxon thresholds ϑ_{k_0}, the definition with the 33 % quantile of $\vec{\vec{\omega}}_{k_0}$ ($\vartheta_{k_0} = Q_{k_0}\left(\vec{\vec{\omega}}_{k_0}\right) = 33\,\%$), leads in palaeoclimate reconstructions to scarcely discernible changes (Fig. B.24).

The analysis of the three-dimensional palaeoclimate reconstruction and the considerations in chapter 4.2.3 motivate to evaluate the contribution of the considered climate state vector.



(a) s.no. 15 **(b)** s.no. 18



(c) s.no. 15 **(d)** s.no. 18

Figure 4.8.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, P_{ANN} and CWD_{ANN} for the Lake Kinneret palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$. (a) and (c) base on s.no. 15 (biome setup 3, $\vec{C} = (T_{DJF}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset) and (b) and (d) on s.no. 18 which is equal to s.no. 15 and differs only in the definition of the climate state vector ($\vec{C} = (T_{DJF}, CWD_{ANN})^T$). The mapping of the figure is identical to Figure 4.4.

In more detail: How large is the effect in the palaeoclimate reconstruction if T_{JJA} is left out in the definition of $\vec{C}$? Figure 4.8 presents the result. The dimension reduction has almost no effect on the marginal pdf for T_{DJF}, only a slight reduction of the variability in the 10 % quantile and more variability in the whole pdf for CWD_{ANN} in the complete profile.

In summary, the lack of difference of these two biomes in climate phase space is obvious: They do not allow a palaeoclimate reconstruction with a climatic change based on the BBM with two biome types, where especially T_{JJA} and CWD_{ANN} do not differ. And with the whole discussion in background it leads to the conclusion that most of the variation especially in the palaeoclimate reconstruction in Figure 4.4 is an impact of $W_{B_{l_0}}$.

4.4. The Pollen-Ratio Model

As discussed in the last section the BBM with two biomes allows no detection of periods with climate change. The clear shift in the lower part of the depth profile of the biome probability $W_{B_{l_0}}$ (Fig. 4.1) indicates a climatic change and motivates to try another method for the palaeoclimate reconstruction: the pollen-ratio model (PRM) as presented in Ohlwein, Wahl (2012) and already mentioned in the introduction. Ohlwein, Wahl (2012) classify/describe the PRM as a "*special case of the modern analogue technique (MAT)*". In a nutshell the (traditional) MAT connects the composition of modern pollen spectra with the modern climate. As similar identified fossil pollen spectra in the sediment core allow to reconstruct the palaeoclimate with an inversion of the afore established connection applied on the fossil pollen spectra. Ohlwein, Wahl (2012) also explain that the MAT and consequently PRM fits into the concept of BHM in the same way as the BBM as presented in chapter 4.1. The PRM uses only the two different pollen counts of the complete available spectrum and applies "*the concept of response surfaces*" (Ohlwein, Wahl 2012). "*A response surface is essentially a smoothed representation of the pollen data associated with a gridded version of the modern (uni- or multi-variate) climate space used for analogue selection in the MAT*" (Ohlwein, Wahl 2012). These response surfaces are the pollen-climate transfer function and are estimated with a GLM which differs from the one described in chapter 3.2.2.

The main difference is that the response variable in the PRM are the aforementioned two different pollen counts p_i. These are in general pollen counts of two different groups i whereat the group definition is not fixed. This grouping step allows to assign specific taxa k into a group i, or in other words to include all pollen counts p_{ik} of important taxa by summation: $p_i = \sum_{k=1}^{N_k} p_{ik}$. With this notation it is not excluded that only two taxa are applied in the PRM, so that $p_k \equiv p_i$. In the BITM-GLM (s. chap. 3.2.2) the response variable is binary (s. p. 35) and hence corresponding distribution is Bernoulli. Based on the just described characteristics of p_i

the required distribution is the Binomial distribution

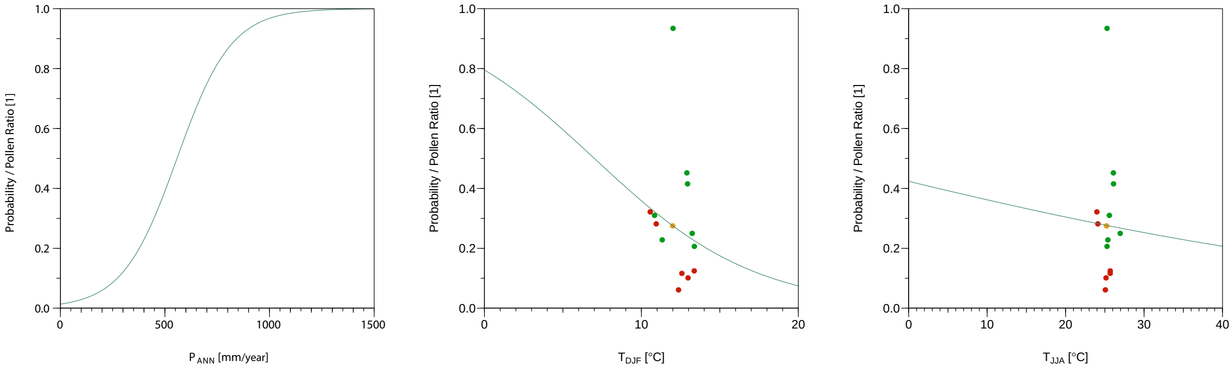
$$\mathrm{Bin}\left(p_i \middle| N_i, p\right) = \binom{N_i}{p_i} p^{p_i} \left(1-p\right)^{N_i - p_i} \tag{4.16a}$$

$$N_i = \sum_{i=1}^{N_i=2} \sum_{k=1}^{N_k} p_{ik}. \tag{4.16b}$$

whereat N_i is the total amount of pollen counts.

Compared to the BITM-GLM in chapter 3.2.2 two important points remain: Point one is ”*that it cannot be used simultaneously to reconstruct more than one climatic variable*” (Ohlwein, Wahl 2012). In the notation used here this means that the climate state vector is one-dimensional. Point two is that GLM with linear terms is applied. Consequently the design matrix $\boldsymbol{X}$ in the case considered here has two columns, which are column one and two of the design matrix defined in equation 3.25. Ohlwein, Wahl (2012) also explain that ”*the name pollen-ratio model originated form an alternative but equivalent description by using the ratio* $\frac{p_{i=1}}{p_{i=1}+p_{i=2}}$”.

Figure 4.9 presents the result of the PRM-GLM estimation based on the E-OBS gridded dataset (s. chap. 2.1.2) and the modern pollen spectra described in chapter 2.2.2. The data points are the pollen ratio of Artemisia and Quercus versus P_{ANN} or T_{DJF} or T_{JJA} on 14 locations (s. Tab. B.6, p. 198). The colour indicates if the location is situated in the Mediterranean, Irano-Turanian or Saharo-Arabian biome area defined in Figure 2.11b. The line is the estimated GLM result based on the assumption that N_i is 1000 on every location.



(a) (b) (c)

Figure 4.9.: PRM-GLM result for P_{ANN}, T_{DJF} and T_{JJA} versus the pollen ratio of Artemisia and Quercus. The data points are the pollen ratios on 14 locations in Israel (s. Tab. B.6, p. 198). The colour indicates if the location is situated in the Mediterranean (green), Irano-Turanian (red) or Saharo-Arabian biome area (orange) defined in Figure 2.11b. The line is the GLM estimation result.

This allows to convert the percentages in Table B.7 and B.8 to the required counts p_i. This is done to enable a first test if it is possible to apply the PRM for a palaeoclimate reconstruction for Lake Kinneret although no information about N_i is available. Beside the fact that data coverage is sparse, the PRM-GLM presented in Figure 4.9 shows one main problem which is independent from N_i: The distribution of the data points covers a to small interval in the climate phase space ($\approx [250;550]\,\frac{\text{mm}}{\text{year}}$, $\approx [10;14]_{T_{DJF}}$ °C and $\approx [23;28]_{T_{JJA}}$ °C) to allow a robust GLM estimation. Especially the absence of low precipitation values is a problem. Beside the datapoint for P_{ANN} for the Saharo-Arabian biome (location Be'er-Sheva) there is no clear boundary in the climate phase space to separate the biomes based on data presented in Figure 4.9. Additionally the distance of location Be'er-Sheva to the assigned E-OBS-grid point is nearly at the maximum distance allowed. Another pollen ratio does not change the distribution of the data points regarding the climate phase space. Therefore a PRM-GLM based on the data described in this section is not inverted and applied in a palaeoclimate reconstruction.

4.5. Ein Gedi - Palaeoclimate Reconstruction Result

The palaeoclimate reconstruction for Ein Gedi presented in Litt *et al.* (2012) is adapted to the requirements of the Jordan Valley reconstruction presented in chapter 6. These in the following discussed adaptations are :

1.) the update of the database for the near surface climate variables on CRU TS 3.1. The results in Litt *et al.* (2012) base on the CRU TS 1.0 dataset (New *et al.* 2000). The update on the mixture of CRU TS 3.1 and E-OBS for the Jordan Valley palaeoclimate reconstruction is discussed in chapter 6.3.

2.) the change of the prior distribution $\pi_{\vec{C}}\left(\vec{c}_0\right)$ for the climate state vector (s. chap. 4.2.1, Tab. 4.1 for the Jordan Valley values). In Litt *et al.* (2012) the palaeoclimate reconstruction is estimated on transferred P_{ANN} values, transferred with the inverse CDF method already mentioned in chapter 3.2.2. Also the prior in Litt *et al.* (2012) is defined on these values. Figure 4.10 shows the marginal pdf of the original Ein Gedi prior (Litt *et al.* 2012), the Jordan Valley prior and the original prior applied in the Birkat Ram palaeoclimate reconstruction presented in Neumann *et al.* (2007). Figure 4.10 shows that the difference between the Jordan Valley prior applied in the palaeoclimate reconstruction presented in this work is small. The area between them is quite small 5×10^{-07} and the difference between the maxima too ($13\,\frac{\text{mm}}{\text{year}}$). It is important to mention that the Jordan Valley prior is defined on P_{ANN}. The change for the component of the T_{DJF} is comparably large (compare $\mu^{Litt}_{T_{DJF}} = 0.0$°C and $\sigma^{Litt}_{T_{DJF}} = 5.0$°C with Tab. 4.1).

3.) the expansion of the dimension of the climate state vector $\vec{C}$ resp. the realisation $\vec{c}$ from two to three dimensions (s. eq. 4.11),

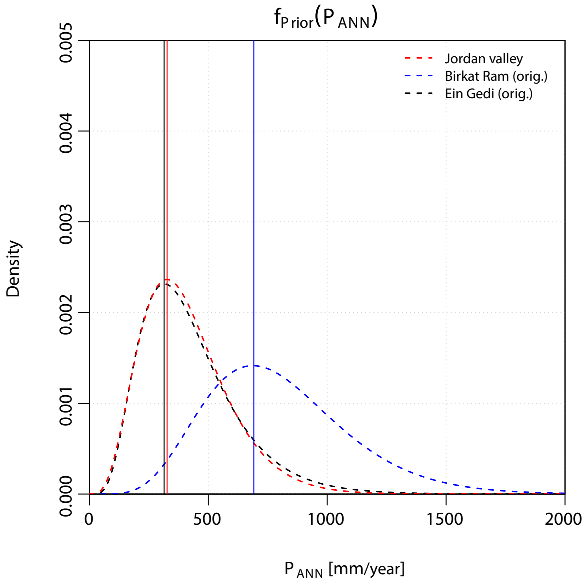
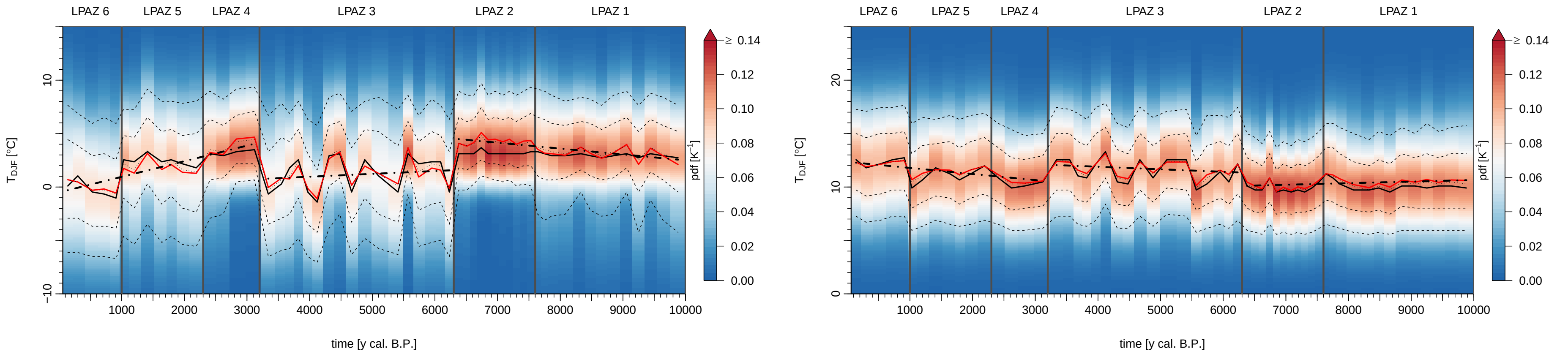


Figure 4.10.: Comparison of marginal prior pdfs in Israel for P_{ANN}. The parameters for the Jordan Valley prior pdf are defined in chapter 4.2.1 Table 4.1 on page 4.1. The values for the Birkat Ram prior pdf are defined in Neumann *et al.* (2007). The parameters for the Ein Gedi prior pdf are defined in Litt *et al.* (2012). All maxima of the pdfs are marked with a vertical line in the respective colour ($P_{ANN}{}^{JordanValley-max} = 327\,\frac{\text{mm}}{\text{year}}$, $P_{ANN}{}^{BirkatRam-max} = 692\,\frac{\text{mm}}{\text{year}}$ and $P_{ANN}{}^{EinGedi-max} = 314\,\frac{\text{mm}}{\text{year}}$).

4.) the change of the biome setup. The palaeoclimate reconstruction in Litt *et al.* (2012) bases of course on biome setup 1. As the different biome setups are in detail discussed in chapter 4.3 here only is presented the analysis what happens if biome setup 3 is applied.

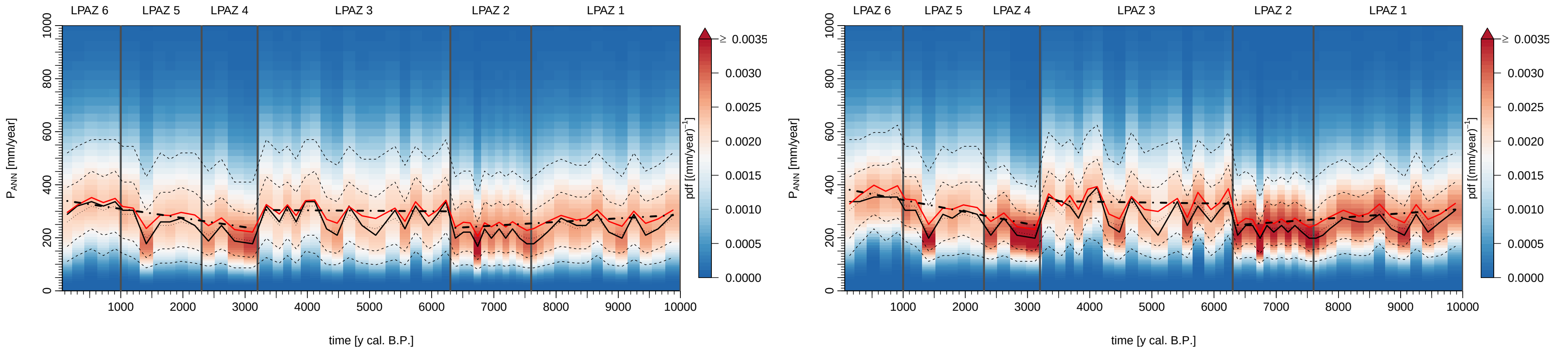
These adaptations are described step by step in the following.

Figure 4.11a and 4.11c show the result with a climate prior distribution very close to the definition in Litt *et al.* (2012): The parameters for the T_{DJF} are identical and for the P_{ANN} the Jordan Valley parameters are applied. A direct comparison with the result in Litt *et al.* (2012) (s. Fig. C.1) reveals that the update of the database from CRU TS 1.0 on CRU TS 3.1 has only little effects on the palaeoclimate reconstruction. The predominant part and the most important features of the marginal pdf depth profiles are conserved: The complete structure and consequently also the structure of the expectation value and three time ranges with different significant trends. A directly visible difference is a reduced variability. Also a shift of about 70 $\frac{\text{mm}}{\text{year}}$ of the complete profile for P_{ANN} to higher precipitation values is identifiable and probably also a result of the small differences of the climate priors discussed above.



(a) $f_{C_0|\vec{P}_0}(c_0 = T_{DJF}|\vec{p}_0)$ estimated with Litt *et al.* (2012) climate prior $\pi_{\vec{C}}(\vec{c}_0)$

(b) $f_{C_0|\vec{P}_0}(c_0 = T_{DJF}|\vec{p}_0)$ estimated with Jordan Valley climate prior $\pi_{\vec{C}}(\vec{c}_0)$



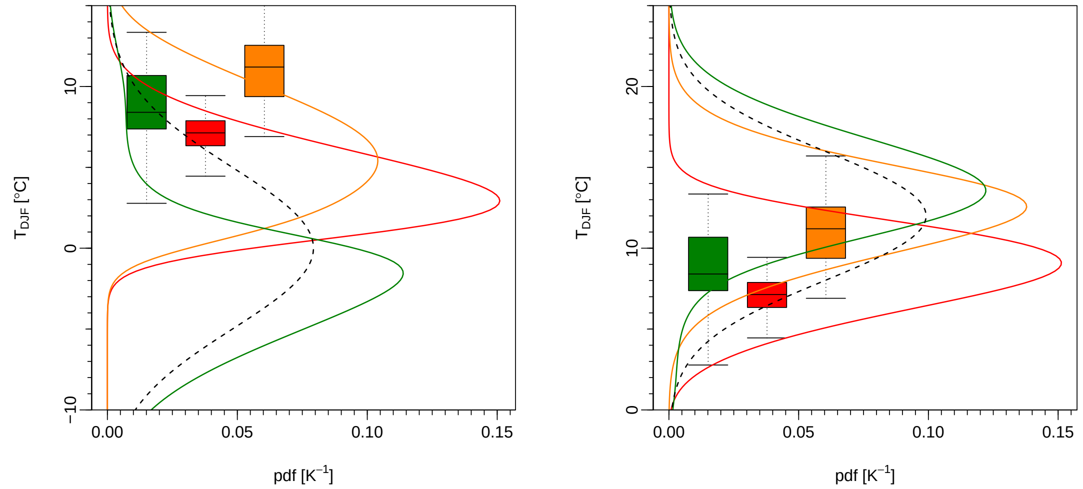
(c) $f_{C_0|\vec{P}_0}(c_0 = P_{ANN}|\vec{p}_0)$ estimated with Litt *et al.* (2012) climate prior $\pi_{\vec{C}}(\vec{c}_0)$

(d) $f_{C_0|\vec{P}_0}(c_0 = P_{ANN}|\vec{p}_0)$ estimated with Jordan Valley climate prior $\pi_{\vec{C}}(\vec{c}_0)$

Figure 4.11.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF} and P_{ANN} for the Ein Gedi palaeoclimate reconstruction estimated with different climate priors $\pi_{\vec{C}}\left(\vec{c}_0\right)$. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$ and s.no. 1 (biome setup 1, $\vec{C} = (T_{DJF}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset). (a) and (c) present the result with a climate prior distribution very close to the definition in Litt *et al.* (2012): The parameters for T_{DJF} are $\mu^{Litt}_{T_{DJF}} = 0.0°\mathrm{C}$ and $\sigma^{Litt}_{T_{DJF}} = 5.0°\mathrm{C}$. The parameters for P_{ANN} are the Jordan Valley parameters. (b) and (d) present the result estimated with the Jordan Valley prior as defined in Table 4.1. The mapping of the figure is identical to Figure 4.4. The x-coordinate in the palaeoclimate reconstruction for Ein Gedi is age in y. cal BP due to the fact that the age-to-depth model established by Migowski *et al.* (2004, 2006) is very reliable due to laminated layers. The LPAZ defined in Litt *et al.* (2012) are also shown.

The complete change of the climate prior to the Jordan Valley Prior has more consequences. Due to the fact that the P_{ANN} component of the prior is identical, the differences between Figure 4.11c and 4.11d are neglectable and a consequence of the other component of the climate prior. The change for T_{DJF} has more impact and emphasizes again how important it is to interpret the result with respect to the prior. The shift of the complete marginal pdf depth profile to higher T_{DJF} values of about 12 °C is a directly identifiable consequence of that since it is in the same order of magnitude as the difference between $\mu^{Litt}_{T_{DJF}}$ and $\mu_{T_{DJF}}$. Nevertheless the same three time ranges with different significant trends are detectable but with other results: On the one hand the gradients of the trends are inverted. On the other hand the gap between LPAZ 3 and the other two time ranges indicates a slight warming instead of significant cooling.

In order to interpret the last mentioned fact it is necessary to regard the marginal distributions $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(b_{l_0}, \vec{c}_0\right)$ (Fig. 4.12) together with the depth Profile of the biome probability $W_{B_{l_0}}$ (Fig. 4.13) as the analysis of the Lake Kinneret palaeoclimate reconstruction in the last section suggests. The box plots in Figure 4.12 are as in Figure 4.5 the visualisation of the input data.



(a) dedicated to Fig. 4.11a

(b) dedicated to Fig. 4.11b

Figure 4.12.: Marginal distributions of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(b_{l_0}, \vec{c}_0\right)$ as defined in eq. 4.15 (solid coloured lines) and the climate prior $\pi_{\vec{C}}\left(\vec{c}_0\right)$ (dashed black line) for T_{DJF} dedicated to the palaeoclimate reconstruction presented in Figure 4.11. The box plots are the visualisation of the input dataset for the estimation of the biome likelihoods (biome setup 1). Green colours represent the Mediterranean biome B_1, red colours the Irano-Turanian biome B_2 and orange colours the Saharo-Arabian biome B_3.

Although they base on biome setup 1 and not on biome setup 3 the discussed results in chapter 4.3 concerning the distinction of B_1 and B_2 persist, because applying biome setup 3 brings along some differences in the data but they are negligible small and not important in this consideration. B_3 compared with B_1 and B_2 based on the box plots for T_{DJF} allows a distinction of the biomes: It differs least 50 % from B_1 and 75 % from B_2. The figures of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathscr{b}_{l_0}, \vec{c}_0\right)$ for P_{ANN} are not shown due to the aforementioned neglectable difference in the palaeoclimate reconstruction. The comparison of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathscr{b}_{l_0}, \vec{c}_0\right)$ shows that the new climate prior moves the position of B_1 curve relative to the others: In Figure 4.12a B_1 is the coolest and in 4.12b the warmest biome. Also the differences between the biomes are overall smaller indicated by the fact that the modes in Figure 4.12b cover only an interval of 4 °C instead of 8 °C in Figure 4.12a. Both realisations of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathscr{b}_{l_0}, \vec{c}_0\right)$ have their errors in describing the input data: In the first case the covered domain compared with the input data is wrong. In the second case the order of the curves in the climate phase space is not correct.

The analysis of the depth profile of the biome probability $W_{B_{l_0}}$ (Fig. 4.13) explains the jump in the palaeoclimate reconstructions in Figure 4.11 at the borders of LPAZ 3. The greatest gradients in the depth profile of $W_{B_{l_0}}$ can be found at these borders. The jump in Figure 4.11a is clearly visible because the best distinguishable biomes B_1 and B_3 in Figure 4.12a have the largest and opposite gradients at these borders. Additionally the values for $W_{B_{l_0}=2}$, which is climatically situated between them but more similar to B_3, goes down resp. up at these borders. The same holds for Figure 4.11b and Figure 4.12b but with changed roles: B_1 and B_2 are the contrasting biomes and B_3 instead of B_2 is the multiplier. But due to the fact that the differences of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathscr{b}_{l_0}, \vec{c}_0\right)$ shown in Figure 4.12b are smaller the jump is less distinctive in 4.11b. Summarized the Mediterranean biome B_1 and its discriminability from the other

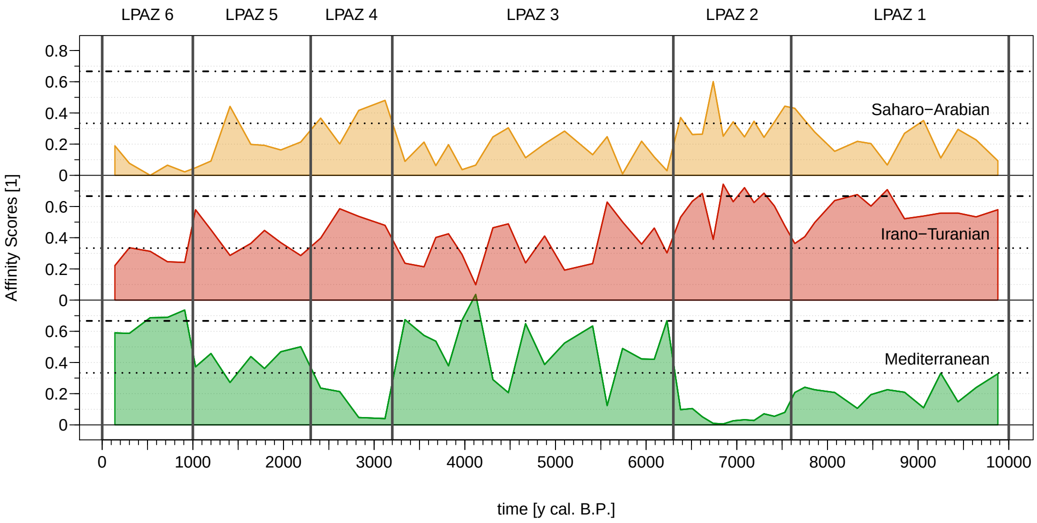
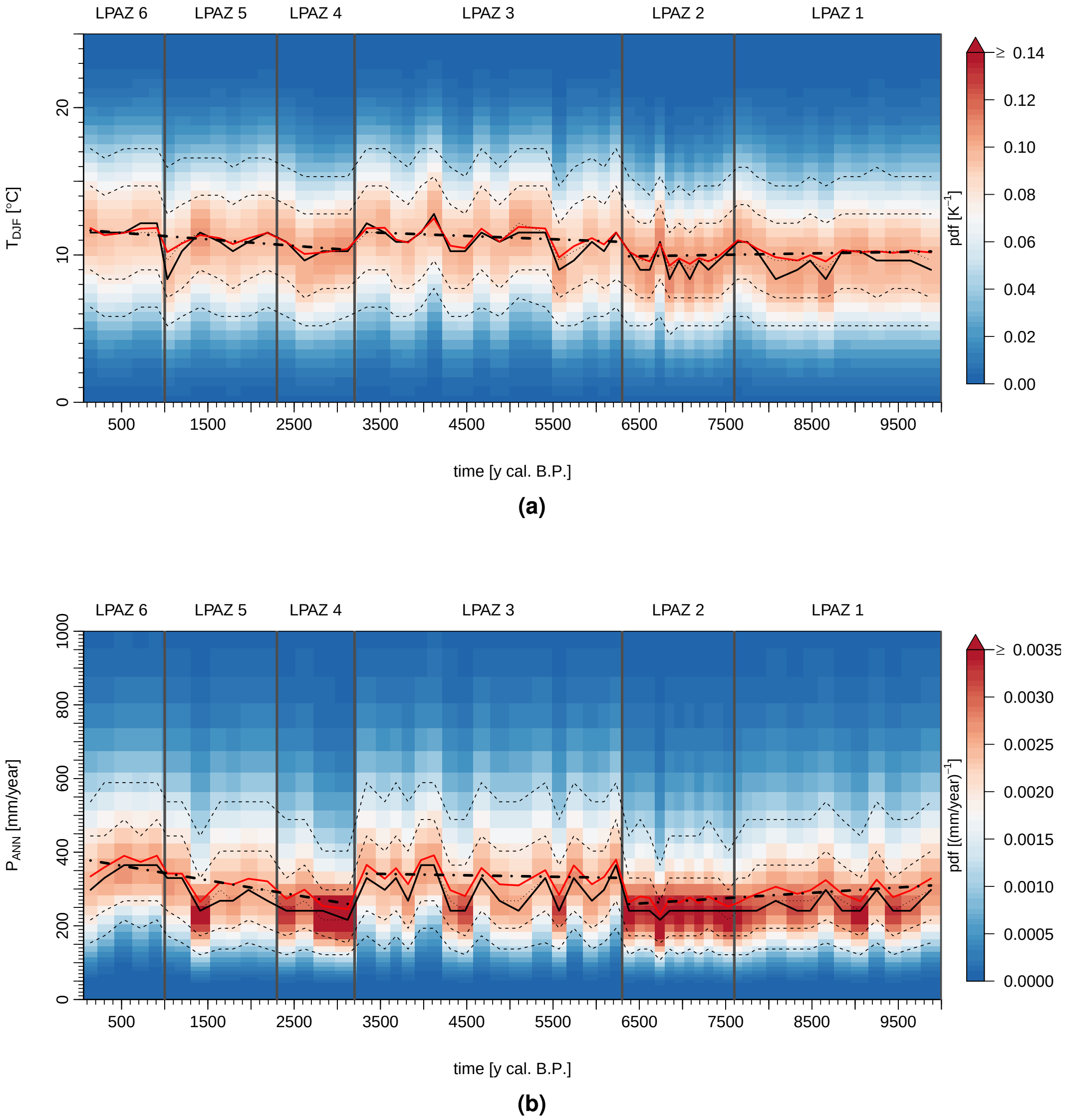


Figure 4.13.: Depth profile of the biome probability $W_{B_{l_0}}$ for Ein Gedi estimated with taxon specific individual thresholds ϑ_{k_0} (redrawn after Litt *et al.* 2012). The curve for $W_{B_{l_0=1}}$ equivalent to the Mediterranean biome is shown in green, for $W_{B_{l_0=2}}$ equivalent to the Irano-Turanian biome in red and $W_{B_{l_0=3}}$ equivalent to the Saharo-Arabian biome in orange. The vertical lines separate the local pollen assemblage zones (LPAZ) of Ein Gedi defined in Litt *et al.* (2012). The x-coordinate here is the age in y. cal BP. The dotted horizontal lines mark the probability of $\frac{1}{3}$ and the dashed-dotted the probability of $\frac{2}{3}$.

biomes B_2 and B_3 together with the high values and gradients of $W_{B_{l_0=1}}$ are the main reasons for the aforementioned change from cooling to warming and also for the detected palaeoclimate change in LPAZ 3.

The next step, as mentioned at the beginning of this chapter is the expansion of the climate state vector from two to three dimensions. Figure 4.14 and 4.15 present the result. The



(a)

(b)

Figure 4.14.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF} and P_{ANN} for the Ein Gedi palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = \mathit{individual}$ and s.no. 7 (biome setup 1, $\overrightarrow{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset). The mapping of the figure is identical to Figure 4.4 resp. 4.11.

differences between the components of the marginal posterior pdfs of the palaeoclimate reconstructions which are available for two and three dimensions, Figure 4.14a versus 4.11b and 4.14a versus 4.11d, are negligible small: For T_{DJF} a shift of the expectation value profile (red solid line) of about 0.5 °C and the 10 % quantile profile (lowest dashed black line) is reduced from about 7 °C to 6 °C. Also the jump at the borders of LPAZ 6 is negligibly reduced. For P_{ANN} the variability in the profile of the modus (black solid line) is reduced and the estimated linear trends is hardly affected. In the marginal pdf profile for T_{JJA} (Fig. 4.15) is no

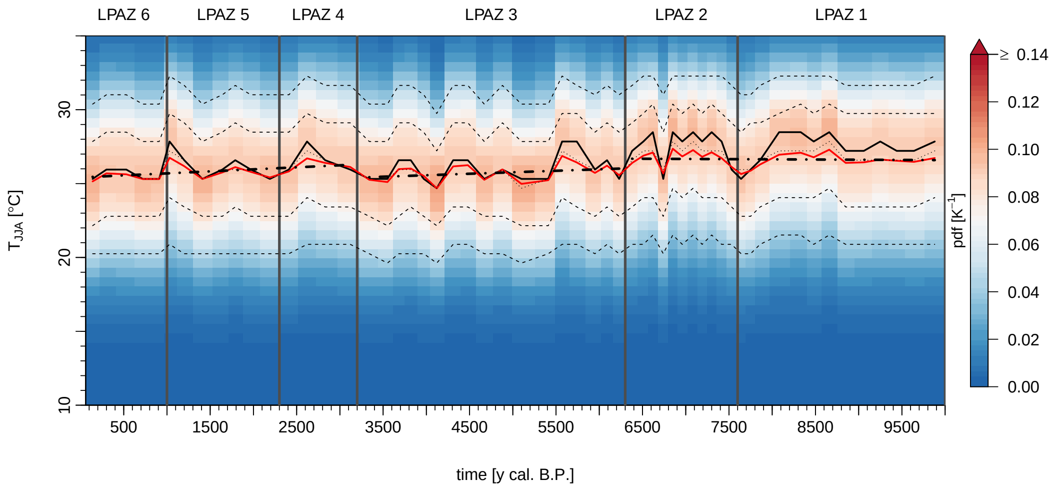


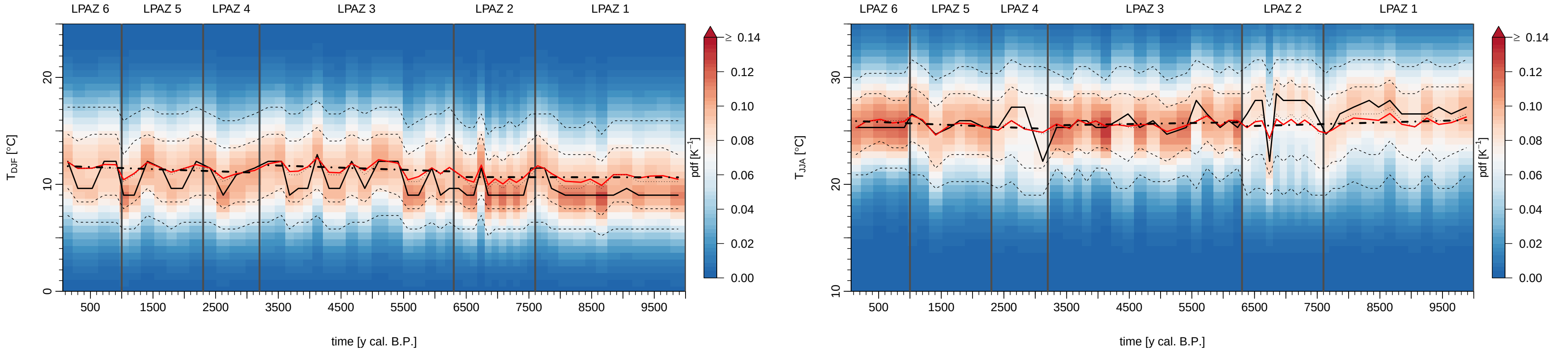
Figure 4.15.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{JJA} for the Ein Gedi palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$ and s.no. 7 (biome setup 1, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset). The mapping of the figure is identical to Figure 4.4 resp. 4.11.

time range identifiable with a climatic change. The explanation for this result is given in the next paragraphs in which the effect of a change of the biome setup from 1 to 3 is discussed. Furthermore all results of the expansion of the climate state vector are in agreement with those in chapter 4.3 based on Figure 4.8.

The change of the biome setup from one to three (Fig. 4.16 versus 4.14 and 4.15) has more impacts on the palaeoclimate reconstruction as the expansion of the climate state vector, especially here in the Ein Gedi T_{DJF} marginal pdf profile: no jump identifiable (Fig. 4.16a versus 4.14a) but much variability in mode (black solid line). The first mentioned fact is supported by an almost continuous line of the linear fit in the profile in Figure 4.16a. The explanation is provided by the analysis of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(b_{l_0},\vec{c}_0)$ shown in Figure 4.17. The modus of the T_{DJF} probability for B_1 is indistinguishable from B_3 and the clear discriminable biome B_2 has not the same pronounced gradients in the biome probability profile $W_{B_{l_0=2}}$ at the borders of LPAZ 3 like the other ones. Additionally $W_{B_{l_0=2}}$ has no high probability values and the other biomes B_1 and B_3 parallel quite low values for a larger time range. In this discussed configuration that would result in a time range in the marginal pdf profile for T_{DJF} with a climate change signal. This hypothesis is supported by the fact that there is at least one timeslice at 6740 y. cal BP situated in LPAZ 2 where the expectation value differs from the rest in LPAZ 2 by a value of about 2 °C.

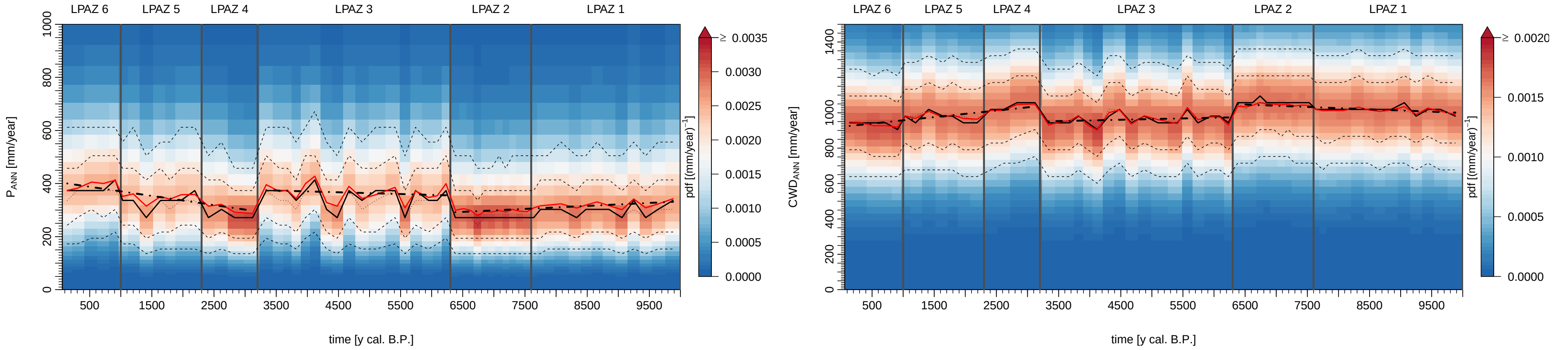
The corresponding T_{JJA} palaeoclimate reconstruction (Figure 4.17b) also shows no time range with a climate change. The curve of the expectation value (red solid line) and mode (black solid line) is similar to 4.15, but the jumps/gradients in the curve of the mode in LPAZ 4 and LPAZ 2 in 4.16b are larger/more pronounced. The reason for no time range with climate change in 4.15 is the same as described in the last paragraph for T_{DJF} but with moved position of the probability $\mathbb{P}_{\mathcal{B}_2,\vec{C}}(b_{2_0},\vec{c}_0)$ relative to the other: B_2 is the warmest biome (Figure not shown). A position of each $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(b_{l_0},\vec{c}_0)$ relative to the others as shown in Figure 4.17b is also no indicator for a climate change signal as the result in Fig. 4.16b demonstrates. The box plot for B_3 and T_{JJA} shows that the input data does not allow any differentiation between B_3 and B_2 but it differs 75 % from B_1, so that the only explanation for the curve progression is a result of covariance structures. So the result for T_{JJA} has to be interpreted with caution. Switched back to the $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(b_{l_0},\vec{c}_0)$ and the palaeoclimate reconstruction: A time frame with climate change in that configuration is only possible if B_1 gets low values as it is the case for the timeslices at 6740 y. cal BP and 8660 y. cal BP ($\equiv$ pollen sample layer 37 and 51).

The box plot for B_3 of the water variables (Fig. 4.17c and 4.17d) shows that the input data allows a differentiation of the biomes due to the fact that it has only a maximum overlap of 25 %. The locations of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(b_{l_0},\vec{c}_0)$ fit only for P_{ANN} so that the result for CWD_{ANN}, just like the others, has to be interpreted with caution. The palaeoclimate reconstructions for the water variables indicate a time range with climate change, the time range associated with LPAZ 3. The result for P_{ANN} is in agreement with the other already presented results (Fig. 4.11c, 4.11d and 4.14b), only negligible changes in the variation of estimated quantile curves (dashed lines), the mode (solid black line) and the expectation value (red solid line). As already mentioned, the palaeoclimate reconstruction of CWD_{ANN} (Fig. 4.16d) shows in LPAZ 3 a jump of 100 $\frac{\text{mm}}{\text{year}}$ to lower values. Despite the just mentioned problems with $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(b_{l_0},\vec{c}_0)$ this is in agreement with the P_{ANN} reconstruction since lower CWD_{ANN} values mean, that more water is available (s. p. 10). As in the Lake Kinneret palaeoclimate reconstruction for CWD_{ANN}, the corresponding figures to Figure 4.16d for T_{DJF} and T_{JJA} are not shown since the result is equal to the shown results in Figure 4.16a and 4.16b.



(a) s.no. 21

(b) s.no. 21

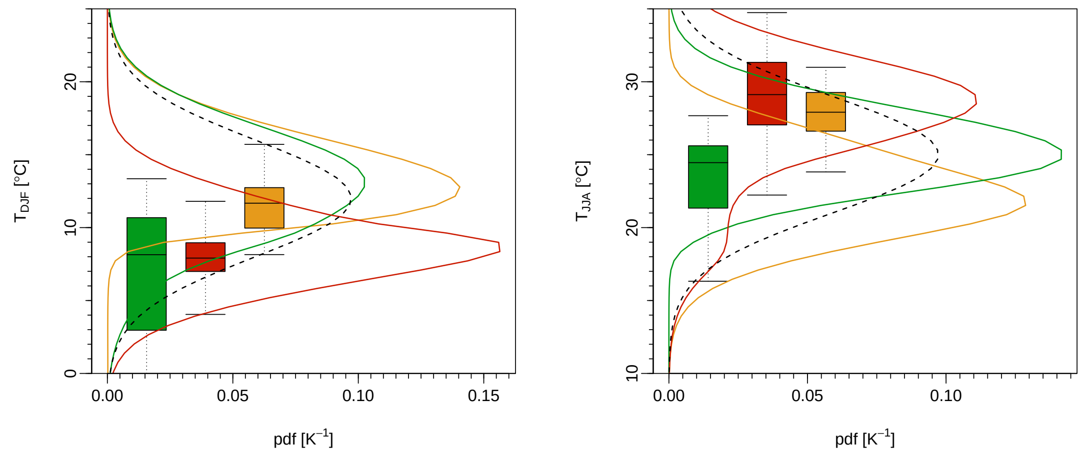


(c) s.no. 21

(d) s.no. 20

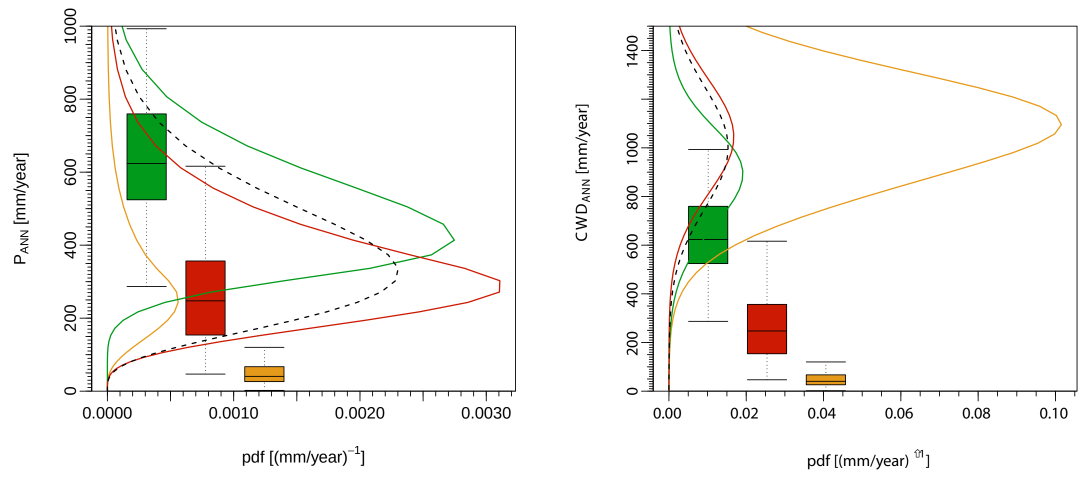
Figure 4.16.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} for the Ein Gedi palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$. (a), (b) and (c) base on s.no. 21 (biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset) and (d) on s.no. 20 which is equal to s.no. 21 and differs only in the definition the climate state vector ($\vec{C} = (T_{DJF}, T_{JJA}, CWD_{ANN})^T$). The mapping of the figure is identical to Figure 4.4 resp. 4.11.

The analysis of the two water variable palaeoclimate reconstructions is also in agreement with the hypothesis that the dominant biome type in the $W_{B_{l_0}}$ profile, here B_1, has to be clearly discriminable in $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathscr{b}_{l_0},\vec{c}_0)$, at least the modes and not in between as it is the case in Figure 4.17b for T_{JJA}. It should also be mentioned that despite some differences all $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathscr{b}_{l_0},\vec{c}_0)$



(a) dedicated to Fig. 4.16a

(b) dedicated to Fig. 4.16b



(c) dedicated to Fig. 4.16c

(d) dedicated to Fig. 4.16d

Figure 4.17.: Marginal distributions of $\mathbb{P}_{\mathcal{B}_l,\vec{C}}\left(\mathscr{b}_{l_0},\vec{c}_0\right)$ as defined in eq. 4.15 (solid coloured lines) and the climate prior $\pi_{\vec{C}}\left(\vec{c}_0\right)$ (dashed black line) for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} dedicated to the palaeoclimate reconstruction presented in Figure 4.16. The box plots are the visualisation of the input dataset for the estimation of the biome likelihoods (biome setup 3) so that for B_1 and B_2 they show the same result as in Figure 4.5. The mapping of the colours is identical to Figure 4.12.

for B_1 and B_2 in Figure 4.17 and 4.5 do not contradict each other with respect to the position of the mode and shape.

4.6. The Bayesian Biome Model - Conclusions

With the BBM it is possible to find a time range with palaeoclimate change if some requirements are fulfilled:

- One biome in the depth profile of $W_{B_{l_0}}$ has to dominate for several pollen sample layers/time slices. Dominating here implies high $W_{B_{l_0}}$ values and low for the other ones.
- Additionally this dominating biome has to be clearly discriminable in the $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathscr{b}_{l_0},\vec{c}_0)$, the clearer the better. If more than two biomes are considered in the BBM then a location of this biome between the other one is critical.
- If the number of considered biomes is greater than two and two of the biomes have an identical mode for $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathscr{b}_{l_0},\vec{c}_0)$, then the biome which is clearly discriminable has to be the dominating one in $W_{B_{l_0}}$. This fact is one reason why it is possible to identify a period with climate change in the reconstruction for Ein Gedi and not in the one for Lake Kinneret.
- The probability $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathscr{b}_{l_0},\vec{c}_0)$ of one biome is always strongly influenced by the climate prior: In the climate phase space which is splitted into several areas by the QDA, this biome is the one which is situated at the edges. In the cases considered in this chapter this is the Mediterranean biome B_1
- The data basis for the QDA is thin/sparse which has the consequence that the QDA result is not robust especially for the Mediterranean biome.

5

Birkat Ram

This chapter presents the result and the mathematical tools that are used for the palaeoclimate reconstruction which bases on the new sediment core data of Birkat Ram. The sediment core and the locations are described in chapter 2.3.2. The basic mathematical principles for this purpose are introduced in chapter 1.2 and are extended in the first section of this chapter to the new Bayesian Indicator Taxa and Biome Model (BITBM) which is applied to the sediment core. The following section describes the estimation of each component of the BITBM with the corresponding results for Birkat Ram. Directly follows the palaeoclimate reconstruction result for Birkat Ram.

5.1. Bayesian Indicator Taxa and Biome Model - Introduction

The palaeoclimate reconstruction for the maar lake Birkat Ram presented in this chapter bases on the pollen counts of the new sediment core as described in chapter 2.3.2.2. The ”old” Birkat Ram core and the corresponding result of the palaeoclimate reconstruction is described in Neumann *et al.* (2007) and the database in detail in Schölzel (2006). Schölzel (2006) also discussed that the identified pollen in the sediment core indicate that the Birkat Ram area was always situated in the Mediterranean biome/climate. Therefore he only applied taxa distribution maps which describe the distribution in Mediterranean biome or more precisely in the direct vicinity of Birkat Ram (s. Schölzel 2006, chap. 3.3.2).

As discovered by Schiebel (2013) this is only the case for the time frame from present time to approximately 10 000 y. cal BP or in other words in the Holocene (s. Schiebel 2013, chap. 6). The ”new” Birkat Ram sediment core covers approximately a time range of 30 000 y. cal BP (s. Schiebel 2013, chap. 5.2.3, p. 49 para. 3). From 17 000 y. cal BP taxa assigned to the Irano-Turanian dominate. To reflect that fact the BITM, as described in chapter 3.2, is extended/expanded to the BITBM. A similar approach is used in Kühl *et al.* (2007, p. 3313) by applying the general condition G (eq. 3.7). The BITM includes individual taxa as well as biome climate by regarding a taxon random variable as defined in equation 3.2, a biome random variable as defined in equation 4.1 and a climate state random variable as defined for example in equation 3.1. Consequently the starting point is the joint probability $\mathbb{P}_{\vec{C},\vec{T},\mathcal{B}_l}\left(\vec{c},\vec{t},b_l\right)$. The goal is to achieve the probability resp. pdf of climate given proxy data (as it is the case for the BITM and the BBM) or in the case considered here taxa and biome data: $\mathbb{P}_{\vec{C}|\vec{T},\mathcal{B}_l}\left(\vec{c}\,|\,\vec{t}\,,\mathscr{b}_l\right)$. As described in chapter 4.1 the reconstructed palaeoclimate is always a mixture signal of different biome types N_l. Hence it has to be summed over the different biome types which is equal to marginalisation and similar to equation 4.3

$$\mathbb{P}_{\vec{C}|\vec{T}}\left(\vec{c}\,\middle|\,\vec{t}\right)=\sum_{l=1}^{N_l}\mathbb{P}_{\mathcal{B}_l|\vec{T}}\left(\mathscr{b}_l\,\middle|\,\vec{t}\right)\cdot\mathbb{P}_{\vec{C}|\vec{T},\mathcal{B}_l}\left(\vec{c}\,\middle|\,\vec{t}\,,\mathscr{b}_l\right)\tag{5.1}$$

To interpret the meaning of $\mathbb{P}_{\vec{C}|\vec{T},\mathcal{B}_l}\left(\vec{c}\,\middle|\,\vec{t}\,,\mathscr{b}_l\right)$ it is necessary to apply again the Bayes theorem:

$$\mathbb{P}_{\vec{C}|\vec{T},\mathcal{B}_l}\left(\vec{c}\,|\,\vec{t}\,,\mathscr{b}_l\right)=\frac{\mathbb{P}_{\vec{C},\vec{T},\mathcal{B}_l}\left(\vec{c}\,,\vec{t}\,,\mathscr{b}_l\right)}{\mathbb{P}_{\vec{T},\mathcal{B}_l}\left(\vec{t}\,,\mathscr{b}_l\right)}=\frac{\mathbb{P}_{\vec{T}|\mathcal{B}_l,\vec{C}}\left(\vec{t}\,|\mathscr{b}_l,\vec{c}\right)\cdot\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_l|\vec{c}\right)\cdot\mathbb{P}_{\vec{C}}\left(\vec{c}\right)}{\mathbb{P}_{\vec{T},\mathcal{B}_l}\left(\vec{t}\,,\mathscr{b}_l\right)}\tag{5.2}$$

The next step is to regard the assumption of pairwise conditional independence of all taxa given a certain climate state and biome type. This is an extension of the BITM assumption already presented in chapter 3.1.2. The result is that $\mathbb{P}_{\vec{T}|\mathcal{B}_l,\vec{C}}\left(\vec{t}\,|\mathscr{b}_l,\vec{c}\right)$ splits into a product analogous to equation 3.5:

$$\mathbb{P}_{\vec{T}|\mathcal{B}_l,\vec{C}}\left(\vec{t}\,|\mathscr{b}_l,\vec{c}\right)=\prod_{k=1}^{N_k}\mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_k|\,\vec{c}\,,\mathscr{b}_l\right)\tag{5.3}$$

The BITBM equation is obtained after the realisation of two further steps: First the results of equation 5.1, 5.2 and 5.3 are summarized in one equation and the common notations for prior and marginal distributions as described in chapter 1.2 (after eq. 1.4) are used since they fulfil the required definitions. Secondly the basic assumption as described in chapter 1.2 (eq. 1.5 on p. 5) and also mentioned in chapter 4.1 in the introduction of the BBM (eq. 4.4 on p. 52) is applied here:

$$\mathbb{P}_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0\,\middle|\,\vec{t}_0\right)=\sum_{l=1}^{N_l}\mathbb{P}_{\mathcal{B}_l|\vec{T}}\left(\mathscr{b}_{l_0}|\,\vec{t}_0\right)\cdot\frac{\left[\prod_{k=1}^{N_k}\mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_{k_0}|\,\vec{c}_0,\mathscr{b}_{l_0}\right)\right]\cdot\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}|\,\vec{c}_0\right)\cdot\pi_{\vec{C}}\left(\vec{c}_0\right)}{m_{\vec{T},\mathcal{B}_l}\left(\vec{t}_0,\mathscr{b}_{l_0}\right)}\tag{5.4}$$

This is in principle the equation for the palaeoclimate reconstructions based on the BITBM. Some more steps have to be done: The multiplication index has to reflect that only taxa which are present in the regarded pollen sample layer ι found in depth d_ι are multiplied ($t_{\iota\, k_0}=1$) as analogous done for the BITM in equation 3.14. And as last step $\mathbb{P}_{\mathcal{B}_l|\vec{T}}\left(\mathscr{b}_{l_0}|\,\vec{t}_0\right)$ is replaced by the biome probability $W_{B_{l_0}}$ analogous to the BBM equation 4.5 (s. chap. 5.2.2 for details):

$$\mathbb{P}_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0|\vec{t}_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right) \cdot \sum_{l=1}^{N_l} W_{B_{l_0}} \frac{\left[\prod_{\substack{k \\ \forall t_{\iota k_0}=1}} \mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_{k_0}|\vec{c}_0, b_{l_0}\right)\right] \cdot \mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\vec{c}_0\right)}{m_{T,\mathcal{B}_l}\left(t_0, b_{l_0}\right)} \tag{5.5}$$

5.2. The BITBM in Detail for Birkat Ram

Analogous to chapter 3.2 and 4.2 all parts of equation 5.5 which contribute to the palaeoclimate reconstruction are discussed in detail in this section. Hereby the focus is set on the differences to BITM and BBM. Also the corresponding result for each part is presented.

5.2.1. $\pi_{\vec{C}}\left(\vec{c}_0\right)$ Prior Distribution for the climate state vector

For the same reasons as in chapter 4.2.1 a climate prior for the whole Jordan Valley is applied. As consequence the same parameters as shown in Table 4.1 are used.

5.2.2. $W_{B_{l_0}}$ Biome Probability

The biome probability $W_{B_{l_0}}$ is estimated in the same way as in chapter 4.2.2 (eq. 4.6 and 4.7) for Lake Kinneret. The principal difference is that only taxa which are used for the indicator taxa matrix $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ (s. chap. 5.2.3 for details) are included in the estimation of $W_{B_{l_0}}$. Due to the fact that Birkat Ram is not far away from Lake Kinneret the same biome types are available/possible. These selected taxa are (T. Litt pers. comm.): *Artemisia* and *Plantago* for the Irano-Turanian biome type B_2 and *Pistacia*, *Olea*, *Quercus ithaburensis type* and *Quercus calliprinos type* for the Mediterranean biome type B_1. For consistency the required taxon specific threshold values ϑ_{k_0} are the same as later on applied for $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ and can be found in Table D.1.

Figure 5.1 presents the resulting depth profile of the biome probability $W_{B_{l_0}}$ for Birkat Ram. The area below the curve represents the Mediterranean biome probability $W_{B_{l_0=1}}$ and between one and the curve represents the Irano-Turanian biome probability $W_{B_{l_0=2}}$. The x-coordinate in all figures presented in this chapter is age in y. cal BP due to the fact that only one reliable age-to-depth model established by Schiebel (2013) is available (s. chap. 2.3.2.2). The hatched area between ≈ 10000 and 17000 y. cal BP (≡ 703 cm and 746 cm) marks the range where Schiebel (2013) assumed a desiccation of Birkat Ram. The shown $W_{B_{l_0}}$ profile is divided in three different zones: 0 to ≈ 10000 y. cal BP, ≈ 10000 to ≈17000 y. cal BP and ≈ 17000 to ≈30000 y. cal BP. The first zone could be named Holocene zone since it lies completely in this geological epoch. It consists of 28 pollen sample layers and it shows a clear preference that Birkat Ram is located in the Mediterranean biome type, which is in agreement with Neumann *et al.* (2007);

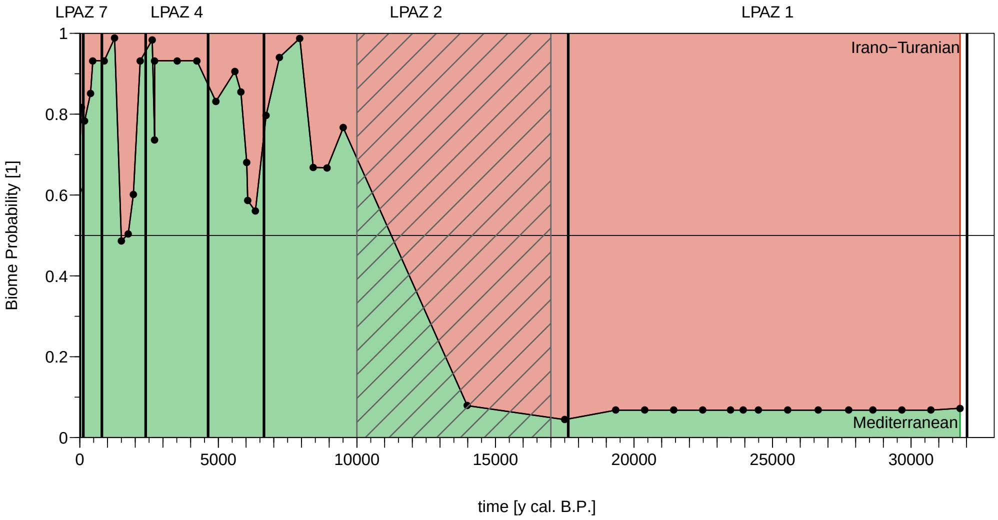


Figure 5.1.: Depth profile of the biome probability $W_{B_{l_0}}$ for Birkat Ram. The area below the curve represents the Mediterranean biome probability $W_{B_{l_0=1}}$ and between one and the curve represents the Irano-Turanian biome probability $W_{B_{l_0=2}}$. The vertical lines separate the local pollen assemblage zones (LPAZ) of Birkat Ram defined in Schiebel (2013, chap. 5.2.3, Tab. 5.2 and Fig. 5.3 $\equiv$ Fig. 2.13 in this work). The x-coordinate is age in y. cal BP due to the fact that only one age-to-depth model established by Schiebel (2013) is available (s. chap. 2.3.2.2). The hatched area between $\approx$ 10 000 and 17 000 y. cal BP ($\equiv$ 703 cm and 746 cm) marks the range where Schiebel (2013) assumed a desiccation of Birkat Ram.

Schiebel (2013); Schölzel (2006). In detail preference means that 93 % of the data points fulfil $W_{B_{l_0=1}} > 0.5$ and 67 % fulfil $W_{B_{l_0=1}} \geq 0.7$.

Only two layers (186 cm $\overset{\approx}{\Leftrightarrow}$ 1 501 y. cal BP and 211 cm $\overset{\approx}{\Leftrightarrow}$ 1 741 y. cal BP) have values of round about 0.5. This is a direct consequence of an increase of pollen counts from the Irano-Turanian taxon *Plantago* (s. Fig. D.1.2 resp. Fig. 5.2, PlSM_BR) and at the same time a decrease of Mediterranean taxa *Olea* (s. Fig. D.1.2 resp. Fig. 5.2, Oleu_BR) and additionally overall low pollen count values for the remaining Mediterranean taxa. Schiebel (2013, chap. 6.4.3) refers this period in the Hellenistic, Roman and Byzantine time period with a strong anthropogenic influence on *Olea*, so that this part of the signal has to be interpreted as human made and it has to be specially regarded in the palaeoclimate reconstruction in chapter 5.3. There are two other similar parts in the signal in the Holocene part: around 514 cm $\overset{\approx}{\Leftrightarrow}$ 6 059 y. cal BP/539 cm $\overset{\approx}{\Leftrightarrow}$ 6 336 y. cal BP and 647 cm $\overset{\approx}{\Leftrightarrow}$ 8 424 y. cal BP/673 cm $\overset{\approx}{\Leftrightarrow}$ 8 927 y. cal BP. The first one has similar circumstances as the pollen sample layer around 186 cm $\overset{\approx}{\Leftrightarrow}$ 1 501 y. cal BP / 211 cm $\overset{\approx}{\Leftrightarrow}$ 1 741 y. cal BP. The main difference is that all other Mediterranean taxa (*Pistacia*, *Quercus ithaburensis type* and *Quercus calliprinos type*) have higher pollen count values, especially *Quercus ithaburensis type* and consequently the assigned probability $W_{B_{l_0=1}}$ has higher values since more taxa are included in the estimation. Schiebel (2013, chap. 6.4.2) relates this with beginning of olive cultivation in the Chalcolithic period so that this part of the signal has also to be regarded in chapter 5.3. The circumstances for the second one are different: Only *Plantago* and *Quercus ithaburensis type* are included in estimation of $W_{B_{l_0}}$ so that each biome is represented by only one taxon. This situation occurs only in LPAZ 2 around the above mentioned sample layers and in the first pollen sample layer. Schiebel (2013) analysed more taxa and concluded that this part of the signal is connected with the so called 8.2-event of rapid climate change with return of the Irano-Turanian steppe vegetation which is in agreement with the biome probability result despite the spare data in the estimation.

The second zone in the $W_{B_{l_0}}$-profile consists of only one pollen sample layer. It is the zone where Schiebel (2013) assumed a desiccation of Birkat-Ram and consequently the age-to-depth model covers a large time interval. Schiebel (2013) also concluded that due to this discontinuity in the sediment layers ”*no conclusion concerning a possible dispersion of Mediterranean vegetation can be drawn*”. This fits with the result for $W_{B_{l_0}}$ presented here since in zone one the Mediterranean biome type is most likely and in the third zone the Irano-Turanian biome with a large jump in the desiccation-zone.

As already mentioned the last zone which consists of 15 pollen sample layers shows a clear preference for the Irano-Turanian biome with $W_{B_{l_0=2}}$ values all in the order of 90 %. It matches also with the result from Schiebel (2013) of dominating ”*steppe-like character of the vegetation in the vicinity of Birkat Ram*” in this period. Summarized the biome probability $W_{B_{l_0}}$ result here is in a good agreement with the results of the pollen analysis of Schiebel (2013) which includes all taxa detected in the Birkat Ram sediment core.

5.2.3. Indicator Taxa and Taxa Transfer Function

The part in the BITBM which accounts the contribution of the individual taxa is the product in the numerator of equation 5.5. It is quasi a BITM which differs from the one applied on the ”old” Birkat-Ram sediment core (s. chap. 2.3.2.2 and Neumann *et al.* 2007). First the transfer functions are directly estimated via GLM in the same way as in chapter 3.2.2 and therefore no further application of the Bayes theorem is required as it is the case for eq. 3.6. Second the distribution map for each single taxa differ. Neumann *et al.* (2007); Schölzel (2006) applied, as mentioned, only Mediterranean distribution maps based on Feinbrun-Dothan, Danin (1998). In the case considered here a summary of them together with the maps presented in chapter 2.2.1 is used.

As indicators six taxa are selected (T. Litt pers. comm.) which are the same as those listed in the estimation of the biome probability $W_{B_{l_0}}$. These taxa are also present in the pollen samples of the ”old” Birkat Ram sediment core. The transformation of pollen counts to presence and absence information summarized in the required indicator taxa matrix $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ is done with the same approach as in chapter 3.2.1. The corresponding figures for the ϑ_{k_0}-estimation are presented in appendix D.1.2. Especially for *Quercus ithaburensis type* the advantage of a possible adaptation of ϑ_{k_0} to include the background knowledge of the palaeontologist is taken:

Quercus ithaburensis type is only present in a contiguous zone in the estimated indicator profile (s. Fig. 5.2 row Qud_BR and Fig. D.1.2).

In contrast the palaeoclimate reconstruction of Lake Prespa only six taxa are the past proxy database so that no mahalanobis distance filtering is required (s. p. 29). The resulting indicator taxa matrix is shown in Figure 5.2 with the same mapping as in Fig. 3.3a/3.3c for the absence information ($t_{\iota k_0} = 0$, two light grey colours) and the presence information ($t_{\iota k_0} = 1$, two dark grey colours). The shortcuts on the vertical axis refer to the full taxon name in Tab. D.1 (p. 205). The hatched area marks as in Fig. 5.1 the range where Schiebel (2013) assumed a desiccation of Birkat Ram. The assignment of the taxa to the biomes is marked by the coloured dots: Mediterranean biome B_1 in green and Irano-Turanian biome B_2 in red (for details also Tab. D.1).

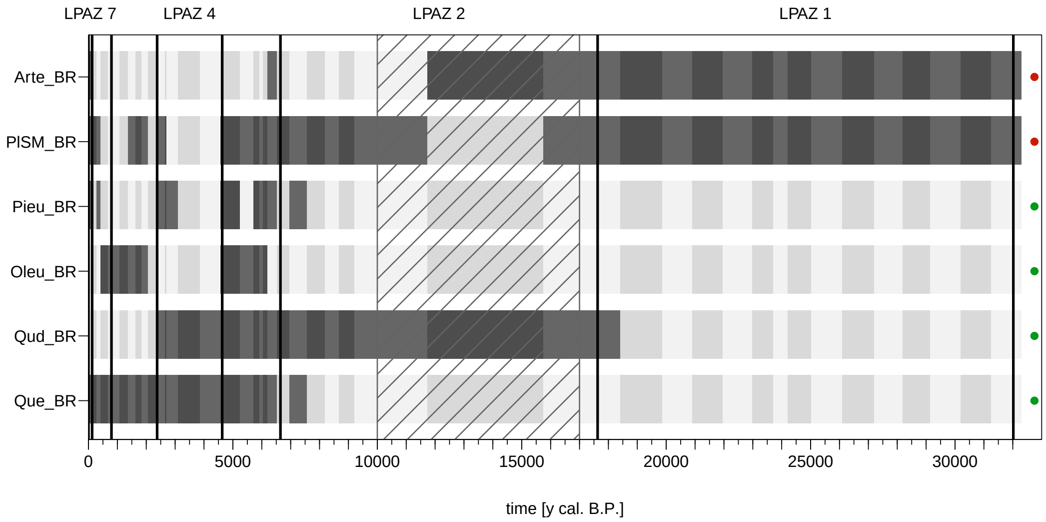


Figure 5.2.: Indicator Taxa Matrix for Birkat Ram $\mathcal{I}_{\vartheta_0}^{N_t \times N_{k_0}}$. The figure shows the absence information ($t_{\iota k_0} = 0$) for each regarded taxon k_0 in two light grey colours and the presence information ($t_{\iota k_0} = 1$) in two dark grey colours. (same mapping as in Fig. 3.3a/3.3c) The shortcuts on the vertical axis refer to the full taxon name in Tab. D.1 (p. 205). The assignment of the taxa to the biomes is marked by the coloured dots: Mediterranean biome B_1 in green and Irano-Turanian biome B_2 in red (for details also Tab. D.1). The vertical lines separate the local pollen assemblage zones (LPAZ) of Birkat Ram defined in Schiebel (2013) (chap. 5.2.3, Fig. 5.3 and Table 5.2 or Fig. 2.13 in this work). The hatched area between $\approx 10\,000$ and $17\,000$ y. cal BP ($\equiv 703$ cm and 746 cm) marks the range where Schiebel (2013) assumed a desiccation of Birkat Ram.

These so called "biome marker" are important: The aforementioned taxon part of the BITBM,

$$\prod_{\substack{k \\ \forall t_{\iota k_0}=1}} \mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_{k_0}|\,\vec{c}_0, b_{l_0}\right), \tag{5.6}$$

includes one further difference to the BITM of chapter 3.2. The probability has two conditions (climate and biome) instead of one (climate), which is considered as follows. If the summation in equation 5.5 is e. g. at biome type $B_{l=1}$ (here equal to the Mediterranean biome) only taxa which are assigned to this biome type ($\mathscr{b}_{\iota\, l_0} = l$) and regarded as present ($t_{\iota\, k_0} = 1$) are multiplied.

Two properties of the indicator taxa matrix $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ presented in Figure 5.2 are important to note: In all pollen sample layers or resp. time slices are at least two taxa regarded as present except 255 cm $\overset{\approx}{\Leftrightarrow}$ 2 180 y. cal BP. Secondly in the first of the three zones identified in Figure 5.1 every layer has at least one Mediterranean taxon regarded as present and the third zone at least two Irano-Turanian taxa. This result is in agreement with the findings of the last chapter since the dominating biome type has always at least one assigned taxon regarded as present.

5.2.4. $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$ Biome Likelihood - Biome Transfer Function

The biome likelihood or biome transfer function $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$ has the same functionality as in the BBM. Chapter 4.2.3 describes in detail the estimation via QDA which is also used here with the same setups for the estimation. Due to the fact that biome setup 3 is rated as the best and selected for the palaeoclimate reconstruction for Lake Kinneret and Ein Gedi it is also applied here. The exact s.no. is specified in chapter 5.3.

5.2.5. $m_{\mathcal{B}_l}\left(\mathscr{b}_{l_0}\right)$ Marginal Probability

The last not yet defined and described part of the BITBM is the definition of the marginal probability $m_{\mathcal{B}_l}\left(\mathscr{b}_{l_0}\right)$. It differs from the normalization of the BBM presented in chapter 4.2.4. Here the normalisation is defined as the sum over the numerator of equation 5.5

$$m_{T,\mathcal{B}_l}\left(t_0, \mathscr{b}_{l_0}\right) := \sum_{i=1}^{N_l} \left[\prod_{\substack{k \\ \forall t_{\iota\, k_0}=1}} \mathbb{P}_{T_k|\vec{C},\mathcal{B}_i}\left(t_{k_0}\middle|\vec{c}_0, \mathscr{b}_{i_0}\right) \right] \cdot \mathbb{P}_{\mathcal{B}_i|\vec{C}}\left(\mathscr{b}_{i_0}\middle|\vec{c}_0\right). \tag{5.7}$$

This normalisation ensures that the result of the BITBM fulfils the mathematical definition for a probability. The results of the palaeoclimate reconstruction presented in the following base on

$$\mathbb{P}_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0\middle|\vec{t}_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right) \sum_{l=1}^{N_l} W_{B_{l_0}} \cdot \frac{\left[\prod_{\substack{k \\ \forall t_{\iota\, k_0}=1}} \mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_{k_0}\middle|\vec{c}_0, \mathscr{b}_{l_0}\right) \right] \cdot \mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)}{\sum_{i=1}^{N_l} \left\{ \left[\prod_{\substack{k \\ \forall t_{\iota\, k_0}=1}} \mathbb{P}_{T_k|\vec{C},\mathcal{B}_i}\left(t_{k_0}\middle|\vec{c}_0, \mathscr{b}_{i_0}\right) \right] \cdot \mathbb{P}_{\mathcal{B}_i|\vec{C}}\left(\mathscr{b}_{i_0}\middle|\vec{c}_0\right) \right\}}, \tag{5.8}$$

which is the summary of equation 5.7 and 5.5 and applied on each pollen sample layer of Birkat Ram.

5.3. Birkat Ram - Palaeoclimate Reconstruction Result

All figures (Fig. 5.3 and 5.4) in this chapter show the estimated result of the palaeoclimate reconstruction for Birkat Ram in terms of a depth profile of one-dimensional marginal probabilities $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$. The term marginal is already defined in chapter 3.2.2. Thereby x-coordinate depth is replaced by age in y. cal BP due to the fact that only one age-to-depth model established by Schiebel (2013) is available (s. chap. 2.3.2.2). The mapping of the figures for the posterior probabilities for the Birkat Ram palaeoclimate reconstruction is in principle the same as that of the Lake Prespa, Lake Kinneret and Ein Gedi figures, which is already described at the beginning of chapter 3.3 (s. p. 40). Only the meaning of the coloured scale changes slightly since the presented results are probabilities: The scale ranges from blue for probability values equal to zero to red for probability values equal to one. Also the mode is not shown in the figures. The hatched area between ≈ 10 000 and 17 000 y. cal BP (≡ 703 cm and 746 cm) marks again as in Fig. 5.1 and 5.2 the range where Schiebel (2013) assumed a desiccation of Birkat Ram.

The palaeoclimate reconstruction presented in Figure 5.3a, 5.3b and 5.3c bases on the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\,\vec{c}_0\right)$ whose setup for the estimation is defined by s.no. 21[1]. Additionally Figure 5.3d shows the marginal probability for CWD_{ANN} whose setup (s.no. 20[2]) differs only in the third component of the climate state vector. The corresponding figures for T_{DJF} and T_{JJA} are not shown. Since differences to the presented marginal probabilities are hardly identifiable. Figure 5.4 presents the palaeoclimate reconstruction of the temperatures at the 850 hPa level. A detailed discussion why biome setup 4 instead of 3 is applied can be found in chapter 4.3 in the discussion of Figure 4.6. Also the connection of the 850 hPa to surface temperatures is described in that chapter.

In all figures a linear trend is estimated for several time slices for a better evaluation of climate changes. This trend is estimated on the numerical expectation value (red line, definition: Fig. 4.4 and chap. 3.3) of the marginal probabilities of each climatic variable c_i as also done in chapter 3.3. For the following time range trends are estimated whereby the most borders coincide with the LPAZ-borders which are defined by threshold exceedance of dominant taxa (Schiebel 2013):

time range 3 LPAZ 7 to 5, in detail pollen sample layer 1 to 11 which is equal to −55 y. cal BP $\overset{\tilde{\approx}}{\Leftrightarrow}$ 0 cm and 2 180 y. cal BP $\overset{\tilde{\approx}}{\Leftrightarrow}$ 255 cm

time range 2 LPAZ 4 to 2 directly before the layer with the jump in the biome probability profile (s. chap. 5.2.2, next to last paragraph), in detail pollen sample layer 12 to 28 which is equal to 2 613 y. cal BP $\overset{\tilde{\approx}}{\Leftrightarrow}$ 299 cm and 9 507 y. cal BP $\overset{\tilde{\approx}}{\Leftrightarrow}$ 703 cm

---

[1] s.no. 21 is equal to biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset

[2] s.no. 20 is equal to biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, CWD_{ANN})^T$ and CRU TS 3.1 as estimation dataset

time range 1 directly subsequent the next layer in LPAZ 2 to 1, in detail pollen sample layer 29 to 44 which is equal to 13 984 y. cal BP $\overset{\approx}{\Leftrightarrow}$ 728 cm and 31 767 y. cal BP $\overset{\approx}{\Leftrightarrow}$ 1 090 cm

These three time ranges (numbers from past to present) differ from the zones mentioned in the discussion of the biome probability profile in chapter 5.2.2. But these zones are identifiable in all palaeoclimate reconstructions since the marginal probability profile before and after the so named transition zone between time range 1 and 2 differs sometimes more sometimes less.

In general the results of the palaeoclimate reconstruction of the surface climate (Fig. 5.3) are in accordance with the results of Schiebel (2013), despite the fact that reconstruction bases only on six selected taxa. In detail Schiebel (2013, p. 6, chap. 2.1, para. 2; p. 15 chap. 3.3 para. 2; p. 72 chap. 7 para. 2) described a change to colder and dryer environmental conditions after the Holocene with the onset of the Younger Dryas (YD). For the CWD_{ANN}-reconstruction this matches also since $CWD_{ANN} > 0$ is equal to a dry environment and therefore a shift to higher values is equal to a dryer environment. More precisely the accordance means that there is a slight climate change signal in LPAZ 1 compared to the rest of the profile which is identifiable in all four subfigures of Figure 5.3.

Table 5.1.: Results for the parameters of the estimated linear trends in Figures 5.3 and 5.4.

$c_i = a \cdot t + b$		$c_i = T_{DJF}$ Fig 5.3a	$c_i = T_{JJA}$ Fig 5.3b	$c_i = P_{ANN}$ Fig 5.3c	$c_i = CWD_{ANN}$ Fig 5.3d	$c_i = T_{DJF}^{850hpa}$ Fig 5.4a	$c_i = T_{JJA}^{850hpa}$ Fig 5.4b	$c_i = P_{ANN}$ Fig 5.4c	$c_i = CWD_{ANN}$ Fig 5.4d
t. r. 3 (R-B 1, o.)	a	-0.0002	-0.0001	-0.0090	0.0025	0.0000	0.0000	-0.0144	0.0019
	$\pm\sigma$	0.0002	0.0001	0.0120	0.0086	0.0000	0.0001	0.0315	0.0087
	b	12.9	24.7	378.1	971.1	5.2	20.8	436.7	1013.1
	$\pm\sigma$	0.3	0.1	14.7	10.5	0.0	0.1	38.7	10.7
t. r. 2 (R-B 2)	a	0.0000	0.0000	-0.0088	0.0128	0.0000	0.0000	-0.0061	0.0013
	$\pm\sigma$	0.0001	0.0000	0.0040	0.0046	0.0000	0.0000	0.0163	0.0015
	b	13.4	23.9	477.0	829.3	4.7	20.3	407.8	993.0
	$\pm\sigma$	0.5	0.3	25.1	28.7	0.2	0.2	101.5	9.3
t. r. 1 (R-B 3, u.)	a	-0.0001	0.0000	-0.0009	-0.0012	0.0000	0.0000	-0.0026	0.0000
	$\pm\sigma$	0.0000	0.0000	0.0003	0.0003	0.0000	0.0000	0.0008	0.0000
	b	11.0	26.7	251.5	1069.4	4.7	21.6	336.6	987.3
	$\pm\sigma$	0.6	0.2	6.4	8.6	0.0	0.2	19.3	0.0

In detail all slopes a of the estimated linear trends in Figure 5.3 ($c_i = a \cdot t + b$) are small with a maximum value of $0.0128 \pm \sigma = 0.0046$ for the time range 2 in Fig. 5.3b so that an acceptable approximation is to consider only the intercept b as a representation of the expectation value in each time range or simple mean of the expectation values in each layer (red line) for a simple hypothesis testing. The result of the linear fits are summarized in Tab. 5.1.

The already mentioned gaps in the $W_{B_{l_0}}$-profile in chapter 5.2.2 are identifiable in all marginal probabilities in Figure 5.3 by a wider probability range resp. larger variance. Indicated is this by larger interquantile distances in the figures and also with jumps in the expectation value (red line). These are much smaller than the jumps in the $W_{B_{l_0}}$ profile. Best visible is this in the marginal probabilities of the water variables in Fig. 5.3c and 5.3d at about 1 700 y. cal BP, 6 300 y. cal BP and 8 500 y. cal BP. This is in accordance with the discussed position of the jumps in $W_{B_{l_0}}$. The influence on the reconstructed surface temperatures T_{DJF} and T_{JJA} is small with $\Delta T \approx \max 0.5\,\mathrm{K}$) and slopes $a \approx 0$. The influence on the reconstructed water variables P_{ANN} and CWD_{ANN} is larger since the estimated linear trends here have slope values with values different from zero (Tab. 5.1).

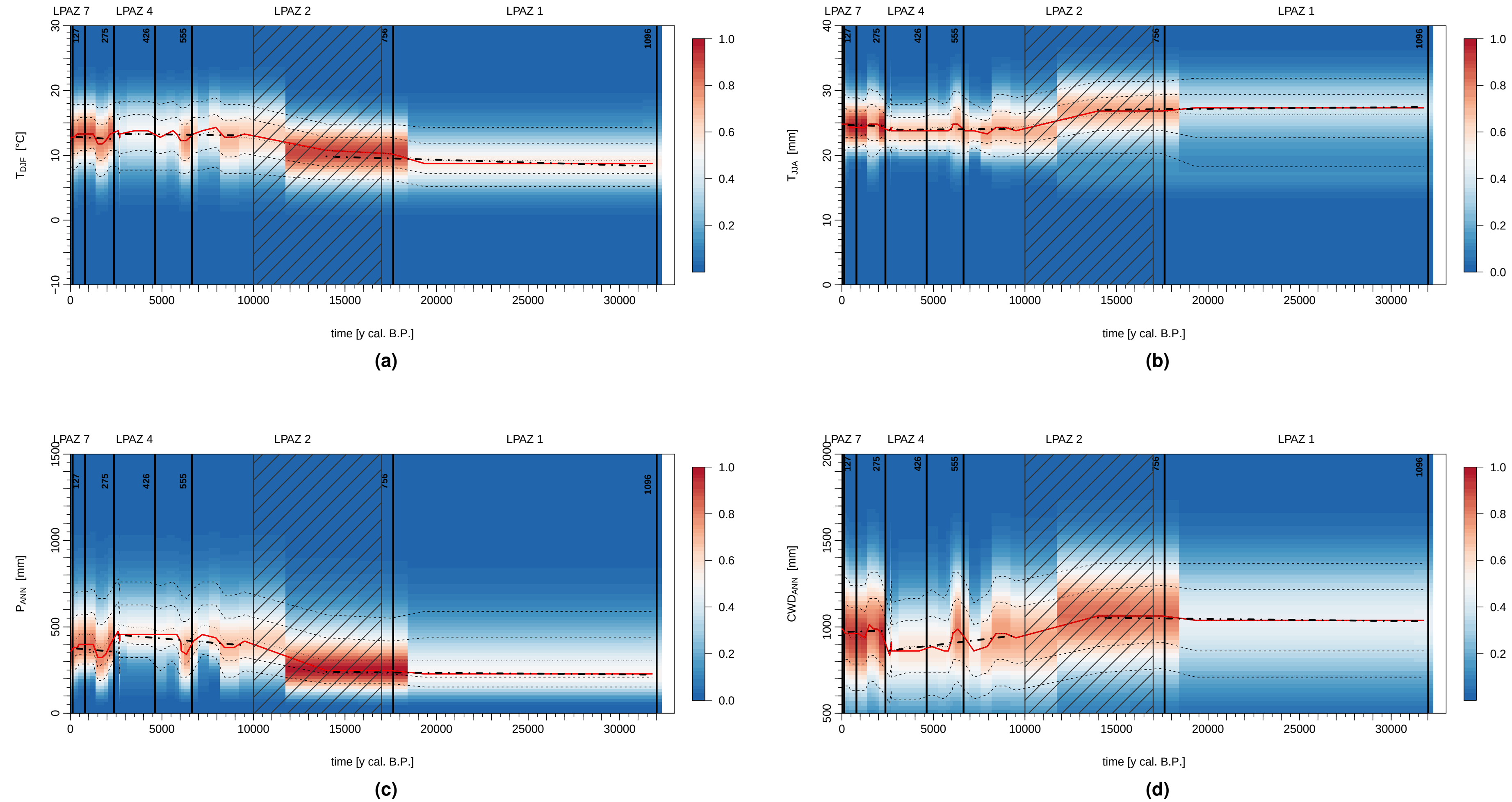
Another result which holds for all four marginal probabilities $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ is the lack of variability in time range 1. This is in accordance with the applied indicator taxa (chap. 5.2.3) since only two taxa (*Artemisia* ⇔ Arte_BR and *Plantago* ⇔ PlSM_BR) are regarded as present and contribute to the palaeoclimate reconstruction. Only the top layer of time range 1 is different: *Quercus ithaburensis type* ⇔ Qud_BR contributes to the result with larger values in the marginal probabilities $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ (more red colours) but not with discernible effects on the rest.

Also indicated in all four figures is that there is a slight, as already mentioned, climate change at the transition zone. This change is identifiable by gradients in all curves in Figure 5.3. The confidence in this climatic change depends on the reliability in this reconstruction in general (only six taxa). But it is also possible to apply a Student's t-test with a significance level of 0.05 as also done in chapter 3.3. This is done with R[3] for the mean of all expectation values (red line) in time range 1 versus time range 3 with null hypothesis that the means are equal and results in p-values between 1×10^{-9} and 1×10^{-16}. This is equal to a rejection (otherwise acceptance) of the null hypothesis so that the hypothesis test supports the assumption of climate change. If only the figures are considered the extrapolation of the estimated linear trend from time zone 1 into time zone 2 indicates that the estimated lines lie in all four figures outside of the 25 % resp. 75 % quantile curves. This is not the case for the discussed jumps in the CWD_{ANN} reconstruction between the time range from 11 000 y. cal BP till 6 000 y. cal BP and for a comparison of time range 2 with time range 3. This fact also supports the climate change hypothesis. Thereby the greatest changes are identifiable for the P_{ANN} reconstruction, best demonstrated by the intercept value b since it is almost doubled. The relative changes of b for the other climate variables are in the same order of magnitude: 120 % for T_{DJF}, 90 % for T_{JJA} and 75 % for CWD_{ANN}.

Between time range 2 and 3 there is slight difference in the reconstructions but not as distinctive as for time range 1 and 2. If the jumps are left out than the main differences are lower values in time range 2 than in 3 for all marginal probabilities presented in Figure 5.3 (less red

---

[3] in more detail with the R-function *t.test()* and its default setup (two-sided, no equal variances) of the base package *stats* of R version 3.0.1, s. Team (2013) for more details



(a) (b) (c) (d)

Figure 5.3.: Marginal posterior probability $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN} for the Birkat Ram palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$. (a), (b) and (c) base on s.no. 21 (biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and CRU TS 3.1 as estimation dataset) and (d) on s.no. 20 which is equal to s.no. 21 and differs only in the definition of the climate state vector ($\vec{C} = (T_{DJF}, T_{JJA}, CWD_{ANN})^T$). The mapping of the figure is identical to Figure 4.4.

colours in the image plot).

The temperature reconstructions additionally shows no substantial difference of the intercepts ($b_{t.r.2} \in [b_{t.r.3} - \sigma; b_{t.r.3} + \sigma]$). The only variation in the red lines are a result of $W_{B_{l_0}}$ with, as just mentioned, no substantial effect on the trends. Only the jump at about 8 500 y. cal BP could be interpreted as not human made (s. discussion of $W_{B_{l_0}}$), but it is not distinguishable from the others. Therefore there is no evidence for a temperature change throughout the Holocene.

The situation for the marginal probability profiles for P_{ANN} and CWD_{ANN} is different but also similar: Similar are the circumstances for 8 500 y. cal BP. It is also not distinguishable from the jumps which are regarded anthropogenic. But the jumps in the expectation value (red line) are comparatively large ($\approx 10\%$ relative change) to the temperature ones ($\approx 3\%$). Due to that fact there is a difference in the estimated linear trends between time range 2 and 3. The Student's t-test with the same setup as the last one results in p-values of 2.3×10^{-4} (P_{ANN}) and 2.3×10^{-5} (CWD_{ANN}) which are five orders of magnitude smaller than those of the last one. Additionally if the whole marginal probability profiles in time range 2 and 3 are considered without the jumps, more precisely if the quantile curves are considered, then the difference in the P_{ANN} reconstruction between these two time ranges are in the order of magnitude of $\approx \max 50 \, \frac{\text{mm}}{\text{year}}$. Thereby the driest part is the top layer and the wettest at the top of LPAZ 4 with an increase in between (equivalent becoming slightly more humid) and a slight decrease after the maximum. For CWD_{ANN} the difference is larger ($\approx \max 150 \, \frac{\text{mm}}{\text{year}}$) between the time ranges. And compared to P_{ANN} time range 2 is separated by distinctive gradients at the borders. The final conclusion for the P_{ANN} and also for the CWD_{ANN} reconstruction is drawn after the discussion of Figure 5.4.

The result for the palaeoclimate reconstructions of the temperatures at the 850 hPa level in Figure 5.4 is different. The ordinate scale in Figure 5.4 is the same as in 5.3 so that one main difference between the 850 hPa temperature reconstruction and those of surface values becomes obvious: A narrow profile for $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ quantifiable in the distance between e. g. the 10% and 90% quantile curves: $T_{DJF}^{850hpa} \approx \max 6\,\text{K}$ and $T_{DJF} \approx \max 11\,\text{K}$, $T_{JJA}^{850hpa} \approx \max 6\,\text{K}$ and $T_{JJA} \approx \max 11\,\text{K}$. The distance for $P_{ANN} \approx \max 450 \, \frac{\text{mm}}{\text{year}}$ and $CWD_{ANN} \approx \max 800 \, \frac{\text{mm}}{\text{year}}$ is the same. The next difference for the marginal profiles $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ is that for CWD_{ANN} and T_{DJF}^{850hpa} compared to the according subfigures in Figure 5.3 there is no more climate change detectable neither at the transition zone nor between time range 2 and 3: All differences for b are within the standard deviation σ (Tab. 5.1) and the quantile curves show almost no variation throughout the profile.

The T_{JJA}^{850hpa} and P_{ANN} result in Figure 5.4b and 5.4c indicate that there is a climate change. The T_{JJA}^{850hpa} result is in accordance with the reconstruction of the surface temperature T_{JJA} since it shows also a warming with the onset of the YD. Thereby the p-value for the Student's t-test for time range 1 versus 2 is in the same order of magnitude (1×10^{-16}). For P_{ANN} the result is a little bit different since the climate change is visible but much less pronounced as

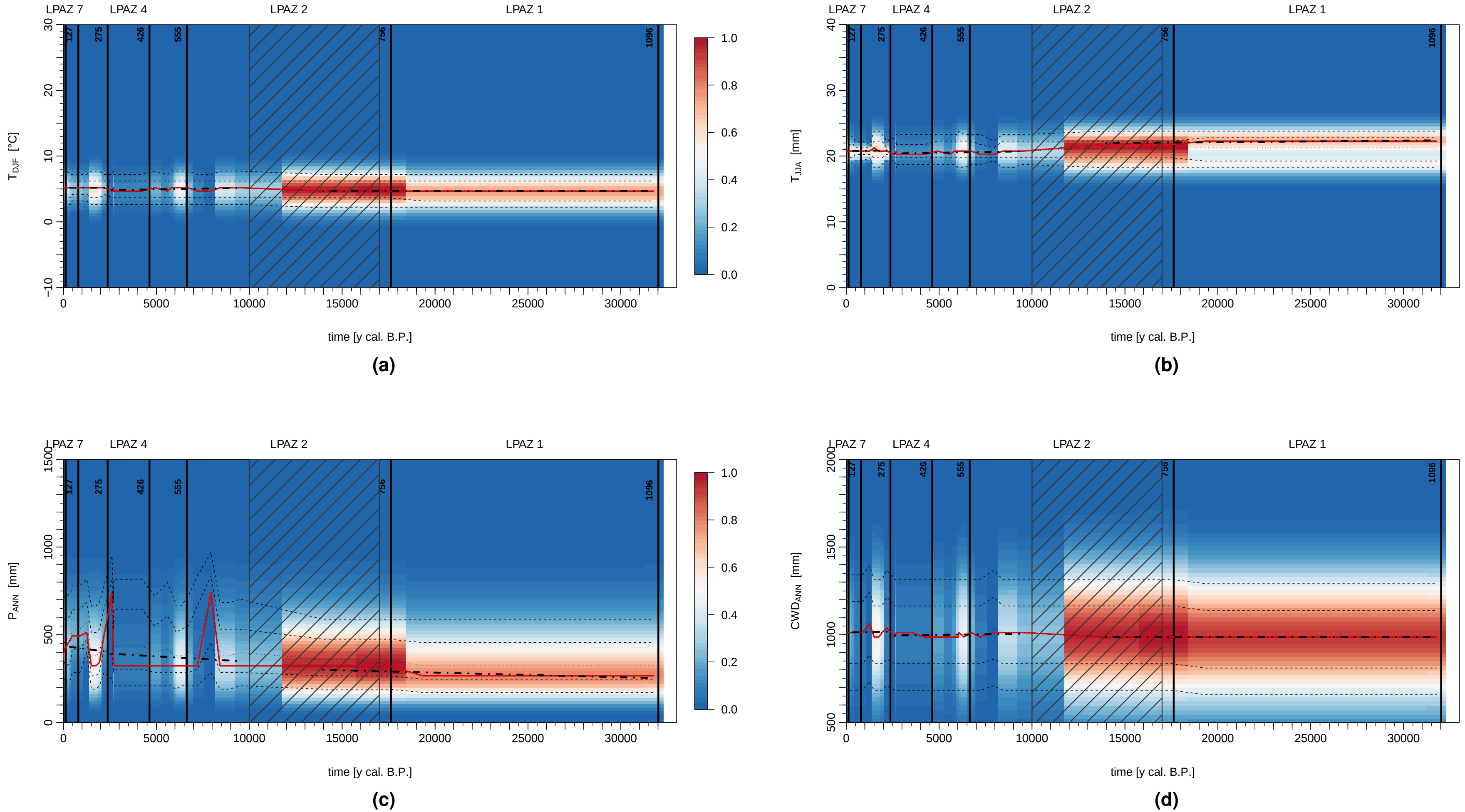


Figure 5.4.: Marginal posterior probability $\mathbb{P}_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN} and CWD_{ANN} for the Birkat Ram palaeoclimate reconstruction. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$. (a), (b) and (c) base on s.no. 56 (biome setup 4, $\vec{C} = \left(T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}\right)^T$ and CRU TS 3.1 as estimation dataset) and (d) on s.no. 55 which is equal to s.no. 56 and differs only in the definition of the climate state vector ($\vec{C} = \left(T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, CWD_{ANN}\right)^T$). The mapping of the figure is identical to Figure 4.4.

in the palaeoclimate reconstruction estimated with the transfer function of the surface values: the p-value is only 0.0099. The intercepts $b \pm \sigma$ differ but only seen from time range 1. But compared to T_{DJF}^{850hpa} and CWD_{ANN} there is a variability identifiable. The shape of the expectation value curve (red line) for P_{ANN} here shows also jumps like the P_{ANN} curve in Figure 5.4c in the same pollen sample layers or resp. time slices, whereas two of them are large. This is caused by Mediterranean biome probability values of almost one ($W_{B_{l_0=1}} \approx 1$) together with fact that *Quercus ithaburensis type* $\Leftrightarrow$ Qud_BR is regarded as present in these pollen sample layers. This taxon induced these problems only in the reconstruction of the temperature at the 850 hPa level but it is too important to omit it in the complete reconstruction. An outcome which is obtained in all components of the reconstruction in Figure 5.4 is that there is no difference visible between time range 1 and 2 if the jumps are excluded (anthropogenic or not) in this consideration. Overall it is important to note that the palaeoclimate reconstruction of P_{ANN} and CWD_{ANN} of Birkat Ram in general are very sensitive to the setup of the climate state vector $\vec{c}$, more precisely: What are the other components (middle troposphere temperatures (850 hPa level) or surface temperatures)?

In summary it is noted that the climate change to higher summer temperatures (T_{JJA} and T_{JJA}^{850hpa}) with the onset of the YD is very reliable so that this hypothesis from Schiebel (2013) is confirmed. The hypothesis of dryer environmental conditions could be only partly confirmed since only the palaeoclimate reconstruction applied with the surface based transfer functions clearly provide this. For the BITBM as a whole it becomes also obvious that it is not as robust against anthropogenic outliers/influences as BITM but not so sensitive as the BBM. But for a reliable conclusion the BITBM has to be applied at a sediment core with more taxa.

6

Jordan Valley

Motivated by the successful climate field reconstruction (CFR) of Gebhardt (2003)/Gebhardt *et al.* (2008) and Simonis (2009)/Simonis *et al.* (2012) which allows a better assessment of possible climate changes in the past this chapter presents a simplified version of their approach for the Jordan Valley. In this chapter the work of Gebhardt (2003)/Gebhardt *et al.* (2008) and Simonis (2009)/Simonis *et al.* (2012) is namend the *previous work*. The basic hypothesis in their approach is the possibility to describe the past climate $\vec{c}_0$ as deviation $\vec{c}\,'$ from the modern climate state $\vec{c} := \vec{c}^{\,mod}$:

$$\vec{c}\,' := \vec{c}_0 - \vec{c} \tag{6.1}$$

Past climate in this context means to evaluate the climate state of a certain defined time slice at all considered locations.

First the Jordan Valley is defined as the analysis area. The following section describes the simplifications of mathematics of the previous work. Directly follows a discussion of the adaptations of the local palaeoclimate reconstructions whereat the results are discussed in detail in chapter 4 resp. 5. The next section describes the selection of contemplable time slices for the CFR. The last part presents the results.

6.1. Jordan Valley - Definition of the analysis ”Area”

The climate data for near surface temperatures T_{DJF} and T_{JJA} and the precipitation P_{ANN} is available on a regular latitude-longitude grid. The location vector $\vec{r}_n$ is defined on the center of the grid boxes. The data is described in detail in chapter 2.1. In this chapter it is the aim to evaluate the palaeoclimate reconstructions along the Dead Sea Rift or more precisely along the Jordan Valley. Therefore it is necessary to specify this area by defining each coordinate combination of the location vector $\vec{r}_{n=1,\dots,N_{GB}}$ situated in the Jordan Valley. For this purpose the underlying orography is applied as shown in Figure 6.1.

Due to the fact that the resolution of the CRU TS 3.1 dataset is too coarse to represent the Jordan Valley (only six grid boxes) alone a mixture with another climate dataset is considered: the E-OBS dataset. Figure 6.1 shows both: the CRU TS 3.1 and E-OBS orography as well as the respectively defined Jordan Valley grid boxes. The later on applied discretisation of the

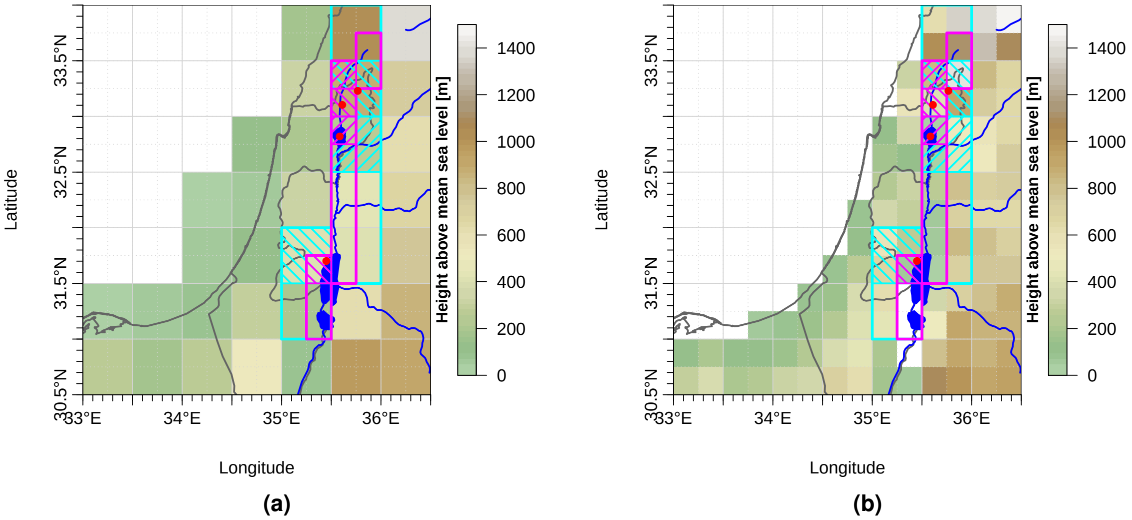


Figure 6.1.: Definition of the Jordan Valley location vector.
(a) shows the CRU TS 3.1 and (b) the E-OBS orography. The sediment cores are marked with red dots. The grid boxes which define the Jordan Valley location vector are marked in cyan colour for the definition based on CRU TS 3.1 and in magenta colour for the E-OBS definition. The hatched boxes mark the assigned boxes of the sediment cores.

Jordan Valley is the E-OBS definition (marked with magenta colour in Fig. 6.1) resulting in $N_{GB} = 13$ grid boxes whereupon the numbering starts with $n = 1$ at the Jordan estuary.

The climate variable CWD_{ANN} is not used in this CFR. There are two reasons for that: One is that for E-OBS there is no equivalent data product available. The other is that especially for Birkat Ram the discussion of the palaeoclimate reconstruction of the surface climate together with the middle troposphere enlightens that this component shows no evidence for a possible climate change.

Just described was the definition of the Jordan Valley. The discretisation of the climate variables as well as the later on presented discrete cost function is analogously defined as in the previous work: They summarized the whole climate state field which is later on applied to one vector $\vec{c}$:

$$\vec{c} = \left({}^{1}T_{DJF} \cdots {}^{n}T_{DJF} \cdots {}^{N_{GB}}T_{DJF} \quad {}^{1}T_{JJA} \cdots {}^{n}T_{JJA} \cdots {}^{N_{GB}}T_{JJA} \quad {}^{1}P_{ANN} \cdots {}^{n}P_{ANN} \cdots {}^{N_{GB}}P_{ANN} \right)^{T} \tag{6.2}$$

If the climate state at location $\vec{r}_n$ is required this is equivalent of using ${}^{n}\vec{c}$ defined as:

$${}^{n}\vec{c} = \vec{c}\,(\vec{r}_n) = \begin{pmatrix} {}^{n}T_{DJF} \\ {}^{n}T_{JJA} \\ {}^{n}P_{ANN} \end{pmatrix} = \left(0 \cdots {}^{n}T_{DJF} \cdots 0\, 0 \cdots {}^{n}T_{JJA} \cdots 0\, 0 \cdots {}^{n}P_{ANN} \cdots 0 \right)^{T} \tag{6.3}$$

The index notation with $^n\ldots = \ldots(\vec{r}_n)$ and the discretisation is also defined for the other applied variables $\vec{c}\,'$ and $\vec{c}_0$. Figure 6.2 shows one of the resulting profiles for the used climate variables in this approach: T_{DJF}. The abscissa represents the aforementioned Jordan Valley grid boxes. The blue points represent the available data at each grid point. The vertical black lines mark the locations of the sediment cores Ein Gedi, Lake Kinneret and Birkat Ram which are also named fossil sites in this chapter. In this approach it is assumed that this profile could be described by a simple GLM with quadratic term. This GLM is estimated independently for each component of the climate state vector with the statistical software R[1]. The result which takes into account the inverse CDF method (chap. 6.2.3) is indicated by the dashed black line in Figure 6.2. The solid black lines are the 99.5 % confidence bands for the estimated GLM. The R package *visreg*[2] provides this functionality.

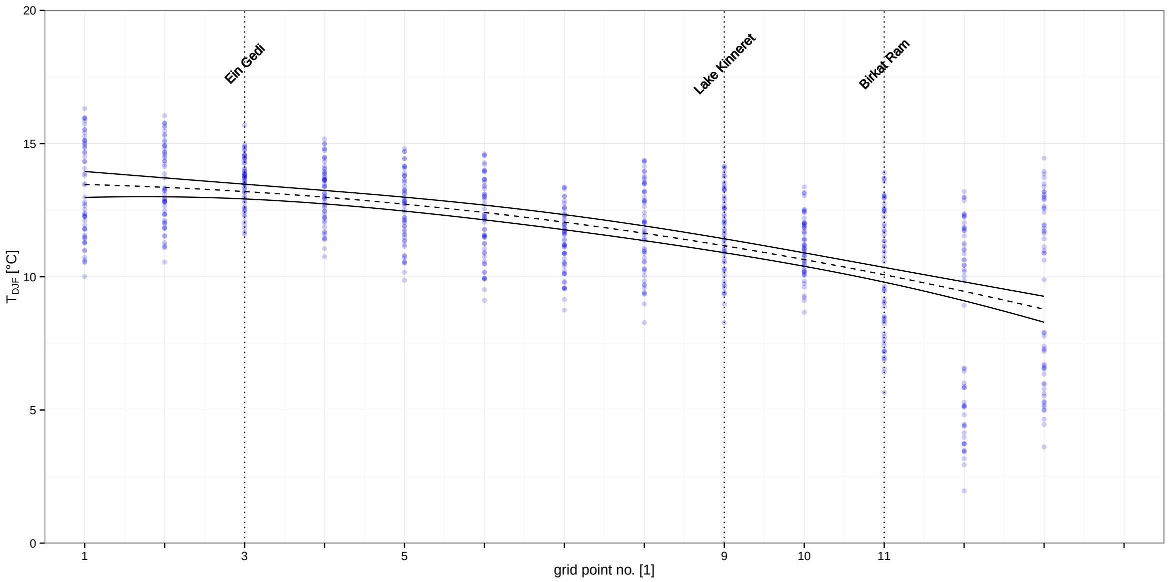


Figure 6.2.: The climate profile in the Jordan Valley for T_{DJF}.
The abscissa represents the Jordan Valley grid boxes. The blue points represent the available data at each grid point (E-OBS and CRU TS 3.1). The vertical black lines mark the locations of the sediment cores Ein Gedi, Lake Kinneret and Birkat Ram. The dashed black line indicates the result of a GLM with quadratic term which takes into account the inverse CDF method discussed in chap. 6.2.3 and the solid black lines are the corresponding 99.5 % confidence bands.

For the application in the later on presented CFR it is necessary to describe this profile for the whole Jordan Valley: The coefficients $\beta_{\ldots}$ which define the profile are known and therefore

---

[1] in more detail with the base package *stats* of R version 3.0.1, s. Team (2013) for more details, and also chap. 3.2.2

[2] in more detail with package version 2.2-2 (2016-02-05) s. Breheny, Burchett (2016) for more details

the expectation value state at each grid point $\vec{r}_n$ could be estimated via

$$ {}^{n}\vec{c} = \begin{pmatrix} {}^{n}T_{DJF} \\ {}^{n}T_{JJA} \\ {}^{n}P_{ANN} \end{pmatrix} = \begin{pmatrix} \beta_{a1} + \beta_{b1} \cdot n + \beta_{c1} \cdot n^2 + {}^{n}\varepsilon_{T_{DJF}} \\ \beta_{a2} + \beta_{b2} \cdot n + \beta_{c2} \cdot n^2 + {}^{n}\varepsilon_{T_{JJA}} \\ \beta_{a3} + \beta_{b3} \cdot n + \beta_{c3} \cdot n^2 + {}^{n}\varepsilon_{P_{ANN}} \end{pmatrix} . \tag{6.4} $$

This notation is refined with some additional definitions: First the coefficients are summarized in one column vector ($\equiv (\ldots)^T$):

$$ \vec{\beta} := \begin{pmatrix} \beta_{a1} & \beta_{a2} & \beta_{a3} & \beta_{b1} & \beta_{b2} & \beta_{b3} & \beta_{c1} & \beta_{c2} & \beta_{c3} \end{pmatrix}^T . \tag{6.5} $$

Secondly the definition of the location vector described by the grid box numbers (eq. 6.6a) and its component wise squared (eq. 6.6b):

$$ \vec{n} := \begin{pmatrix} 1 & \cdots & n & \cdots & N_{GB} \end{pmatrix}^T , \tag{6.6a} $$

$$ \vec{n^2} := \begin{pmatrix} 1^2 & \cdots & n^2 & \cdots & N_{GB}^2 \end{pmatrix}^T . \tag{6.6b} $$

And thirdly the analysis area matrix $\boldsymbol{R}$:

$$ \boldsymbol{R} := \begin{pmatrix} \vec{1} & \vec{0} & \vec{0} & \vec{n} & \vec{0} & \vec{0} & \vec{n^2} & \vec{0} & \vec{0} \\ \vec{0} & \vec{1} & \vec{0} & \vec{0} & \vec{n} & \vec{0} & \vec{0} & \vec{n^2} & \vec{0} \\ \vec{0} & \vec{0} & \vec{1} & \vec{0} & \vec{0} & \vec{n} & \vec{0} & \vec{0} & \vec{n^2} \end{pmatrix} \tag{6.7} $$

Thereby $\vec{0}$ and $\vec{1}$ are column vectors of length N_{GB} with only zeros and resp. ones as entries. $\boldsymbol{R}$ summarizes the definition of the analysis area. With this definitions it is now possible to rewrite equation 6.4 for the whole climate state vector of the Jordan Valley (eq. 6.2) with

$$ \vec{c} := \boldsymbol{R} \cdot \vec{\beta} + \vec{\varepsilon} . \tag{6.8} $$

$\vec{\varepsilon}$ is analogously defined to $\vec{c}$ and summarizes the error ε in equation 6.4 at each grid point and is later on discussed in detail in chapter 6.2.1.

6.2. The Cost Function $\mathcal{J}$

The presented approach in this study is a simplified version of the CFR approach presented in the previous work which bases on the minimization of a cost function $\mathcal{J}$:

$$ \mathcal{J} = \mathcal{J}_{veg} + \mathcal{J}_{Mod} . \tag{6.9} $$

This cost function summarizes the contributions of several paleoclimate reconstructions at different locations and one time slice in a vegetational part $\mathcal{J}_{veg}$ and a model part in $\mathcal{J}_{Mod}$. This is a common approach in Meteorological data assimilation. Its result is a CFR which fits best to the palaeoclimate reconstruction at all locations. Its mathematical concept is presented in details in Gebhardt (2003). In this study the focus is set on the required adaptation for the quasi linear profile which is used as model part. Therefore the starting point in this study is the already discretised cost function.

6.2.1. The Model Cost Function $\mathcal{J}_{Mod}$

The basic hypothesis in this approach and also in the previous work is that it is always possible to describe the past climate $\vec{c}_0$ as deviation $\vec{c}\,'$ from the modern climate state $\vec{c}$. In this work this hypothesis is extended to the assumption that the profile in the past time frame also follows a quasi linear profile due to the fact that this deviation is small. Gebhardt (2003, chap. 6.1) also assumes that this deviation is small, a fact which is supported by the local reconstructions presented in this work. This requires first to rewrite equation 6.1 with the help of equation 6.8:

$$\begin{aligned} \vec{c}\,' &= \overbrace{\boldsymbol{R}\cdot\vec{\beta}_0+\vec{\varepsilon}_0}^{=:\vec{c}_0} - \overbrace{\left(\boldsymbol{R}\cdot\vec{\beta}+\vec{\varepsilon}\right)}^{=:\vec{c}} && \text{(6.10a)}\\ &= \boldsymbol{R}\cdot\underbrace{\left(\vec{\beta}_0-\vec{\beta}\right)} + \underbrace{\left(\vec{\varepsilon}_0-\vec{\varepsilon}\right)} && \text{(6.10b)}\\ &= \boldsymbol{R}\cdot\vec{\beta}\,' \qquad + \vec{\varepsilon}\,' && \text{(6.10c)} \end{aligned}$$

Equation 6.10c expresses the climate state deviation dependent on the quasi linear profile deviation $\vec{\beta}\,'$. The second part is to express this (deviation small and quasi linear profile) mathematically. This is done by

$$\vec{\beta}\,'\cdot\vec{\beta}\,'^{T} = 0 \Leftrightarrow \left(\vec{\beta}_0-\vec{\beta}\right)\cdot\left(\vec{\beta}_0-\vec{\beta}\right)^{T} = 0 \qquad (6.11)$$

This method is called Ridge Regression and therefore the parameter $\vec{\varepsilon}\,'$ which summarizes the errors is set to zero. The model part following Gebhardt (2003) is then given by

$$\mathcal{J}_{Mod}\left(\vec{\beta}_0\right) = \gamma_M\cdot\left(\vec{\beta}_0-\vec{\beta}\right)\cdot\left(\vec{\beta}_0-\vec{\beta}\right)^{T} \qquad (6.12)$$

The factor γ_M is a weighting factor based on several minimization runs. The applied γ_M later on is the one which fulfils $\frac{\mathcal{J}_{veg}}{\gamma_M\mathcal{J}_{Mod}} \approx 1$ at the cost function minimum $\mathcal{J}_R$. In other words both parts are equally weighted at the minimum. This as also done in the previous work. The used values γ_M for the palaeoclimate reconstruction in this work can be found in Table E.1 in appendix E.1.

6.2.2. The Vegetational Cost Function $\mathcal{J}_{veg}$

The discretised cost function applied in Gebhardt (2003, eq. 6.35) transferred to the nomenclature of this study is defined by

$$\mathcal{J}_{veg} = -\sum_{s=1}^{S}\left(\ln\left[{}^{s}f_{\vec{C}|\vec{P}}\left(\overrightarrow{{}^{s}c}_{0}\,|\,\vec{p}\right)\right]\Delta\sin\phi_{s}\right) \tag{6.13}$$

In equation 6.13 the evaluation of each palaeoclimate reconstruction ${}^{s}f_{\vec{C}|\vec{P}}\left(\ldots\right)$ is written with dependence on the past climate $\overrightarrow{{}^{s}c}_{0}$ at each fossil site $\vec{r}_{s}$. To write it dependent on the whole field the previous work used a simple matrix multiplication:

$$\overrightarrow{{}^{s}c}_{0} = {}^{s}\boldsymbol{P}\cdot\vec{c}_{0} = {}^{s}\boldsymbol{P}\cdot\left(\boldsymbol{R}\cdot\vec{\beta}_{0}\right). \tag{6.14}$$

The matrix ${}^{s}\boldsymbol{P}$ contains only zeros and ones at the position of the fossil sites. The computation of $\Delta\sin\phi_{s}$ is not changed compared to the previous work and is the difference between the sine of the latitude of the northern and the southern grid box boundary.

Also mentioned in the previous work is the problem that the palaeoclimate reconstructions are estimated with data gained at the height of the fossil site h_{site}. But the quasi linear profile is estimated at the mean grid box height h_{mean}, so that strictly spoken the current form of the vegetational cost function (eq. 6.13) is evaluated with the past climate $\overrightarrow{{}^{s}c}_{0,h_{site}}$. Therefore the previous work applied a simple height correction for temperature based on the temperature gradient of the standard atmosphere ($\gamma_{d} = 6.5\times10^{-3}\frac{K}{m}$)[3] by a linear extrapolation. For precipitation no height correction is reasonable so that with ${}^{s}P_{ANN,0,h_{site}} \equiv {}^{s}P_{ANN,0,h_{mean}}$ the past climate state is estimated as defined by equation 6.15a (equal for the present climate state).

$$\overrightarrow{{}^{s}c}_{0,h_{site}} = \begin{pmatrix} {}^{s}T_{DJF} \\ {}^{s}T_{JJA} \\ {}^{s}P_{ANN} \end{pmatrix}_{0,h_{site}} = \overbrace{\begin{pmatrix} {}^{s}T_{DJF} \\ {}^{s}T_{JJA} \\ {}^{s}P_{ANN} \end{pmatrix}_{0,h_{mean}}}^{=\,\overrightarrow{{}^{s}c}_{0,h_{mean}}} + \overbrace{\gamma_{d}\left({}^{s}h_{site} - {}^{s}h_{mean}\right)}^{=:\,{}^{s}\Delta T}\overbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}^{=:\,\vec{e}_{h}} \tag{6.15a}$$

$$= {}^{s}\boldsymbol{P}\left(\boldsymbol{R}\vec{\beta}_{0}\right) + {}^{s}\Delta T\;\vec{e}_{h} \tag{6.15b}$$

Equation 6.15b is derived by inserting equation 6.1, 6.10c and 6.14 into 6.15a. The discrete vegetational cost function $\mathcal{J}_{veg}$ dependent on quasi linear profile for the past time slice $\vec{\beta}_{0}$ is then given by:

$$\mathcal{J}_{veg}\left(\vec{\beta}_{0}\right) = -\sum_{s=1}^{S}\left(\ln\left[{}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\boldsymbol{P}\left(\boldsymbol{R}\vec{\beta}_{0}\right) + \gamma_{d}\left({}^{s}h_{site} - {}^{s}h_{mean}\right)\vec{e}_{h}\,|\,\vec{p}\right)\right]\cdot\Delta\sin\phi_{s}\right). \tag{6.16}$$

---

[3] s. also chap. 4.2.3 below eq. 4.13 on p. 58

6.2.3. Inverse CDF Method for The Linear Profile

One problem remains which has already been addressed by Schölzel (2006): It has to be ensured that the minimum of $\mathcal{J}$ is estimated only for reasonable precipitation values ($P_{ANN} \geq 0$). Schölzel (2006) and Simonis (2009) solve this with the inverse CDF method (already mentioned in chap. 3.2.2). The inverse CDF for the climate state vector ${}^{n}\vec{c}$ in the Jordan Valley at location $\vec{r}_n$ is estimated with

$$
{}^{n}\vec{u} = \begin{pmatrix} {}^{n}u_1 \\ {}^{n}u_2 \\ {}^{n}u_3 \end{pmatrix} = \vec{\varphi}_{prior}\left({}^{n}\vec{c}\right) = \begin{pmatrix} \frac{{}^{n}T_{DJF} - \mu^{prior}_{T_{DJF}}}{\sigma^{prior}_{T_{DJF}}} \\ \frac{{}^{n}T_{JJA} - \mu^{prior}_{T_{JJA}}}{\sigma^{prior}_{T_{JJA}}} \\ F^{-1}_{\mathcal{N}(0,1)}\left(F_{\mathcal{G}(\nu^{prior},\lambda^{prior})}\left({}^{n}P_{ANN}\right)\right) \end{pmatrix} \tag{6.17}
$$

In this equation the values of the climate prior of the Jordan Valley (s. chap. 4.2.1 for details) are used for the transformation. This implicates the hypothesis that the climate prior is valid/applicable and identical for all past time slices. A fact which is already used in the palaeoclimate reconstructions presented in chapter 4 and 5. The GLM result for T_{DJF} indicated in Figure 6.2 already takes the inverse CDF method into account.

All considerations for the quasi linear profile in chapter 6.1, 6.2.1 and 6.2.2 for $\vec{c}_{0,h_{mean}}$ are also valid for $\vec{u}_{0,h_{mean}}$. If $\vec{u}_{0,h_{mean}} := \boldsymbol{R} \cdot \vec{\beta}_0$ replaces equation 6.10 then ${}^{s}\vec{c}_{0,h_{site}}$ is given by

$$
{}^{s}\vec{c}_{0,h_{site}} = \vec{\varphi}^{-1}_{prior}\left({}^{s}\boldsymbol{P}\left(\boldsymbol{R}\vec{\beta}_0\right)\right) + {}^{s}\Delta T\,\vec{e}_h \tag{6.18}
$$

This equation also shows why it is necessary to write the cost function dependent on $\vec{\beta}_0$ instead of $\vec{\beta}'$, as done in the previous work, since the necessary constraint

$$
q\left({}^{n}\vec{c}_0 - {}^{n}\vec{c}\right) = q\left({}^{n}\vec{c}_0\right) - q\left({}^{n}\vec{c}\right) \tag{6.19}
$$

is not fulfilled for $\vec{\varphi}_{prior}(\cdot)$ resp. $\vec{\varphi}^{-1}_{prior}(\cdot)$. The final version of the discrete cost function which is minimized with respect to $\vec{\beta}_0$ is than given by

$$
\begin{aligned}
\mathcal{J}\left(\vec{\beta}_0\right) \quad = \quad & -\sum_{s=1}^{S}\left(\ln\left[{}^{s}f_{\vec{C}|\vec{P}}\left(\vec{\varphi}^{-1}_{prior}\left({}^{s}\boldsymbol{P}\left(\boldsymbol{R}\vec{\beta}_0\right)\right) + \gamma_d\left({}^{s}h_{site} - {}^{s}h_{mean}\right)\vec{e}_h \middle| \vec{p}\right)\right] \cdot \Delta \sin\varphi_s\right) \\
& + \gamma_M \cdot \left(\vec{\beta}_0 - \vec{\beta}\right) \cdot \left(\vec{\beta}_0 - \vec{\beta}\right)^T .
\end{aligned} \tag{6.20}
$$

The minimum at $\vec{\beta}_{R,0}$ of this nine dimensional cost function $\mathcal{J}$ is calculated by finding the roots of the gradient $\vec{\nabla}_{\vec{\beta}_0}\mathcal{J}\Big|_{\vec{\beta}_0 = \vec{\beta}_{R,0}} = 0$ (necessary condition) and checking the definiteness of the Hessian matrix $\boldsymbol{\mathcal{H}}_{\vec{\beta}_0}(\mathcal{J})$ (sufficient condition). In detail this means that the Hessian matrix has to be positive definite for $\vec{\beta}_{R,0}$.

The j component of this gradient is defined by the partial derivative of $\mathcal{J}$ with respect to j-th element of $\vec{\beta}_0 =: \beta_j$:

$$\begin{aligned}\frac{\partial \mathcal{J}}{\partial \beta_j} \quad = \quad & -\sum_{s=1}^{S}\left\{\left[{}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)\right]^{-1}\cdot\frac{\partial\, {}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)}{\partial\, {}^{s}\vec{c}_{0,h_{site}}}\cdot\frac{\partial \vec{\varphi}^{\,-1}_{prior}\left({}^{s}\vec{u}_{0,h_{mean}}\right)}{\partial\, {}^{s}\vec{u}_{0,h_{mean}}}\cdot \boldsymbol{S}\vec{e}_{\beta_j}\right\} \\ & +\gamma_M\cdot 2\left(\beta_{0j}-\beta_j\right).\end{aligned} \tag{6.21}$$

At this juncture the notation is shortened by the applying equation 6.18 (${}^{s}\vec{c}_{0,h_{site}} = \ldots$). Additionally two shortcuts are defined: First the product ${}^{s}\boldsymbol{PR}$ defines the matrix $\boldsymbol{S}$. Second $\vec{e}_{\beta_j}$ is a column vector with the same dimension as $\vec{\beta}_0$ in which only the j-th component is one and the rest is zero.

The i-th and j-th element of the Hessian matrix $\boldsymbol{\mathcal{H}}_{\vec{\beta}_0}(\mathcal{J})$ is defined by the second derivative $\frac{\partial^2}{\partial\beta_i\partial\beta_j}$:

$$\begin{aligned}\frac{\partial^2 \mathcal{J}}{\partial \beta_i \partial \beta_j} \quad = \quad & \sum_{s=1}^{S}\left\{\left[{}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)\right]^{-2}\cdot\left[\frac{\partial\, {}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)}{\partial\, {}^{s}\vec{c}_{0,h_{site}}}\cdot\frac{\partial \vec{\varphi}^{\,-1}_{prior}\left({}^{s}\vec{u}_{0,h_{mean}}\right)}{\partial\, {}^{s}\vec{u}_{0,h_{mean}}}\cdot \boldsymbol{S}\vec{e}_{\beta_j}\right]^2\right. \\ & -\left[{}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)\right]^{-1}\cdot\frac{\partial^2\, {}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)}{\partial\left({}^{s}\vec{c}_{0,h_{site}}\right)^2}\cdot\left[\frac{\partial \vec{\varphi}^{\,-1}_{prior}\left({}^{s}\vec{u}_{0,h_{mean}}\right)}{\partial\, {}^{s}\vec{u}_{0,h_{mean}}}\cdot \boldsymbol{S}\vec{e}_{\beta_j}\right]^2 \\ & \left.-\left[{}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)\right]^{-1}\cdot\frac{\partial\, {}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)}{\partial\, {}^{s}\vec{c}_{0,h_{site}}}\cdot\frac{\partial^2 \vec{\varphi}^{\,-1}_{prior}\left({}^{s}\vec{u}_{0,h_{mean}}\right)}{\partial\left({}^{s}\vec{u}_{0,h_{mean}}\right)^2}\cdot\left[\boldsymbol{S}\vec{e}_{\beta_j}\right]^2\right\} \\ & +\gamma_M\cdot 2\delta_{ij}\end{aligned} \tag{6.22}$$

The main problem remaining is that the derivatives of $\vec{\varphi}^{\,-1}_{prior}\left({}^{s}\vec{u}_{0,h_{mean}}\right)$ and ${}^{s}f_{\vec{C}|\vec{P}}\left({}^{s}\vec{c}_{0,h_{site}}\right)$ are not analytically evaluable. Due to that fact the minimization is executed numerically with the R-routine optim. In details the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm for box-constrained optimization (L-BFGS-B)[4] is applied to estimate the minimum of $\mathcal{J}$ (eq. 6.20) without the factor $\Delta \sin\varphi_s$. On the one hand the difference between the three involved factors is negligible. On the other hand as shown by equation 6.21 the derivation of this factor yields no contribution to the gradient. Since the numerical estimation of the minimum also applies the gradient this factor is omitted.

Due to numerical uncertainties it is also necessary to shrink the minimization to values of $\vec{\beta}_0$ which result in physically reasonable values. This is done by restricting the range for $\vec{\beta}_0$ with the upper and lower limit of

$$\vec{\beta}_{upper,0} = \begin{pmatrix}9.000 & 9.000 & -3.400 & 0.500 & 0.500 & 0.010 & 0.025 & 0.025 & 0.001\end{pmatrix}^T, \tag{6.23a}$$

$$\vec{\beta}_{lower,0} = -\begin{pmatrix}9.000 & 9.000 & 9.000 & 0.500 & 0.500 & 0.010 & 0.025 & 0.025 & 0.001\end{pmatrix}^T. \tag{6.23b}$$

---

[4] s. Byrd *et al.* (1995) for details

This is for the normal distributed components of ${}^{n}\vec{u}$ (T_{DJF} and T_{JJA}) a range of approximately nine sigma. The transformation of the third component of ${}^{n}\vec{u}$ is not symmetric and therefore the upper range has to be adapted. Figure E.1 in appendix E.1 presents the resulting possible ranges for the profile transferred to the climate phase space.

6.3. CRU TS 3.1-E-OBS Mixture- Palaeoclimate Reconstruction Results

The preceding pages introduce a CFR for the Jordan Valley based on a mixture of the CRU TS 3.1 and E-OBS dataset. The local palaeoclimate reconstructions ${}^{s}f_{\vec{C}|\vec{P}}(\ldots)$ presented in chapter 4 and 5 have to reflect that. This is done by a modification of the input dataset for the estimation of the transfer functions. Both, GLM and QDA transfer functions are estimated based on a dataset which is a combination of CRU TS 3.1 and E-OBS.

All these palaeoclimate reconstructions show structures indicated in the marginal distributions which are very similar to the reconstructions based on the CRU TS 3.1 dataset alone. This is Best visible in the characteristics of the quantile curves and the distances between them. Therefore only the marginal distributions which show important differences are discussed here. The remaining figures for marginal distributions can be found in the appendix at which for

Ein Gedi Figure E.2 is the counterpart for 4.16a, 4.16b and 4.16c,

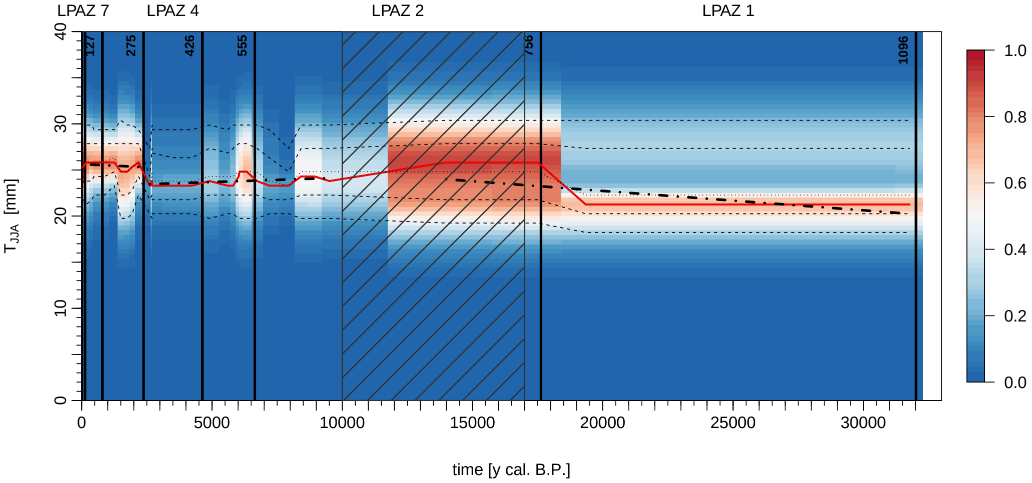
Lake Kinneret Figure E.3 is the counterpartfor 4.4a, 4.4band 4.4c,

Birkat Ram Figure E.4 is the counterpart for 5.3a, 5.3b and 5.3c.

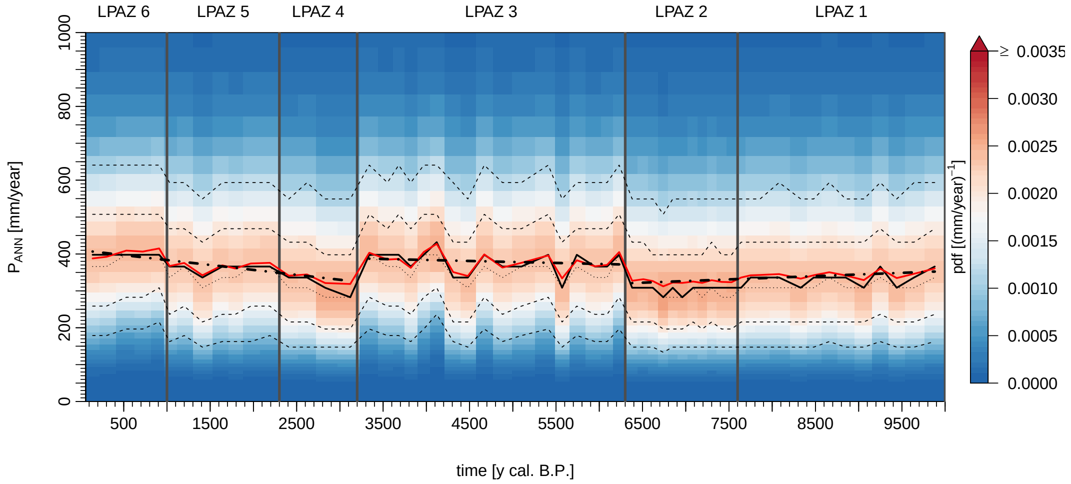
The most important characteristic is that the T_{JJA} palaeoclimate reconstruction for Birkat Ram shows no warming with the onset of the Younger Dryas (YD) (Fig. 6.3a). It is identifiable that the 10 % resp. 90 % quantile curves show no variation if the jump positions discussed in chapter 5.3 are excluded. Only the expectation value curve (red line) indicates a large jump with the onset of the YD (at about 17 000 y. cal BP). The result for the marginal distribution for P_{ANN} (shown in the appendix) is similar to the differences between the CRU TS 3.1 and ERA-Interim based palaeoclimate reconstruction: The climate change is visible but much less pronounced as in the palaeoclimate reconstruction estimated with the transfer function based on the CRU TS 3.1 dataset alone.

For Ein Gedi there is a change detectable in the marginal distribution for P_{ANN} (Fig 6.3b) but the time frame (LPAZ 6) with larger values is still identifiable. The difference is that the jumps at the borders of this zone are less distinct.

One possible explanation for the changes in the palaeoclimate reconstructions is that there are differences in the climate dataset due to the fact that the investigation area is situated at the



(a) Marginal posterior probability ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ for Birkat Ram



(b) Marginal posterior pdf ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ for Ein Gedi

Figure 6.3.: Marginal posterior distributions ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\overrightarrow{p}_0\right)$ for Birkat Ram (a) with $c_0 = T_{JJA}$ and Ein Gedi (b) with $c_0 = P_{ANN}$. In detail all base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$ and s.no. 84 (biome setup 3, $\vec{C} = \left(T_{DJF}, T_{JJA}, P_{ANN}\right)^T$ and a combination of CRU TS 3.1 and E-OBS as estimation dataset). The mapping of the figure is identical to Figure 4.4.

borders of E-OBS-dataset region. Especially the problems with T_{JJA} of Birkat Ram ($\equiv {}^{BR}T_{JJA}$) in the discussion/conclusions of the CFR later on have to be regarded since ${}^{BR}T_{JJA}$, ${}^{BR}P_{ANN}$ and ${}^{EG}P_{ANN}$ are the only marginal distributions which indicate climate changes. In the following the index ${}^{BR}\ldots$ is a shortcut for Birkat Ram, ${}^{EG}\ldots$ for Ein Gedi and ${}^{LK}\ldots$ for Lake Kinneret.

6.4. Time slices Selection

For an application of the CFR introduced in chapter 6.1 and 6.2 it is necessary to select contemplable time slices. For this aim it is required to choose the fossil site with the most reliable age-to-depth model. In this case this is Ein Gedi due to the fact that it has a laminated sediment core structure. Schiebel (2013, chap. 6, Fig. 6.2) also selected Ein Gedi as the "reference" core in the detailed pollen analysis of the fossil sites located in the Jordan Valley. The term "dated pollen samples" in the following paragraphs reflects the fact that for the regarded pollen sample layer an age-to-depth model is available.

Therefore it is tested if there are dated pollen samples in the Lake Kinneret and the Birkat Ram sediment core given an age of an Ein Gedi dated pollen sample. In the process it is assumed that the error of the Ein Gedi age-to-depth model is zero and those of the two others include errors. The second step is to filter this result. Figure 6.4 shows the result of that process.

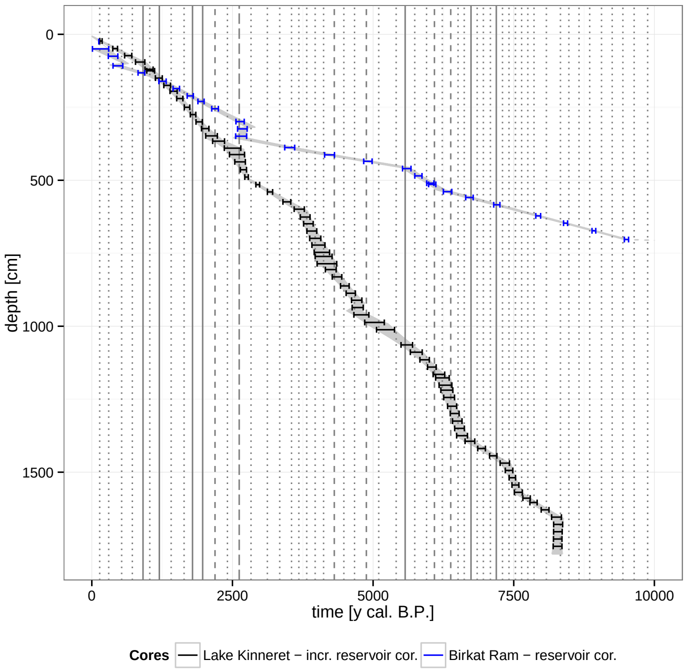


Figure 6.4.: Age-to-depth models for the time slices selection for the Jordan Valley palaeoclimate reconstruction:
Shaded grey areas: age-to-depth models established by Schiebel (2013).
Blue error bars: age of pollen sample layers of Birkat Ram. Black error bars: age of pollen sample layers of Lake Kinneret.
Vertical lines: pollen sample layers of the Ein Gedi. The line type marks the assignment of the dated pollen sample layers: solid: one-to-one and onto (s. Tab. 6.1); long dashed: more than one contemplable; short dashed: Birkat Ram has one but Lake Kinneret has more than one contemplable; dotted: remaining dated pollen samples of Ein Gedi.

The shaded grey areas mark the complete available age-to-depth models of both fossil sites

Table 6.1.: Selected time slices for the Jordan Valley palaeoclimate reconstruction. All Pollen sample layers with a one-to-one and onto assigned dated pollen sample layer for the selected time slice of the Ein Gedi pollen sample.

No.	Ein Gedi			Lake Kinneret			Birkat Ram		
1	5	≡	910 y. cal BP	5	≡	95 cm	6	≡	132 cm
2	7	≡	1200 y. cal BP	8	≡	150 cm	7	≡	161 cm
3	10	≡	1790 y. cal BP	13	≡	275 cm	9	≡	211 cm
4	11	≡	1970 y. cal BP	15	≡	323 cm	10	≡	230 cm
5	29	≡	5570 y. cal BP	44	≡	1064 cm	18	≡	460 cm
6	37	≡	6740 y. cal BP	58	≡	1394 cm	23	≡	559 cm
7	41	≡	7190 y. cal BP	60	≡	1444 cm	24	≡	584 cm

which are established by Schiebel (2013). The blue error bars mark the age of all available pollen sample layers of Birkat Ram and the black ones those of Lake Kinneret. All pollen sample layers of the Ein Gedi sediment core are marked with vertical lines. The solid lines mark those where the just described process results in one Ein Gedi, one Lake Kinneret and one Birkat Ram dated pollen sample or in other words with a one-to-one and onto assignment of the considered dated pollen sample. These pollen sample layers are listed in Table 6.1. For the age marked with a long dashed line (2620 y. cal BP) Birkat Ram and Lake Kinneret have more than one contemplable dated pollen sample. For the ages marked with short dashed lines Birkat Ram has one but Lake Kinneret has more than one contemplable dated pollen sample (2 190 y. cal BP, 4 310 y. cal BP, 4 880 y. cal BP, 6 090 y. cal BP and 6 380 y. cal BP). The remaining dated pollen samples of Ein Gedi are marked with dotted lines. For the seven time slices presented in Table 6.1 the Jordan Valley CFR method as introduced in chapter 6.1 and 6.2 is applied. With some considerations about the mathematical implementation it could be possible to reconstruct more than these time slices, but first it has to be analysed if the introduced approach works as the following chapter 6.5 tries to answer. After that it has to be evaluated if it is worth to apply this method to other time slices.

6.5. Jordan Valley - Palaeoclimate Reconstruction Result

This chapter presents the result of the minimization of $\mathcal{J}$. Since this palaeoclimate reconstruction has to be analysed for the whole Jordan Valley the past climate is determined by

$$\vec{c}_{R,0,h_{mean}} = \vec{\varphi}^{-1}_{prior}\left(\boldsymbol{R}\vec{\beta}_{R,0}\right), \tag{6.24a}$$

$$^{s}\vec{c}_{R,0,h_{mean}} = \vec{\varphi}^{-1}_{prior}\left({}^{s}\boldsymbol{P}\left(\boldsymbol{R}\vec{\beta}_{R,0}\right)\right) \tag{6.24b}$$

and not with the height corrected version as defined by equation 6.18. The difference between equation 6.24a and 6.24b is the matrix multiplication with ${}^{s}\boldsymbol{P}$ to obtain the palaeoclimate for one fossil site.

A detailed analysis of the Jordan Valley palaeoclimate reconstruction also implicates errors. This is done in the same way as in the previous workby utilizing the connection between the Hessian matrix at the minimum $\vec{\beta}_{R,0}$ and the covariance matrix $\boldsymbol{\Sigma}_{\vec{\beta}_{R,0}}$.

$$\frac{1}{2}\boldsymbol{\mathcal{H}}_{\vec{\beta}_{R,0}} = \boldsymbol{\Sigma}^{-1}_{\vec{\beta}_{R,0}}. \tag{6.25}$$

This approach originates from Gauthier (1992) and Rabier, Courtier (1992). The square roots of diagonal elements of $\boldsymbol{\Sigma}_{\vec{\beta}_{R,0}}$ yield than to the error bars of the palaeoclimate reconstruction as e. g. presented in Figure 6.5, 6.6a or 6.6b for one time slice. This is the CFR for the time slice 5570 y. cal BP.

6.5.1. Jordan Valley Palaeoclimate Reconstruction for 5570 y. cal BP

This time slice is chosen because the Ein Gedi reconstruction shows an increase in the annual precipitation amount in the local pollen assemblage zones (LPAZ) 3. This is discussed in chapter 4.5 where it is also suggested as the variable to evaluate in this application. The mapping of Figure 6.5, 6.6a and 6.6b is identical to 6.2. In all three figures the palaeoclimate reconstruction values of the profile of the past time slice (solid line) and the confidence band (dashed lines) based on equation 6.25 are added in orange. Thereby these confidence band represents the $68.3\% = \pm\sigma_{u_3} = \pm\sigma_{u_{P_{ANN}}}$ interval.

The results for all fossil sites and climate variables (${}^{s}\vec{c}_{R,0,h_{mean}}$) are summarized in Table E.2 (apx. E.3.2). One result of all CFR is, as shown exemplarily in Figure 6.5 for $c_{i=3,R,0,h_{mean}} := P_{ANN,R,0,h_{mean}}$, that the uncertainty of the palaeoclimate profile grows with increasing distance to grid point no. 1 or in other words the intercept. This result is not surprising and is in accordance with a general characteristic of regression problems: The error growth with growing distance to the intercept which is here grid point no. 1.

Another result for $P_{ANN,R,0,h_{mean}}$ in Figure 6.5 is that the estimated palaeoclimate profile for $P_{ANN,R,0,h_{mean}}$ is almost constant since the value for $\beta_{c3} \approx -0.0003$ (s. eq. 6.4 for the definition of $\beta_{c\ldots}$) is very small. The profile for the modern time slice shows quadratic behaviour

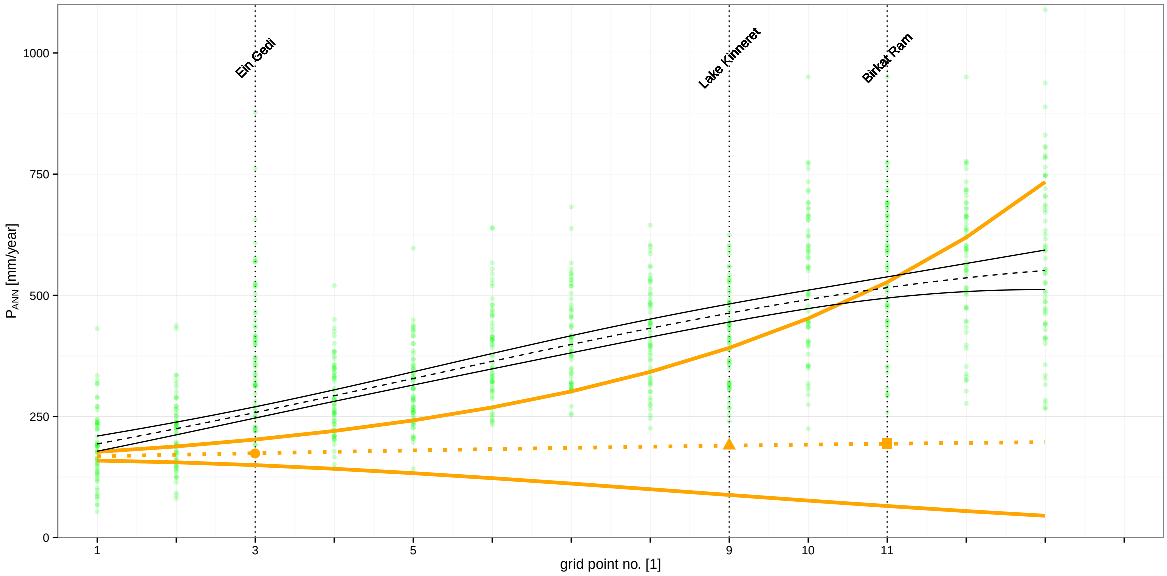


Figure 6.5.: The profile reconstruction in the Jordan Valley for $P_{ANN,R,0,h_{mean}}$ for the time slice 5570 y. cal BP.
The mapping of the figure is defined analogously to Fig. 6.2: The abscissa represents the Jordan Valley grid boxes. The green points represent the available data at each grid point (E-OBS and CRU TS 3.1). The vertical black lines mark the locations of the sediment cores Ein Gedi, Lake Kinneret and Birkat Ram. The dashed black line indicates the result of a GLM with quadratic term and the solid black lines are the corresponding 99.5 % confidence bands. In orange the palaeoclimate profile (minimum of $\mathcal{J}$) with the same mapping of the line types as the modern time slice GLM except that the confidence bands represent the 68.3% interval based on $\pm\sigma_{u_3} = \pm\sigma_{u_{P_{ANN}}}$. The values for the fossil sites are marked with an orange dot (EG = Ein Gedi), a triangle (LK = Lake Kinneret) and a square (BR = Birkat Ram).

($\beta_{c3} = -0.0027$). One consequence is that the climate anomaly defined by

$$^{s}c_i' := {}^{s}c_{i,R,0,h_{mean}} - {}^{s}c_{i,h_{mean}} \tag{6.26}$$

is larger for larger grid point numbers which is mathematically precise $\left|{}^{BR}P_{ANN}'\right| > \left|{}^{LK}P_{ANN}'\right| > \left|{}^{EG}P_{ANN}'\right|$. This characteristic is nearly valid for all time slices and variables as presented in Table 6.2. In a nutshell one result for all time slices is, despite the just mentioned explanation for ${}^{s}P_{ANN}'$ in Figure 6.5:

$$\left|{}^{BR}c_i'\right| > \left|{}^{LK}c_i'\right| > \left|{}^{EG}c_i'\right|. \tag{6.27}$$

The only exception is ${}^{s-5570}T_{DJF}'$ with a value of zero as indicated in Figure 6.6a which is discussed in detail in the following. For ${}^{s}P_{ANN}'$ equation 6.27 is even valid without the absolute value and therefore the result for the reconstructed palaeoclimate based on the minimum of

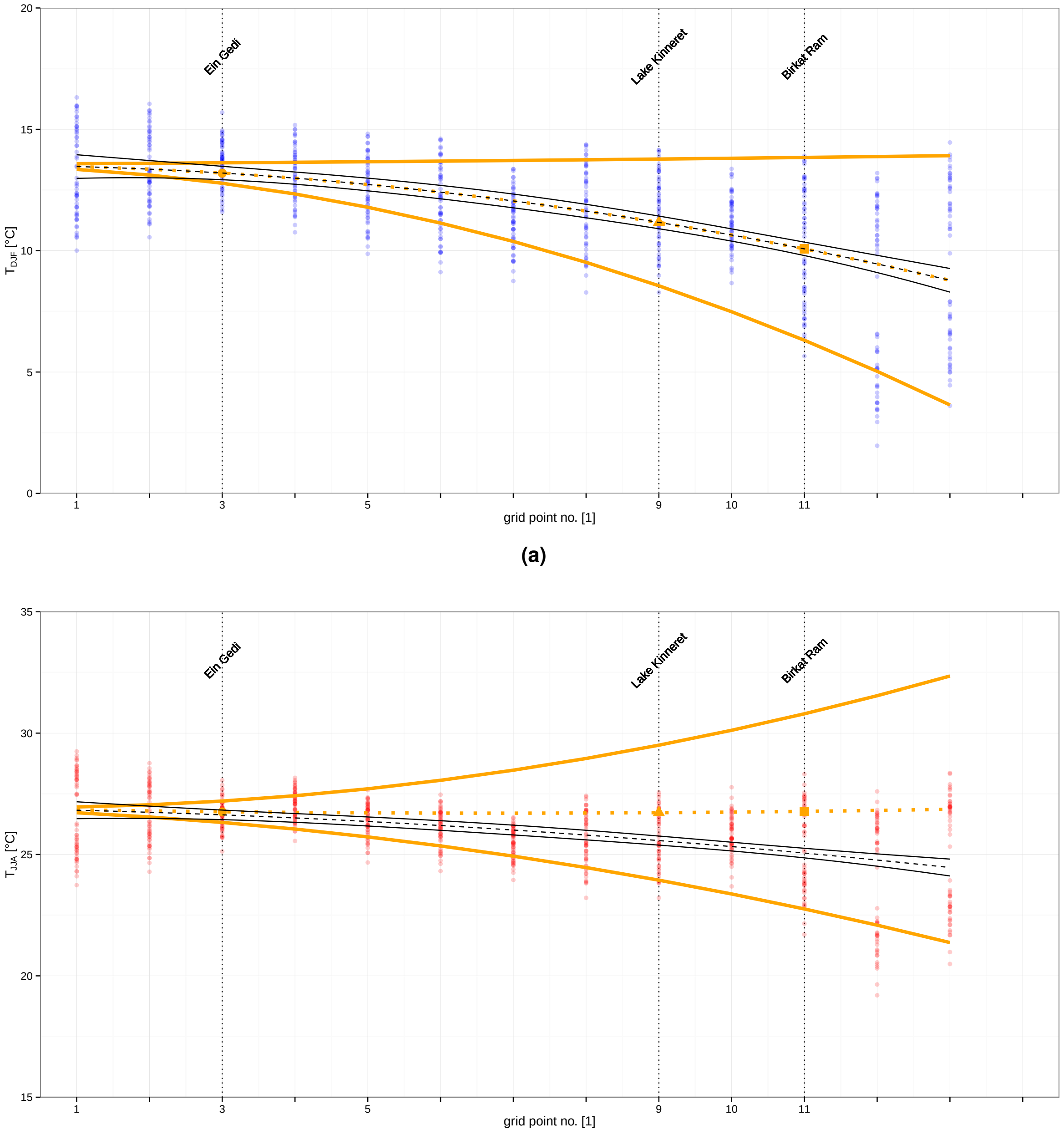
Table 6.2.: The Jordan Valley CFR result for the climate anomaly $^{s}c_i'$ for all considered time slices and fossil sites s. The first row in this table contains $^{s}c_{i,h_{mean}}$, the values for the modern time slice and the other rows the values for the climate anomaly $^{s}c_i'$ as defined by equation 6.26.

No.	time slice [y. cal BP]	Ein Gedi			Lake Kinneret			Birkat Ram		
		$^{EG}T_{DJF}{}'$ [°C]	$^{EG}T_{JJA}{}'$ [°C]	$^{EG}P_{ANN}{}'$ $[\frac{mm}{year}]$	$^{LK}T_{DJF}{}'$ [°C]	$^{LK}T_{JJA}{}'$ [°C]	$^{LK}P_{ANN}{}'$ $[\frac{mm}{year}]$	$^{BR}T_{DJF}{}'$ [°C]	$^{BR}T_{JJA}{}'$ [°C]	$^{BR}P_{ANN}{}'$ $[\frac{mm}{year}]$
0	0	13.2	26.6	259.5	11.1	25.6	464.3	10.1	25.1	514.9
1	910	0.0	0.0	-81.0	0.3	0.0	-230.5	0.4	0.0	-251.1
2	1200	0.2	0.1	-81.2	2.1	0.9	-230.7	3.1	1.3	-251.3
3	1790	-0.1	0.0	-81.4	-0.7	0.0	-230.9	-1.0	0.0	-251.6
4	1970	0.1	0.0	-81.3	1.3	0.0	-230.8	1.8	-0.0	-251.5
5	5570	0.0	0.1	-85.4	0.0	1.2	-274.4	-0.0	1.7	-321.1
6	6740	-0.6	-0.3	-81.4	-5.1	-3.1	-231.0	-7.6	-4.6	-251.7
7	7190	0.0	-0.0	-81.4	0.1	-0.1	-231.0	0.1	-0.2	-251.7

the cost function in a first assessment is that the past climate is dryer as the modern time slice. Keep in mind that the presented uncertainty range of the palaeoclimate profile in Figure 6.5 is the confidence interval of 68.3 % and not the 99.5 % (modern time slice) it leads to the conclusion that this just mentioned climatic change for $P_{ANN,R,0,h_{mean}}$ for time slice 5570 y. cal BP varies in certainty: In addition to the aforementioned growing uncertainty ranges the palaeoclimate and the modern confidence bands only show an overlap for Birkat Ram. This is not the case for Lake Kinneret and Ein Gedi so that the certainty of a climate change based on the minimum of the cost function is higher for Lake Kinneret and Ein Gedi than for Birkat Ram. A detailed analysis follows in the discussion of the time series in chapter 6.5.2.

Figure 6.6a presents the CFR for $T_{DJF,R,0,h_{mean}}$. It indicates that there is no climate change identifiable for all $T_{DJF,R,0,h_{mean}}$ as shown in Table 6.2 even if the uncertainty range $\sigma_{u_{T_{DJF}}}$ is incorporated in this consideration. The remaining palaeoclimate reconstruction profile $T_{JJA,R,0,h_{mean}}$ shows that there is a slight change to higher summer temperatures with growing grid point number (Fig. 6.6b). Due to the fact that this derivation is small and within the 68.3 % confidence interval a climate change is relatively uncertain.

Neither the winter nor the summer temperatures indicate a palaeoclimate change for this period a fact which is supported by the local palaeoclimate reconstructions discussed in the chapter 4.3, 4.5, 5.3 and 6.3 and in the following chapter 6.5.2.



(a)

(b)

Figure 6.6.: The profile reconstruction in the Jordan Valley for $T_{DJF,R,0,h_{mean}}$ and $T_{JJA,R,0,h_{mean}}$ for the time slice 5570 y. cal BP.
The mapping of the figures is defined analogously to Fig. 6.5 the only difference is the colour of the available data points: for T_{DJF} it is blue and for T_{JJA} red.

6.5.2. Jordan Valley Palaeoclimate Reconstruction - Time Series

The most important aim of the Jordan Valley CFR is to evaluate the local palaeoclimate reconstructions and the possible indicated climate changes there. Therefore the profiles for the remaining time slices can be found in appendix E.3.1. The analysis of the CFR is now continued by time series plots which are comparable to the marginal posterior distributions ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ of the local palaeoclimate reconstructions as for example presented in Figure 6.3. These time series plots are the visualisation of Table 6.2 or resp. E.2 .

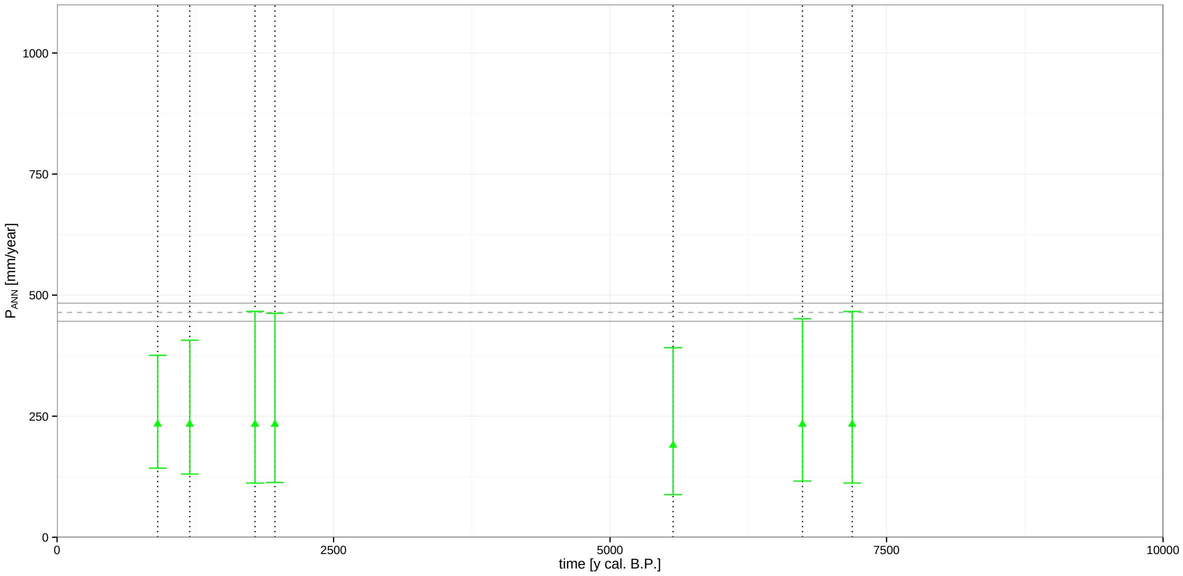


Figure 6.7.: The palaeoclimate CFR result ${}^{LK}P_{ANN\,R,0,h_{mean}}$ for Lake Kinneret summarized as a time series. The abscissa represents the age in y. cal BP. In green the palaeoclimate CFR result based on an application of equation 6.24b. The values ${}^{LK}P_{ANN\,R,0,h_{mean}}$ (minimum of $\mathcal{J}$) are marked with a triangle and the error bars represent the 68.3% interval based on $\pm\sigma_{u_3} = \pm\sigma_{u_{P_{ANN}}}$ (s. eq.6.25). The mapping for the modern time slice GLM value for Lake Kinneret is the same as in Figure 6.2 except that here it only represents the value at grid point no. 9.

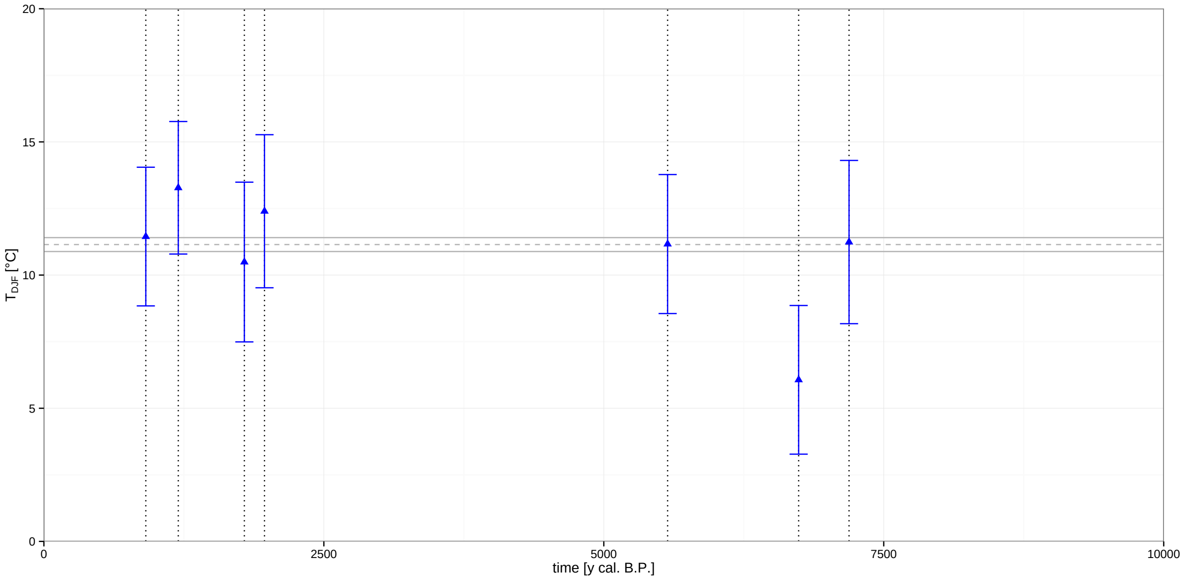
The aforementioned result of a palaeoclimate for annual precipitation amount dryer than the modern climate persists at first glance for all fossil sites and analysed time slices as Figure 6.7, E.11 and E.12 or the corresponding data in Table 6.2/E.2 indicate: For Ein Gedi these are values between 174 and 179 $\frac{\text{mm}}{\text{year}}$ with small errors between $+\sigma_{u_{P_{ANN}}} = 30$ and $-\sigma_{u_{P_{ANN}}} = 25\,\frac{\text{mm}}{\text{year}}$ compared to them of the two other fossil sites. The result for Lake Kinneret lies between 189 and 234 $\frac{\text{mm}}{\text{year}}$ (errors between $+\sigma_{u_{P_{ANN}}} = 233$ and $-\sigma_{u_{P_{ANN}}} = 122\,\frac{\text{mm}}{\text{year}}$) and for Birkat Ram between 193 and 264 $\frac{\text{mm}}{\text{year}}$ (errors between $+\sigma_{u_{P_{ANN}}} = 417$ and $-\sigma_{u_{P_{ANN}}} = 129\,\frac{\text{mm}}{\text{year}}$). The overlap of the error bars with the modern time slice varies beginning with no overlap for Ein Gedi (Fig. E.11), half of them overlap for Lake Kinneret (Fig. 6.7) and all overlap for Birkat Ram (Fig. E.12). In summary all results of this Jordan Valley CFR yield to values including the error bars of around 200 $\frac{\text{mm}}{\text{year}}$ for all fossil sites and time scales. If the overlap as well as the range of the error bars is taken as a measure of uncertainty of the palaeoclimate reconstruction Ein Gedi is the one with the most reliable result followed by Lake Kinneret and than by Birkat Ram.

If the result is regarded in detail this has to be treated with caution due to the following facts: The minimization of $\mathcal{J}$ yields to the Jordan Valley profile which fits best to the modern profile and the marginal posterior distributions ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ for all components of the climate state vector. If the values for ${}^{s}P_{ANN,h_{site}} \equiv {}^{s}P_{ANN,h_{mean}}$ are analysed together with the marginal distribution figures for ${}^{s}P_{ANN,0,h_{site}}$ (s. e. g. Fig. 6.3 or apx. E.2) it becomes obvious that the modern profile for ${}^{s}P_{ANN,h_{site}}$ lies almost in the 50% interquantile range of the marginal distributions for all three Figures but for ${}^{EG}P_{ANN,h_{site}}$ it is too dry and for ${}^{LK}P_{ANN,h_{site}}$ and ${}^{BR}P_{ANN,h_{site}}$ too wet compared to the expectation value of the complete profile ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ (red lines in the considered figures in apx. E.2). This includes the modern time slice marginal distribution (assigned to the first pollen sample layer) which can be regarded as poor representation of the modern time slice since the dating is not precise as Figure 6.4 indicates.

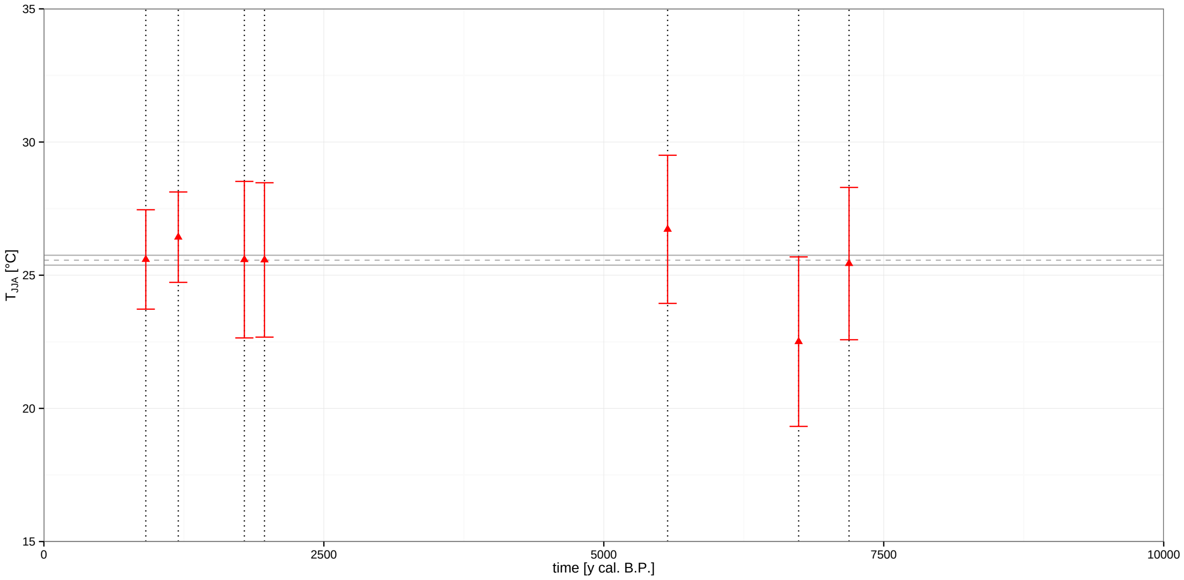
Since the minimization of $\mathcal{J}$ results in ${}^{s}P_{ANN,R,0,h_{mean}}$ values, as just mentioned, which fit best to both parts $\mathcal{J}_{veg}$ and $\mathcal{J}_{Mod}$ of the cost function (s. eq. 6.9), the result fits to this approach. But why is there a difference between the modern time slice value of the GLM fit and the expectation value of the local reconstructions for the first pollen sample layer? On the one hand it is possible that the modern Jordan Valley profile described by a GLM with quadratic terms is mathematical not the best representation for the modern Jordan Valley climate for ${}^{s}P_{ANN,h_{site}}$. One conceivable solution for palaeoclimate reconstructions is an application of the method developed by Simonis (2009). It is based on a simple model for the annual precipitation amount which is an application of the omega equation. On the other hand the dataset for the estimation of the modern Jordan Valley profile should be assessed critically. The approach presented here uses a mixture of two datasets to enlarge the geographical resolution of the Jordan Valley (chap. 6.1). The result for the local palaeoclimate reconstructions are discussed in chapter 6.3 in which one problem is already addressed: the investigation area is situated at the borders of E-OBS-dataset region. But it is not possible to exclude the effect if the borders based on the available data and analyses presented in this work for the Jordan Valley. Later on, after the discussion of the temperature reconstructions, the mixing of the datasets is regarded in more detail.

As mentioned in chapter 6.3, ${}^{BR}T_{JJA,0,h_{site}}$, ${}^{BR}P_{ANN,0,h_{site}}$ and ${}^{EG}P_{ANN,0,h_{site}}$ are the only marginal posterior distributions which indicate climate changes in the past as discussed in the previous chapters (4.3, 4.5 and 5.3). But the palaeoclimate change for ${}^{BR}T_{JJA}$ appear with the onset of the YD at about 17000 y. cal BP which is outside of the time range allowing a CFR for the Jordan Valley with the three considered fossil site data in this work here as Figure 6.4 indicates. Hence all Jordan Valley CFR for the temperatures should indicate no palaeoclimate change.

This is the case as the time series plots (Fig. 6.8 and E.11 and E.12) show at a first glance for all considered time slices except 6740 y. cal BP. Thereby the result for the 6740 y. cal BP time slice for Ein Gedi fits since both palaeoclimate reconstructions (${}^{EG}T_{DJF,0,h_{mean}}$ and ${}^{EG}T_{JJA,0,h_{mean}}$) overlap with the modern time slice value and the difference between the minimization results is relatively small (s. Tab. E.2 for the absolute values). The difference between them (${}^{EG}T_{DJF}{}' = -0.6\,°\mathrm{C}$ and ${}^{EG}T_{JJA}{}' = -0.3\,°\mathrm{C}$) indicate a small cooling for both temperature re-



(a) $T_{DJF,R,0,h_{mean}}$ in blue



(b) $T_{JJA,R,0,h_{mean}}$ in red

Figure 6.8.: The palaeoclimate CFR results ${}^{LK}T_{DJF,R,0,h_{mean}}$ and ${}^{LK}T_{JJA,R,0,h_{mean}}$ for Lake Kinneret summarized as a time series.
The mapping of the figures is defined analogously to Fig. 6.7 the only difference is the colour: for T_{DJF} it is blue and for T_{JJA} red.

constructions but within the confidence range. For Lake Kinneret and Birkat Ram the Jordan Valley CFR result for 6740 y. cal BP indicates cooling with no overlap of the error bars for ${}^{LK}T_{DJF}{}' = -5.1\,°\mathrm{C}$ and ${}^{BR}T_{DJF}{}' = -7.6\,°\mathrm{C}$ and only a small for for ${}^{LK}T_{JJA}{}' = -3.1\,°\mathrm{C}$ and ${}^{BR}T_{JJA}{}' = -4.6\,°\mathrm{C}$.

To get an explanation for the difficulties with time slice 6740 y. cal BP it is necessary, as done for ${}^{s}P_{ANN,0,h_{mean}}$, to consider the marginal posterior distributions ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ together with the modern profile values for the temperatures. For this, the temperature height correction has to be applied as mentioned e. g. on page 100. Therefore not ${}^{s}T_{\dots,h_{site}} \not\equiv {}^{s}T_{\dots,h_{mean}}$ is valid but ${}^{s}T_{\dots,h_{site}} \equiv {}^{s}T_{\dots,h_{mean}} + {}^{s}\Delta T$ (based on eq. 6.18) with the values ${}^{EG}\Delta T = 3.3\,°\mathrm{C}$, ${}^{LK}\Delta T = 3.5\,°\mathrm{C}$, and ${}^{BR}\Delta T = -4.1\,°\mathrm{C}$. If this is done, similar results to those of the annual precipitation amount become obvious: The modern profile lies in all considered figures in the 80 % interquantile range of the marginal distributions (s. apx. E.2) and for Lake Kinneret actually in the 50 % interquantile range. But for Ein Gedi and Lake Kinneret the modern profiles indicate too high temperatures, compared to the expectation value of the complete profile and for Birkat Ram too cold temperatures. The discussion of this bad representation of the modern time slice marginal distributions in connection with the annual precipitation amount (GLM and input dataset) can be transfered to the result for the temperatures. For the temperatures there is an additional possible explanation: Is the temperature correction really necessary for this simple kind of CFR? If it is leaved out in this comparison with the marginal distributions, than the modern profile value is not that bad representation of the modern time slice for Ein Gedi and Lake Kinneret. Due to the fact that this result is not consistent for all fossil sites it is not followed up by a minimization of the cost function without the height correction.

For a complete explanation of the result of time slice 6740 y. cal BP it has again to be kept in mind that the cost function summarized contributions from $\mathcal{J}_{veg}$ and $\mathcal{J}_{Mod}$. Here it is necessary to consider the differences to the other considered time slices: First for Ein Gedi there is a very small distance between the modern time slice value and the maximum of the probability density function (pdf) of ${}^{EG}T_{DJF,0,h_{site}}$. Second ${}^{EG}T_{JJA,0,h_{site}}$ has the highest pdf value in the marginal profile resulting in a small interquantile range. And third for Birkat Ram the distance between the modern time slice value and the pdf-maximum of ${}^{BR}T_{DJF,0,h_{site}}$ is also smaller than for the other time slices. Since the grid point of Ein Gedi is situated near the grid point which represents the intercept differences, they have a large impact on the whole Jordan Valley CFR profile. In other words the pdf values of the fossil site Ein Gedi move the whole profile to lower temperatures.

In the discussion of Figure 6.7 the difficulties of mixing two different datasets was shortly addressed. For a complete analysis of this the Jordan Valley profiles for T_{DJF} und T_{JJA} (Fig. 6.2 resp. 6.6a and 6.6b) have to be included in the considerations. Some important facts have to be mentioned before: First both datasets use different orography (s. e. g. Fig. 6.1). The CRU TS 3.1 temperature data is interpolated on the height of the E-OBS orography with the already introduced height correction based on the temperature gradient of the standard atmosphere γ_d (s. e. g. chap. 6.2.2). Furthermore the CRU TS 3.1 dataset has no border since it is available for the complete earth so that effects due to borders can be excluded for this dataset. Also the geographical extension of the E-OBS dataset is larger for the temperature than for the precipitation amount dataset (s. Webpage). Second the interpolation is not available for the annual precipitation amount.

Analysing Figure 6.5 it is recognizable that the distribution of the input data covers a large interval for grid point no. 3 (EG) and from 10 to 13 of $\approx 700\,\frac{\text{mm}}{\text{year}}$ compared to the rest ($\approx 300\,\frac{\text{mm}}{\text{year}}$). In Figure 6.6 this is only identifiable for grid point no. 12 and 13 ($\approx 11\,^\circ$C) compared to $\approx 5\,^\circ$C for the other ones. Additionally there is a gap in the distribution of the data points for these two grid points which is the main reason for the large spread. This visible grouping of the data points can be assigned to the two datasets used (not visualized in the figures). If a large spread is linked with difficulties in the mixing of the datasets the following can be concluded: The height correction based on the incorporated orography is not unproblematic since the problems appear not at the same grid points where the precipitation amount profiles indicate a large spread. Additionally the boundary of the temperature dataset, as mentioned in the last paragraph, is not that near as for the precipitation dataset, so therefor the mixing of the dataset and also the CFR has to be regarded with care.

6.5.3. Jordan Valley palaeoclimate reconstruction - Summary and Conclusion

In summary the Jordan Valley CFR results in a dryer palaeoclimate than the modern climate with P_{ANN} values around $200\,\frac{\text{mm}}{\text{year}}$ for all fossil sites and considered time slices. If the overlap as well as the range of the error bars is taken as a measure of uncertainty of the palaeoclimate reconstruction Ein Gedi is the one with the most reliable result followed by Lake Kinneret and than by Birkat Ram. Also the result of the local reconstructions in the previous chapters of no palaeoclimate change in the summer and winter temperatures is confirmed. For each of the reconstructed Jordan Valley profiles (s. apx. E.3.1 or Fig. 6.5 and 6.6) the uncertainty of the palaeoclimate profile grows with increasing distance to grid point no. 1 or in other words the intercept. This is in accordance with a general characteristic of regression problems: The error growth with increasing distance to the intercept which is here grid point no. 1. Also valid is that the climate anomaly defined by equation 6.26 fulfils

$$\left| {}^{BR}c_i' \right| \geq \left| {}^{LK}c_i' \right| \geq \left| {}^{EG}c_i' \right| \tag{6.28}$$

for all time slices.

But all quantitative results for the temperatures and the annual precipitation amount have to be handled with care since the Jordan Valley CFR approach is problematic. Mathematically and numerically the method works but there are some problems with the input data additionally to problems of the local reconstructions as discussed in the analyses of figures in the last section. If this discussion is extended one question respective the input datasets arises: Why isn't E-OBS used allone as input climate dataset for the palaeoclimate reconstructions? One answer is that the covered area isn't large enough to estimate all required transfer functions alone, neither the biome nor the taxon specific transfer functions. For the biome this is especially the extended distribution area of the Irano-Turanian which lies partly outside of the E-OBS dataset area. Another one is that a re-digitalization of the distribution maps has to be done in which the question of representation arises: Is the distribution represented by the maps really sensitive enough for the high geographical resolution of the E-OBS dataset? With the mixing of both dataset in the estimation of the transfer functions this problem is partly avoided since this is quasi equal to an averaging of both. The only effect of the averaging is a slightly reduced $BSS_{\dddot{B}_l}$. Additionally it is mentioned in chapter 4.2.3 that the main characteristics of the marginal distributions persist. Due to the other problems in all parts of the palaeoclimate reconstructions in the Jordan Valley this re-digitalization is not done.

Another question is: Why isn't another dataset used? First the CRU TS 3.1 data alone has a too coarse spatial resolution (s. chap. 6.1). Another gridded dataset based on observations with a high resolution for at least the complete Eurasian continent and the Mediterranean area, which is required, isn't available (as known by the author). Additionally all known available reanalyses with a geographical high resolution have the same problem that the required area is not available or is situated in the fringe.

If this availability of geographical high resolution climate data is solved there are also other difficulties: One main problem is that the local reconstructions, especially Birkat Ram and Lake Kinneret, have to be optimized (details in chap. 6.5.3). Switched back to the Jordan Valley CFR approach there are also methodical problems: The selection resp. definition of the intercept position has a large influence on the whole result. An equal weighting of all fossil sites is not possible, even if the factor $\Delta \sin \varphi_s$ (s. chap. 6.2.3) is incorporated in the minimization. Another approach, as for example the CFR introduced by Gebhardt (2003) could solve this problem. If this is not done the question of representation of modern Jordan Valley climate by the GLM profile as introduced in chapter 6.1 has then to be solved.

7

Summary, Conclusion and Outlook

The first part of this chapter gives an short overview over all results of this work and assesses them in connection with the ojectives mentioned in the introduction. In chapter 7.2 an outlook to further work is given.

7.1. Summary and Conclusion

General Summary

All local palaeoclimate reconstructions presented in this work are marginal distributions of a reconstructed multivariate (almost all with three dimensions) probability density function (pdf) (chap. 3 and 4) or probability (chap. 5). For all estimated marginal distributions the associated marginal quantiles are estimated as well as the weighted mean which is interpreted as a numerical expectation value. These quantities are determined to allow a better evaluation and interpretation of the corresponding marginal distributions. For Lake Prespa and Birkat Ram a linear trend is estimated for several time slices for a better evaluation of climate changes. This trend is estimated on the numerical expectation value of the marginal pdfs of each climatolgical variable c_i. For Lake Prespa there are four and for Birkat Ram there are three time ranges. The numbers are from past to present and time range X is shortened from now on with t. r. X. These t. r. are identifiable in all palaeoclimate reconstructions since the marginal distribution profiles before and after differ more or less. Some of these time ranges are compared by application of a Student's t-test with a significance level of 0.05 to evaluate the climate change.

For all considered fossil sites marginal distributions for the temperature of the coldest season/month (T_{Jan} for Lake Prespa resp. T_{DJF} for the other ones) and the warmest season/month (T_{Jul} resp. T_{JJA}) are available. Due to the fact, as mentioned in the introduction, that the Ein Gedi palaeoclimate reconstruction is only adapted to the Jordan Valley situation for this fossil site no middle troposphere (850 hPa level) temperature palaeoclimate reconstruction is estimated. For the other fossil sites 850 hPa temperature reconstructions of the coldest season/month (T_{Jan}^{850hpa} resp. T_{DJF}^{850hpa}) and the warmest season/month (T_{Jul}^{850hpa} resp. T_{JJA}^{850hpa}) are discussed. Surface temperatures based on the middle troposphere temperatures could be determined with a simple height correction (s. e. g. Stolzenberger 2011), based on the temperature gradient of the standard atmosphere ($6.5 \times 10^{-3} \frac{K}{m}$), the mean height of the 850 hPa level (1500 m) and the height above mean sea level (AMSL) of the drilled sediment core which are also referred as fossil sites. As also mentioned in chapter 2.1.1 it is an ongoing discussion which and which combination of environmental parameters is the best to be reconstructed.

Also available for all fossil sites are palaeoclimate reconstructions of climate variables which describe the water availability: The "classical" annual precipitation amount P_{ANN} and the climatic water deficit (CWD) accumulated for one year (CWD_{ANN}).

All presented palaeoclimate temperature reconstructions except Lake Kinneret share that the surface and the middle troposphere temperature reconstructions are in accordance. All of them validate that more (Lake Prespa and Birkat Ram) or less (Lake Kinneret) the temperature height correction mentioned on page 40 is correct. Also shown is that the CWD_{ANN} palaeoclimate reconstruction works and does not contradict P_{ANN}. The marginal distribution for CWD_{ANN} is for example for Lake Prespa a reconstructed palaeoclimate variable which allows more identifiable variation than in P_{ANN}. The results of each fossil site is summarized in more detail in the following.

Lake Prespa

For the sediment core of Lake Prespa the Bayesian Indicator Taxa Model (BITM) is selected as palaeoclimate reconstruction method. Therefore the pollen counts are transferred with a threshold definition to presence and absence information. This method allows to take expert knowledge into account and the result is the indicator taxa matrix. This indicator taxa matrix is then filtered with an algorithm which ensures that taxa with to similar transfer functions are excluded in the palaeoclimate reconstruction (s. chap. 3.2.1 for details). The transfer functions for the required taxa base on generalized linear model (GLM) estimated for a three dimensional climate state vector. The setup is in detail described in chapter 3.2.2. The required climate prior for the BITM is a subjective prior defined by a product of the marginal distributions of the multivariate prior distribution (s. chap. 3.2.3 and eq. 3.32 for details). The parameter definition bases on histograms of the modern climate at the fossil site. Both results are then applied to estimate four different palaeoclimate reconstructions of a three dimensional climate state vector $\vec{C} := (T_{Jan}, T_{Jul}, P_{ANN})^T$, $(T_{Jan}, T_{Jul}, CWD_{ANN})^T$, $\left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, P_{ANN}\right)^T$ or $\left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, CWD_{ANN}\right)^T$ (s. eq. 3.1 for details).

The estimated marginal distributions are presented in Figure 3.6 and 3.7 and discussed in chapter 3.3. Briefly summarized the results are: There are four discriminable time ranges identifiable in the surface temperature palaeoclimate reconstructions for T_{Jan} and T_{Jul} with greater climate changes for T_{Jan} than for T_{Jul}. For T_{Jan} the greatest change between two time ranges is identifiable between t. r. 2 and 3. Time range 3 differs also significantly from the modern climate. For T_{Jul} only one time range is identifiable with a significant difference to the modern T_{Jul} climate for a longer time scale: t. r. 2. The palaeoclimate reconstruction of the middle troposphere temperatures T_{Jan}^{850hpa} and T_{Jul} (Fig. 3.7a resp. 3.7b) supports this results: In general the differences between the time ranges are persistent respective to the significance and the direction of the change compared to the surface temperature results presented in Figure 3.6a resp. 3.6b. The most clear quantifiable change is that T_{Jan}^{850hpa} is on average in t. r. 2 more similar to t. r. 3 as it is the case for T_{Jan}. Also the marginal pdfs for T_{Jan}^{850hpa} and T_{Jul}^{850hpa} are slightly narrower than those of T_{Jan} and T_{Jul}. As a consequence the simple height correction mentioned in the beginning is also confirmed. The result for the palaeoclimate reconstructions of the water variables P_{ANN} and CWD_{ANN} is different from that: One important result is that different definitions of the climate state vector yield to palaeoclimate reconstruction results which are in general not inconsistent. This means that the sign of the differences between the group means of the time ranges is the same. For both P_{ANN} reconstruction results, Figure 3.6c and 3.7c, only the climatic oscillations in t. r. 1 are reliable. Compared to the P_{ANN} palaeoclimate reconstruction both CWD_{ANN} results (Figure 3.6d and 3.7d) indicate four clearly discriminable time ranges. These differ significantly among themselves with only one time range with a really visible and significantly different climate to the modern time slice: time range 2.

Lake Kinneret

For Lake Kinneret the Bayesian Biome Model (BBM) is the selected palaeoclimate reconstruction method since Lake Kinneret has a large catchment area (s. chap. 2.3.2.1 for details). For the BBM the taxa are assigned to the Mediterranean or Irano-Turanian biome. Then the pollen taxa counts of all taxa are summarized to the according biome probability $W_{B_{l_0}}$ (s. chap. 4.2.2 for details). The required thresholds ϑ_{k_0} for the so called noise reduction are determined with different methods: $\vartheta_{k_0} = 0$, $\vartheta_{k_0} = Q_{k_0}\left(\vec{\tilde{\omega}}_{k_0}\right) = 33\,\%$ the $33\,\%$ quantile of $\vec{\tilde{\omega}}_{k_0}$ and a taxon specific individual definition. The influence of the different thresholds is evaluated: Compared to $\vartheta_{k_0} = 0$ there is an increase of the probability for the Mediterranean biome in the upper part of the record and a decrease in lower part (s. Fig. 4.1). It is clearly deducible that the trend of the biome probability $W_{B_{l_0}}$ varies in the same way as the ratio trees & shrubs/grasses & herbs defined by Schiebel (2013). This fits very well with the fact that the Irano-Turanian biome is a steppe biome type dominated by grasses & herbs (Schiebel 2013). Also identifiable is that only in the lower part of the sediment core the Irano-Turanian biome is preferred, since the biome probability always varies around 0.5 in the rest of the profile. Later on also the influence on the palaeoclimate reconstruction is examined. The climate prior distribution is estimated analogously to the Lake Prespa climate prior with histograms of the modern climate of the Jordan Valley (s. chap. 4.2.1 for details). This is done due to the fact that the Lake Kinneret palaeoclimate reconstruction result is used later on in the Jordan Valley climate field reconstruction (CFR) (s. chap. 6) and the aim in chapter 6 is to learn about climate changes in the Jordan Valley from the data. The biome transfer functions or biome likelihoods $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}|\,\vec{c}_0\right)$ (s. chap. 4.2.3 for details) are estimated with various setups: different biome distributions, different climate datasets and definitions of the climate state vector. The result is verified with the Brier skill score BSS in order to discover the most promising combination of the parameters. The conclusion from the verification result allone is that a three dimensional climate state vector has to be applied. The selection of the biome setup depends on the aim of the palaeoclimate reconstruction: biome setup 3 is selected for surface temperature reconstructions and biome setup 4 for middle troposphere temperature reconstructions.

The palaeoclimate reconstruction result is that none of setups allows a reconstruction of a period with climate change which is identifiable in any of the presented marginal distributions. The only identifiable result is that there is some variation dependent on the biome probability $W_{B_{l_0}}$. The conclusion of the application of the BBM with two biomes is: The lack of difference of the two regarded biomes in climate phase space does not allow a palaeoclimate reconstruction with a climatic change based on the BBM, where especially T_{JJA} and CWD_{ANN} do not differ. Thereby the CWD_{ANN} and P_{ANN} palaeoclimate reconstruction results do not contradict themselves. The middle troposphere temperature reconstructions do not completely contradict the surface temperature reconstructions since the sign of the height correction is correct. But the structure (almost no variation) and the absolute value of the height correction does not fit.

The problems of the BBM with two biomes motivates to test the pollen-ratio model (PRM) as presented in chapter 4.4. In a nutshell the PRM introduced by Ohlwein, Wahl (2012) connects the composition of modern pollen spectra described by a ratio of two pollen taxa with the modern climate. The inversion of this connection allows a palaeoclimate reconstruction. The already mentioned E-OBS climate dataset and the pollen spectra presented in chapter 2.2.2 form the database for the estimation of the PRM-GLM. The result presented in Figure 4.9 shows that the distribution of the data points covers a too small interval in the climate phase space to allow a robust GLM estimation. Especially the absence of low precipitation values is a problem. Therefore a PRM-GLM based on this data is not inverted and applied in a palaeoclimate reconstruction.

Ein Gedi

Due to the fact that the Jordan Valley CFR (s. chap. 6) requires an Ein Gedi palaeoclimate reconstruction with other settings as in Litt *et al.* (2012) the BBM is applied with a three dimensional climate state vector, CRU Time Series (CRU TS) 3.1, biome setup 3 and the Jordan Valley prior distribution (s. chap. 4.5). The estimated result is compared to Litt *et al.* (2012): The change of the biome setup from 1 to 3 and the change of the climate prior has the most impacts on the palaeoclimate reconstruction since the period with a climate change in the summer temperature T_{DJF} is not identifiable. The climate change period for P_{ANN} is robust since it is identifiable in all presented marginal distributions but with different absolute values (s. chap. 4.5 for details). Also the palaeoclimate reconstruction result for CWD_{ANN} is consistent to P_{ANN}: A period with climate change to a more wet climate is identifiable.

The conclusions which analyse the circumstances when it is possible to find a time range with palaeoclimate change are repeated here (s. chap. 4.6):

- One biome in the depth profile of $W_{B_{l_0}}$ has to dominate for several pollen sample layers/time slices. Dominating here implies high $W_{B_{l_0}}$ values and low for the other ones.
- Additionally this dominating biome has to be clearly discriminable in the $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathcal{b}_{l_0},\vec{c}_0)$, the clearer the better. If more than two biomes are considered in the BBM then a location of this biome between the other one is critical.
- If the number of considered biomes is greater than two and two of the biomes have an identical mode for $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathcal{b}_{l_0},\vec{c}_0)$, then the biome which is clearly discriminable has to be the dominating one in $W_{B_{l_0}}$. This fact is one reason why it is possible to identify a period with climate change in the reconstruction for Ein Gedi and not in the one for Lake Kinneret.
- The probability $\mathbb{P}_{\mathcal{B}_l,\vec{C}}(\mathcal{b}_{l_0},\vec{c}_0)$ of one biome is always strongly influenced by the climate prior: In the climate phase space which is splitted into several areas by the quadratic discriminant analysis (QDA), this biome is the one which is situated at the edges. In the cases considered in this chapter this is the Mediterranean biome B_1
- The data basis for the QDA is thin/sparse which has the consequence that the QDA result is not robust especially for the Mediterranean biome.

Birkat Ram

For Birkat Ram a new method is developed which is situated between the BBM and BITM since the requirements for an application of these methods are not exactly given (s. chap. 5.1 for details). The Bayesian Indicator Taxa and Biome Model (BITBM) combines biome (BBM) and taxon (BITM) information by an application of the Bayes theorem on the joint probability $\mathbb{P}_{\vec{C},\vec{T},\mathcal{B}_l}\left(\vec{c},\vec{t},\mathcal{b}_l\right)$. Presented in equation 5.5 is the applied BITBM:

$$\mathbb{P}_{\vec{C}_0|\vec{T}_0}\left(\vec{c}_0|\vec{t}_0\right) = \pi_{\vec{C}}\left(\vec{c}_0\right)\cdot\sum_{l=1}^{N_l} W_{B_{l_0}} \frac{\left[\prod_{\substack{k\\ \forall t_{\iota\, k_0}=1}} \mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_{k_0}|\vec{c}_0,\mathcal{b}_{l_0}\right)\right]\cdot\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathcal{b}_{l_0}|\vec{c}_0\right)}{m_{T,\mathcal{B}_l}\left(t_0,\mathcal{b}_{l_0}\right)}.$$

The BITBM is applied on six taxa (T. Litt pers. comm.) of which four are assigned to the Mediterranean biome type and two are assigned to the Irano-Turanian biome type. The required biome probability $W_{B_{l_0}}$ is estimated in the same way as for Lake Kinneret. For consistency the required taxon specific threshold values ϑ_{k_0} are the same as later on applied for $\mathcal{I}_{\vartheta_0}^{N_\iota\times N_{k_0}}$. The resulting depth profile of the biome probability $W_{B_{l_0}}$ is divided in three different zones which are also identifiable in the palaeoclimate reconstruction. The first zone (0 to $\approx$ 10000 y. cal BP) shows a clear preference that Birkat Ram is located in the Mediterranean biome type. The second zone ($\approx$ 10000 to $\approx$17000 y. cal BP) is the zone where Schiebel (2013) assumed a desiccation of Birkat-Ram. The third zone ($\approx 17\,000$ to $\approx 30\,000$ y. cal BP) shows a clear preference for the Irano-Turanian biome . The biome probability $W_{B_{l_0}}$ result here is in a good agreement with the results of the pollen analysis of Schiebel (2013) which includes all taxa detected in the Birkat Ram sediment core. For the same reasons as for Lake Kinneret and Ein Gedi the Jordan Valley climate prior is applied. The part in the BITBM which describes the contribution of the individual taxa is the product in the numerator of equation 5.5 which is a quasi nested BITM. The transfer functions $\mathbb{P}_{T_k|\vec{C},\mathcal{B}_l}\left(t_{k_0}|\,\vec{c}_0, b_{l_0}\right)$ are directly estimated via GLM in the same way as for Lake Prespa (s. chap. 5.2.3 resp. 3.2.2). The indicator taxa are the same as those used in the estimation of the biome probability $W_{B_{l_0}}$. The transformation of pollen counts to presence and absence information summarized in the required indicator taxa matrix $\mathcal{I}_{\vartheta_0}^{N_\iota \times N_{k_0}}$ is done with the same approach as for Lake Prespa (s. chap. 5.2.3 resp. 3.2.1). In contrast to the palaeoclimate reconstruction of Lake Prespa only six taxa are the past proxy database so that no mahalanobis distance filtering is required. The biome likelihood or biome transfer function $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}|\,\vec{c}_0\right)$ has the same functionality as in the BBM and is also estimated with a QDA (s. chap. 5.2.4 resp. 4.2.3). Due to the fact that biome setup 3 is rated as the best and selected for the palaeoclimate reconstruction for Lake Kinneret and Ein Gedi it is also applied here. The marginal probability $m_{\mathcal{B}_l}\left(b_{l_0}\right)$ in equation 5.5 is defined as the sum over the numerator of this equation. This normalisation ensures that the BITBM result fulfils the mathematical definition for a probability.

All parts of the BITBM are applied to estimate four different palaeoclimate reconstructions of a three dimensional climate state vector $\vec{C} := \left(T_{Jan}, T_{Jul}, P_{ANN}\right)^T$, $\left(T_{Jan}, T_{Jul}, CWD_{ANN}\right)^T$, $\left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, P_{ANN}\right)^T$ or $\left(T_{Jan}^{850hpa}, T_{Jul}^{850hpa}, CWD_{ANN}\right)^T$. The resulting marginal distributions of the Birkat Ram palaeoclimate reconstruction are presented in Figure 5.3 and 5.4. In summary it is noted that the climate change to higher summer temperatures (T_{JJA} and T_{JJA}^{850hpa}) with the onset of the Younger Dryas (YD) is very reliable so that this hypothesis from Schiebel (2013) is confirmed. The hypothesis of dryer environmental conditions could be only partly confirmed since only the palaeoclimate reconstruction applied with the surface based transfer functions clearly provide this. For the BITBM as a whole it becomes also obvious that it is not as robust against anthropogenic outliers/influences as BITM but not so sensitive as the BBM. But for a reliable conclusion the BITBM has to be applied at a sediment core with more taxa. Overall it is important to note that the palaeoclimate reconstructions of P_{ANN} and CWD_{ANN} of Birkat Ram in general are very sensitive to the setup of the climate state vector $\vec{c}$, more precisely: What are the other components (middle troposphere temperatures (850 hPa level) or surface temperatures)?

Jordan Valley

The successful CFR of Gebhardt (2003) and Simonis (2009) allows a better assessment of possible climate changes in the past. This motivates to apply a simplified version of their variational approach for the fossil sites situated in the Jordan Valley: Ein Gedi (EG), Lake Kinneret (LK) and Birkat Ram (BR). The basic hypothesis in their approach and the approach presented here is the possibility to describe the past climate $\vec{c}_0$ as deviation $\vec{c}\,'$ from the modern climate state $\vec{c} := \vec{c}^{\,mod}$ (s. eq. 6.1). Past climate in this context means to evaluate the climate state of a certain defined time slice at all considered locations. The analysis area is the Jordan Valley. Its definition bases on the orography of the E-OBS climate dataset (s. chap. 6.1 for details). The CFR for the Jordan Valley bases on the minimization of a cost function. In this study the CFR is a interpolation of the local palaeoclimate reconstructions. The simplifications of mathematics of this cost function from Gebhardt (2003) and Simonis (2009) are described in chapter 6.2. The cost function in the approach applied here summarizes the contributions of the paleoclimate reconstruction of Ein Gedi, Lake Kinneret and Birkat Ram for one time slice in a vegetational part $\mathcal{J}_{veg}$ and the derivation $\vec{c}\,'$ in a model part $\mathcal{J}_{Mod}$. The minimization result of the cost function is a CFR which fits best to the palaeoclimate reconstruction at all locations and also quasi linear profile which is used as model part. This quasi linear profile is a GLM with quadratic term which takes into account the modern climate by application of $\vec{c}\,'$.

Directly after these defintions follows a discussion of the adaptations of the local palaeoclimate reconstructions. These adaptations are that all palaeoclimate reconstructions applied in this variational approach use transfer functions which are estimated with a mixture of two climate datasets: E-OBS and CRU TS 3.1. All of these palaeoclimate reconstructions show structures indicated in the marginal distributions which are very similar to the reconstructions based on the CRU TS 3.1 dataset alone. The most important features of the local palaeoclimate reconstructions persist. Seven time slices are selected for the CFR: 910 y. cal BP, 1200 y. cal BP, 1790 y. cal BP, 1970 y. cal BP, 5570 y. cal BP, 6740 y. cal BP and 7190 y. cal BP. For this selection the most reliable age-to-depth model of the Ein Gedi sediment core is choosen as the "reference" date. Then all pollen sample layers with an one-to-one and onto assignment of their age-to-depth model to the Ein Gedi date are selected.

For these time slices the cost funtion is minimized. Summarized the Jordan Valley CFR results in a dryer palaeoclimate than the modern climate with P_{ANN} values around 200 $\frac{\mathrm{mm}}{\mathrm{year}}$ for all fossil sites and considered time slices. If the overlap as well as the range of the error bars is take as a measure of certainty of the palaeoclimate reconstruction Ein Gedi is those with the most reliable result followed from Lake Kinneret and than from Birkat Ram. Also the result of the local reconstructions in the previous chapters of no palaeoclimate change in the summer and winter temperatures is confirmed. For each of the reconstructed Jordan Valley profiles the uncertainty of the palaeoclimate profile grows with increasing distance to the intercept. This is in accordance with a general characteristic of regression problems: The error growth with growing distance to the intercept. Also valid is that the climate anomaly defined by equation 6.26:

$$^{s}c_i' := {}^{s}c_{i,R,0,h_{mean}} - {}^{s}c_{i,h_{mean}}$$

fulfils equation 6.28

$$\left|{}^{BR}c_i'\right| \geq \left|{}^{LK}c_i'\right| \geq \left|{}^{EG}c_i'\right|$$

for all time slices.

But all quantitative results for the temperatures and the annual precipitation amount have to be handled with care since the Jordan Valley CFR approach is problematic. Mathematically and numerically the method works but there are some problems with the input data additionally to problems of the local reconstructions as discussed in the analyses of figures in chapter 6.3. If this discussion is extended one question respective the input datasets arises: Why isn't E-OBS used allone as input climate dataset for the palaeoclimate reconstructions? One answer is that the covered area isn't large enough to estimate all required transfer functions alone, neither the biome nor the taxon specific transfer functions. For the biome this is especially the extended distribution area of the Irano-Turanian which lies partly outside of the E-OBS dataset area. Another one is that a re-digitalization of the distribution maps has to be done in which the question of representation arises: Is the distribution represented by the maps really sensitive enough for the high geographical resolution of the E-OBS dataset? With the mixing of both dataset in the estimation of the transfer functions this problem is partly avoided since this is quasi equal to an averaging of both. The only effect of the averaging is a slightly reduced $BSS_{\ddot{\dddot{B}}_l}$. Additionally it is mentioned in chapter 4.2.3 that the main characteristics of the marginal distributions persist. Due to the other problems in all parts of the palaeoclimate reconstructions in the Jordan Valley this re-digitalization is not done.

Another question is: Why isn't another dataset used? First the CRU TS 3.1 data alone has a too coarse spatial resolution (s. chap. 6.1). Another gridded dataset based on observations with a high resolution for at least the complete Eurasian continent and the Mediterranean area, which is required, isn't available (as known by the author). Additionally all known available reanalyses with a geographical high resolution have the same problem that the required area is not available or is situated in the fringe.

If this availability of geographical high resolution climate data is solved there are also other difficulties: One main problem is that the local reconstructions, especially Birkat Ram and Lake Kinneret, have to be optimized (details in chap. 6.5.3). Switched back to the Jordan Valley CFR approach there are also methodical problems: The selection resp. definition of the intercept position has a large influence on the whole result. An equal weighting of all fossil sites is not possible, even if the factor $\Delta \sin \varphi_s$ (s. chap. 6.2.3) is incorporated in the minimization. Another approach, as for example the CFR introduced by Gebhardt (2003) could solve this problem. If this is not done the question of representation of modern Jordan Valley climate by the GLM profile as introduced in chapter 6.1 has then to be solved.

7.2. Outlook

The outlook and suggestion for further work is directly connected with the last mentioned difficulties of representation of modern Jordan Valley climate by the GLM profile. Another possibility of a CFR is a Bayesian Hierarchical Model (BHM). There have been great efforts in the science of climate reconstructions especially in the last decade as mentioned in the introduction part on page XI. The paper of Tingley *et al.* (2012) which presents the concept of BHM with multiple levels for palaeoclimate reconstructions brought this science on a mathematically robust level. As pointed out by Tingley *et al.* (2012), it is possible to reconstruct climatic fields if this aim or more precisely level is incorporated in the definition of the applied BHM. One part of the ongoing Paleo Modelling (PalMod) project WG3 is exactly the implementation of the approach from Gebhardt (2003) as a so called process level in a BHM (Nils Weitzel, pers. comm.), so that it is possible to reconstruct physically consistent climatic fields. Another aim in this PalMod project is to apply a complex BHM on real proxy data (Nils Weitzel, pers. comm.). In contrast Tingley *et al.* (2012) presents the application of BHM with simulated proxy data. Nils Weitzel (pers. comm.) also explores the extension of the so called data level in the BHM: He designed Markov Chain Monte Carlo (MCMC) to estimate the regression parameters of tranfer functions instead of the maximum likelihood estimation (MLE) applied in this work. Another suggestion, which is also investigated in the PalMod project, is the incorporation of different proxies in the BHM, called classically multiproxy approach. If this is sucessfully included it is also possible to reconstruct climate parameters which are available over land and ocean. At the moment it is only feasible to consider either land or the ocean.

Another suggestion is the extension of the verification part of the Diplomathesis of Stolzenberger (2011). Stolzenberger (2011) verifies transfer functions based on a two dimensional climate state vector $\vec{C} := (T_{Jan}, T_{Jul})^T$ for some selected taxa. Her research was to find the best climate input dataset for the estimation of transfer functions. Schult (2013) verifies transfer functions estimated from European taxa distribution maps (cf. chap. 2.2.1) with taxa distribution maps available for Northern America with the same tools as Stolzenberger (2011). Some of this research can also be done with the R-package biomod2[1]. This research has to be extended especially for the complete taxapool mentioned in chapter 2.2.1 and also the biome distributions. The ongoing Bachelor thesis of Lisa Wacker (supervised by Prof. Dr. Andreas Hense and Prof. Dr. Thomas Litt) looks at the spatial variation of biomes in the Dead Sea basin when changing the amount of precipitation. If this sensitivity analysis is finished and in the best case extended by temperature it would be also a promising task to apply the BITBM for another contemplable sediment core with more taxa. Sophie Stolzenberger explores in her ongoing PhD thesis the comparison between model simulations which are taken from the PMIP3 database and GLM- based climate reconstructions based on pollen data. The aim is to optimize the model data by including the probabilistic information of the occuring taxa. A clearly

---

[1] `https://cran.r-project.org/package=biomod2` for details

change in summer temperatures can be detected for the Mid Holocene over Europe (Sophie Stolzenberger pers. comm.).

One other suggestion is that the time scale of the available pollen proxy records in Israel has to be extended backwards in the past to investigate also the suggested climate change in the summer temperature (T_{JJA} and T_{JJA}^{850hpa}) with the onset of the YD in the Birkat Ram palaeoclimate reconstruction with a CFR. Stein *et al.* (2011a,b) extend the time scale for Ein Gedi with a new sediment core drilled within the Dead Sea Deep Drilling Project (DSDDP) campaign 2010/2011 which is part of the International Continental Scientific Drilling Program (ICDP). The established age-to-depth model from Torfstein *et al.* (2015, 2013) suggests that the core covers approximately the last150 000 y. cal BP. One part of Sophie Stolzenbergers PhD thesis (in preparation) is the palaeoclimate reconstruction by using the BBM and bases on the pollen counts (PhDthesis by Andrea Miebach and Chunzhu Chen (in preparation)) of this sediment core. In the P_{ANN} palaeoclimate reconstruction there is a clear identifiable climate change period in the time period between approximate 120 000 y. cal BP and 133 000 y. cal BP. Her results also confirm the conclusions of chapter 4.6 for the BBM for the condition when it is possible to reconstruct a period with climate change (Sophie Stolzenberger pers. comm.).

Mentioned in the introduction of this thesis was the motivation of the Collaborative Research Centre (CRC) 806: The answer of the question of origin of modern man and their way to Europe. This work provides a contribution to that by exploring periods with climate change on the suggested Eastern migration corridor. Climate change is one possible driver for movements of modern man since climate influences the availability of food. Food in turn is the main driver for migration. Raichlen *et al.* (2014) indicates a very interesting mathematically description of movements of human hunter-gatherers. It would be interesting to combine this model with the probabilistic climate reconstruction results of this thesis and the results from Sophie Stolzenberger and Nils Weitzel.

Appendix

Table of appendices

A

Lake Prespa

This appendix chapter presents additional information and figures for the palaeoclimate reconstruction of Lake Prespa which bases on the BITM.

A.1. Thresholds for the Determination of the Indicator Taxa

The first section presents the table with all defined threshold values ϑ_{k_0} as required for the definition of the indicator taxa matrix $\mathcal{I}_{\vartheta_0}^{N_t \times N_{k_0}}$ (s. chap. 3.2.1). The second part presents the depth profile figures of the relative abundance $\vec{\widetilde{\omega}}_{k_0}$ and the corresponding empirical cumulative distribution function (ECDF) for each regarded taxon listed in Table A.1.

A.1.1. Threshold and Taxa Table

Table A.1.: Taxa used for the palaeoclimate reconstruction of Lake Prespa and the corresponding thresholds ϑ_{k_0}

No.	Taxon name	Shortcut	Threshold [%]
1	*Abies*	Abies_Prm	2.00
2	*Quercus cerris*	Quce_Prm	4.60
3	*Quercus robur*	Quru_Prm	5.50
4	*Quercus ilex*	Quil_m	10.5
5	*Fagus*	Fasy_f	0.90
6	*Alnus*	Algl_m	1.10
7	*Betula*	Betu_m	1.10
8	*Corylus*	Coav_m	2.50
9	*Fraxinus*	Frax_Prm	1.25
10	*Tilia*	Tilia_Prm	1.00
11	*Ulmus*	Ulmu_Prmf	0.75
12	*Phillyrea*	Phla_m	0.70
13	*Olea*	Oleu_m	0.35

continued on next page. . .

. . . continuation

No.	Taxon name	Shortcut	Threshold [%]
14	*Cornus mas*	Coms_m	0.00
15	*Hippophae*	Hirh_Prm	0.50
16	*Buxus*	Buse_m	0.00
17	*Polygonaceae undiff.*	Poav_Prh	0.00
18	*Polemoniaceae*	Poca__Prm	0.00
19	*Alisma cf. A. plantago-aquatica*	Alpl_Prm	0.00
20	*Typha*	Tyla_Prh	0.65
21	*Sambucus*	Sani_m	0.00
22	*Hedera*	Hehe_m	0.00
23	*Nymphaea*	Nyal_m	0.00
24	*Pistacia*	Pista_Pr	0.28
25	*Ephedra*	Ephed_Pr	0.60
26	*Chenopodium*	Cheno_Pr	2.40

A.1.2. Threshold ϑ_{k_0} Definition

The mapping of all following figures in this chapter is the same as for Figure 3.2 for taxon *Alnus*: Shown on the left side is the depth profile of the relative abundance $\vec{\widetilde{\omega}}_{k_0}$ and on the right side the corresponding ECDF. The depth profile of the relative abundance $\widetilde{\omega}_{\iota\,k_0}$ in each pollen sample ι is shown in dark grey. The light grey area is the ten times enlarged and the very light grey the 100 times enlarged profile. The threshold value ϑ_{k_0} is visualized by the different shaded areas in the depth profile and by the vertical bold black line in the ECDF. The resulting presence/absence information is visualized in the depth profile with black dots (=presence) and circles (=absence).

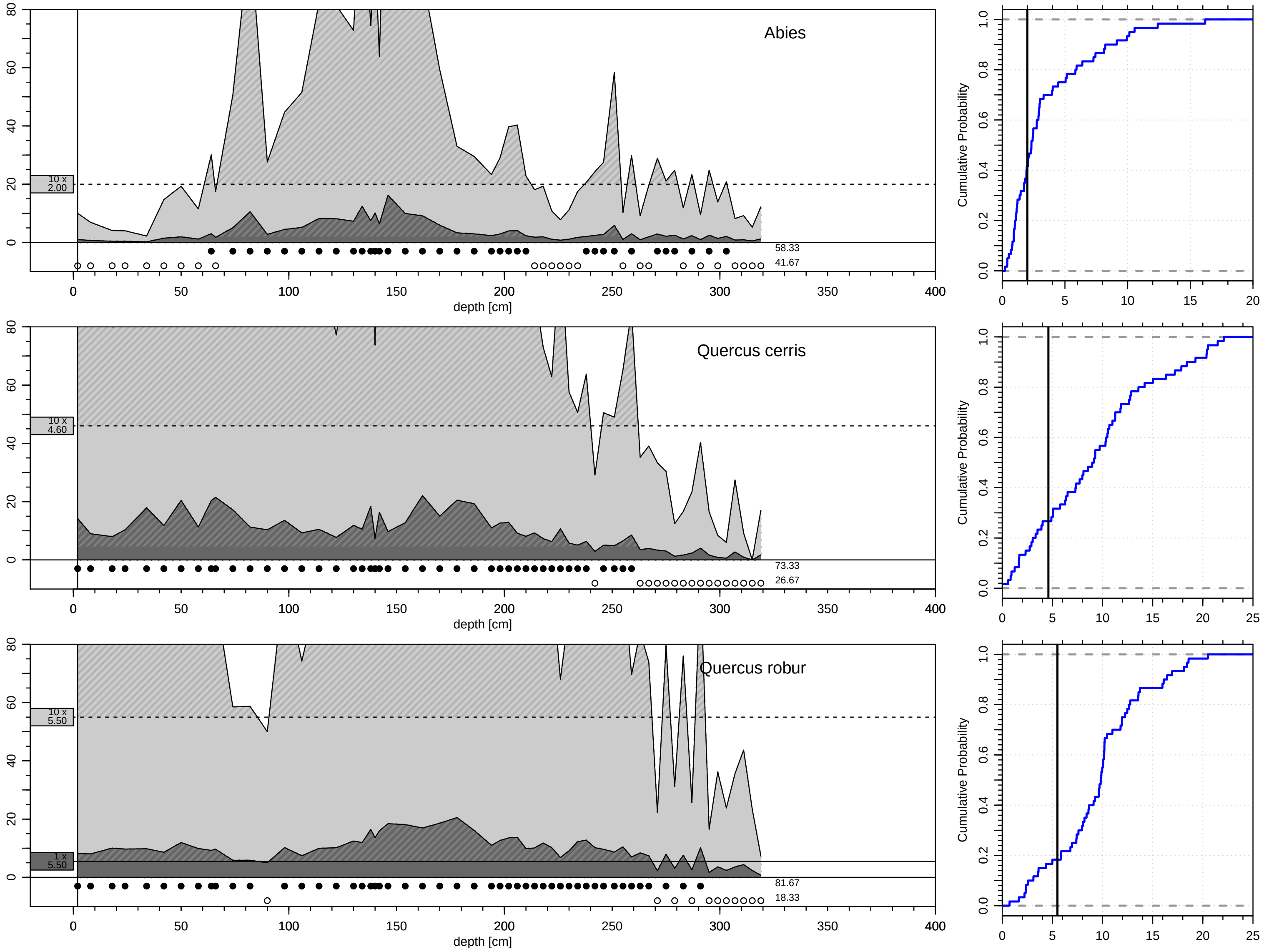


Figure A.1.: Lake Prespa, definitions of the thresholds (part 1).

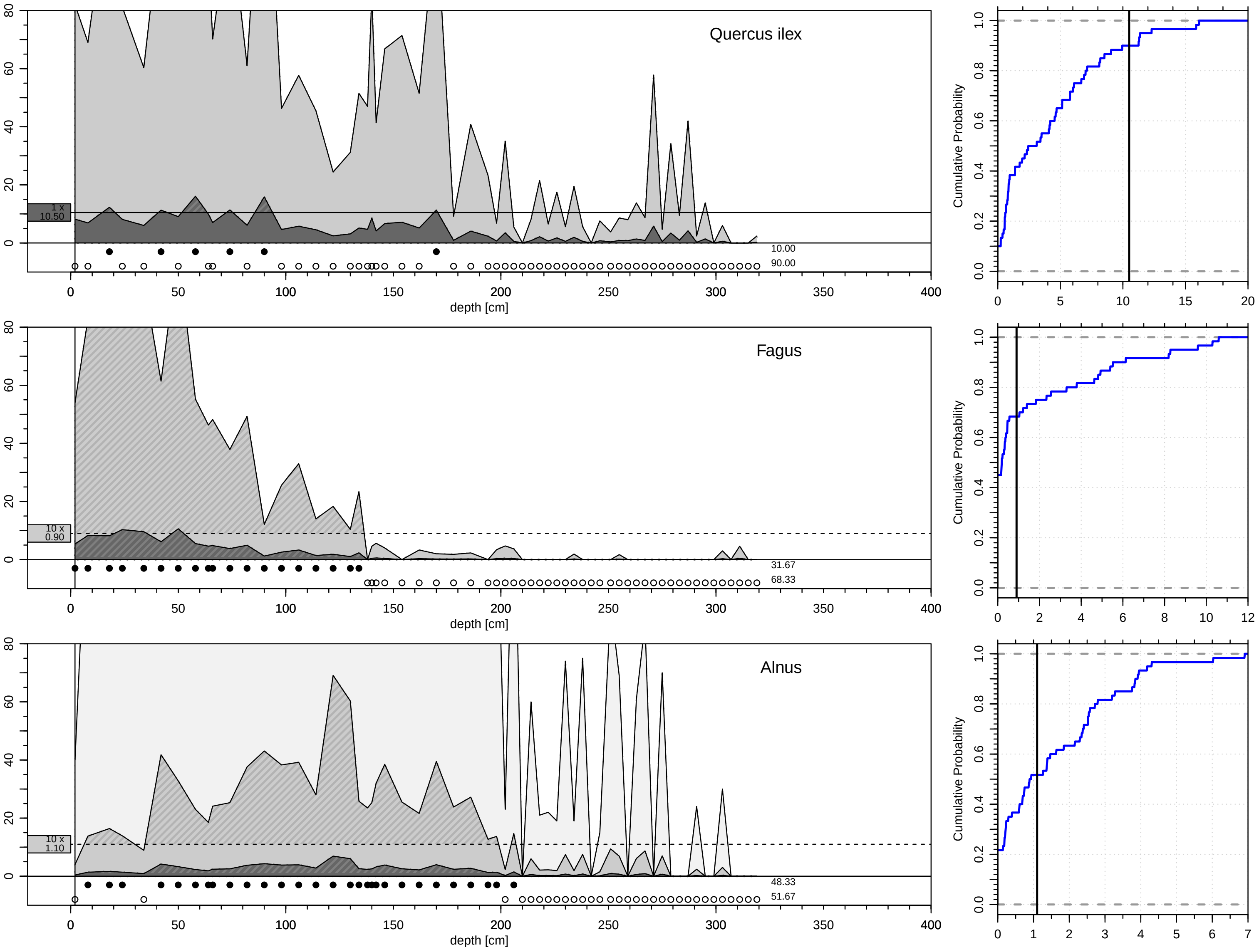


Figure A.2.: Lake Prespa, definitions of the thresholds (part 2).

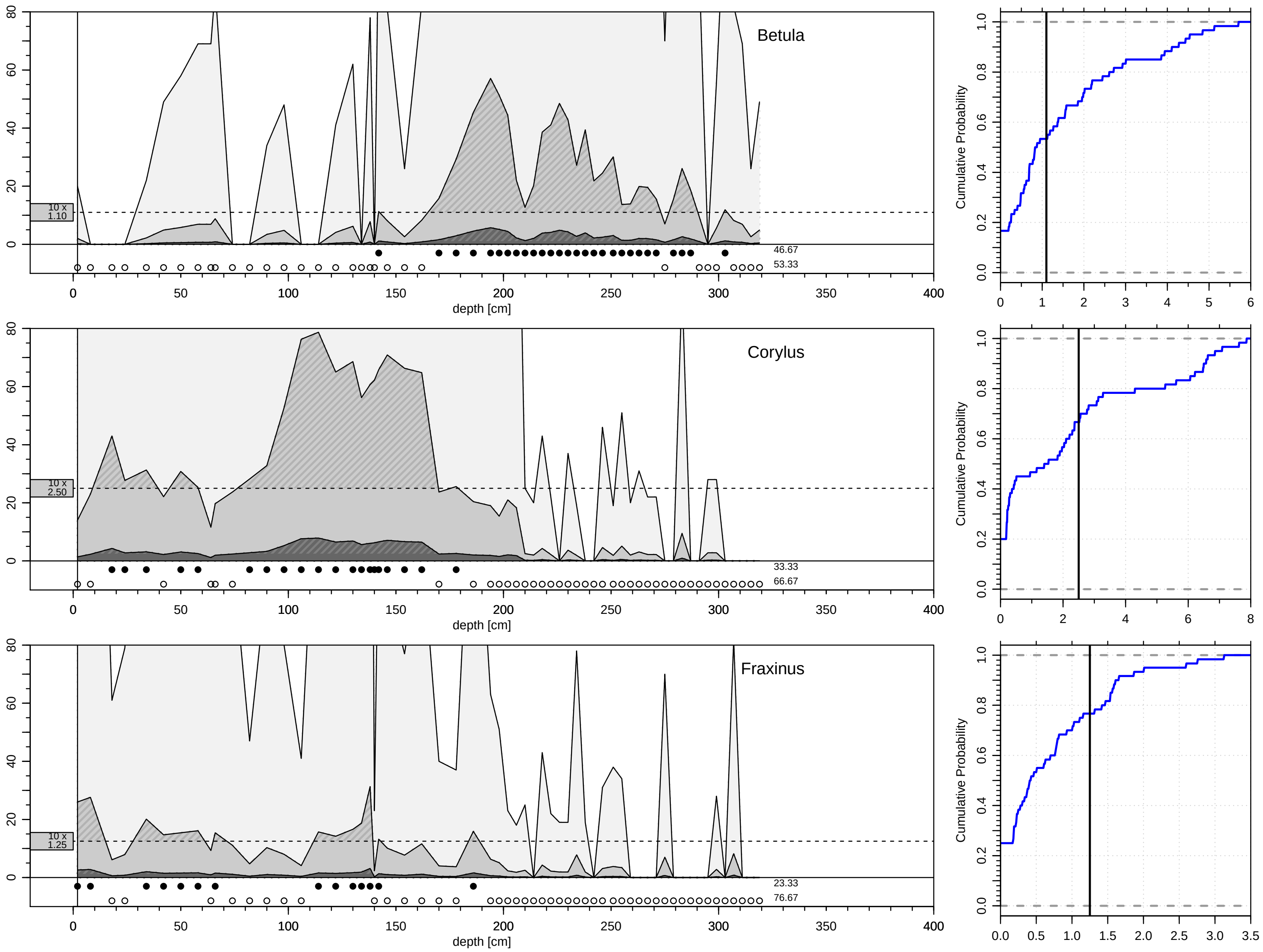


Figure A.3.: Lake Prespa, definitions of the thresholds (part 3).

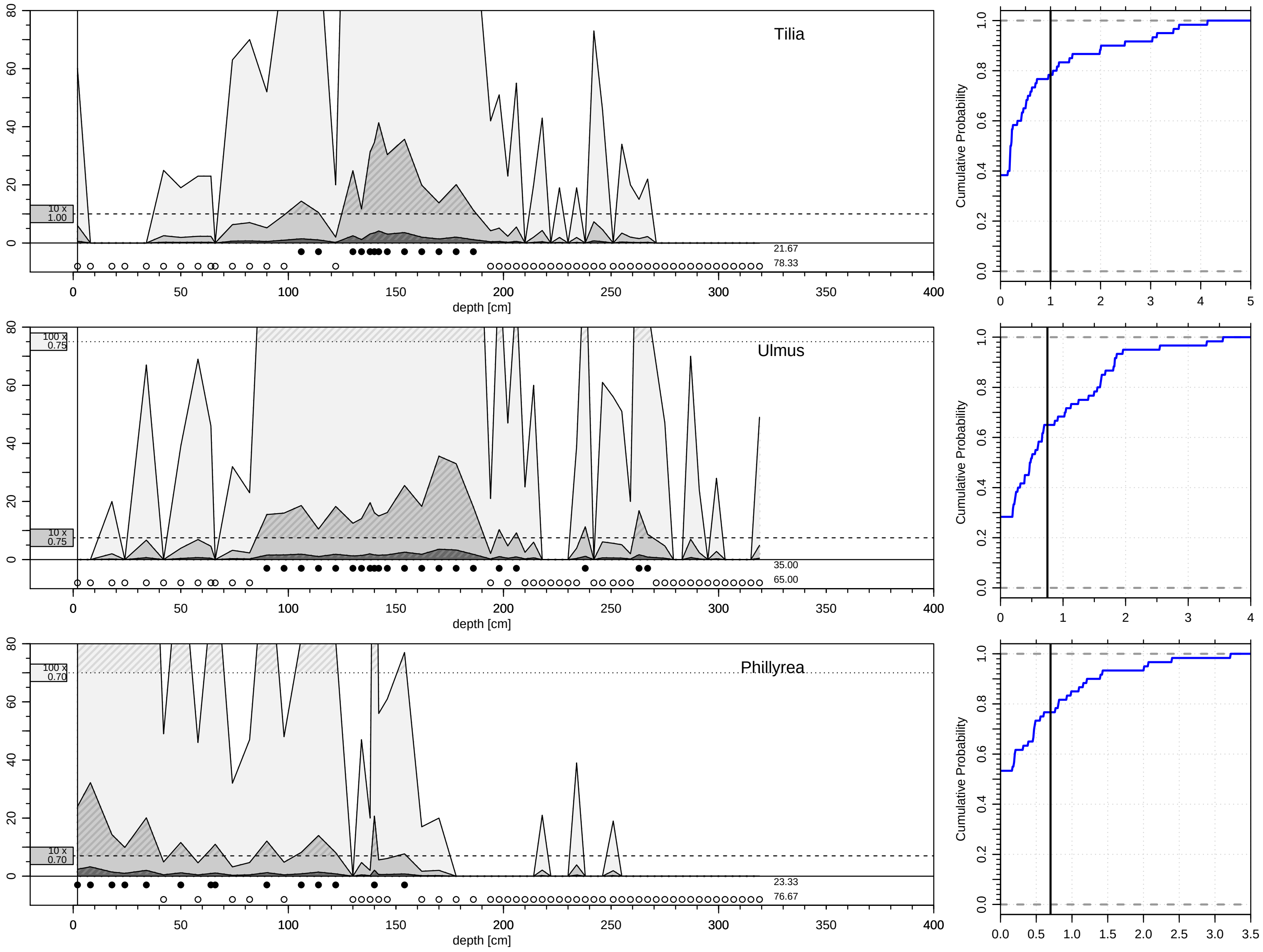


Figure A.4.: Lake Prespa, definitions of the thresholds (part 4).

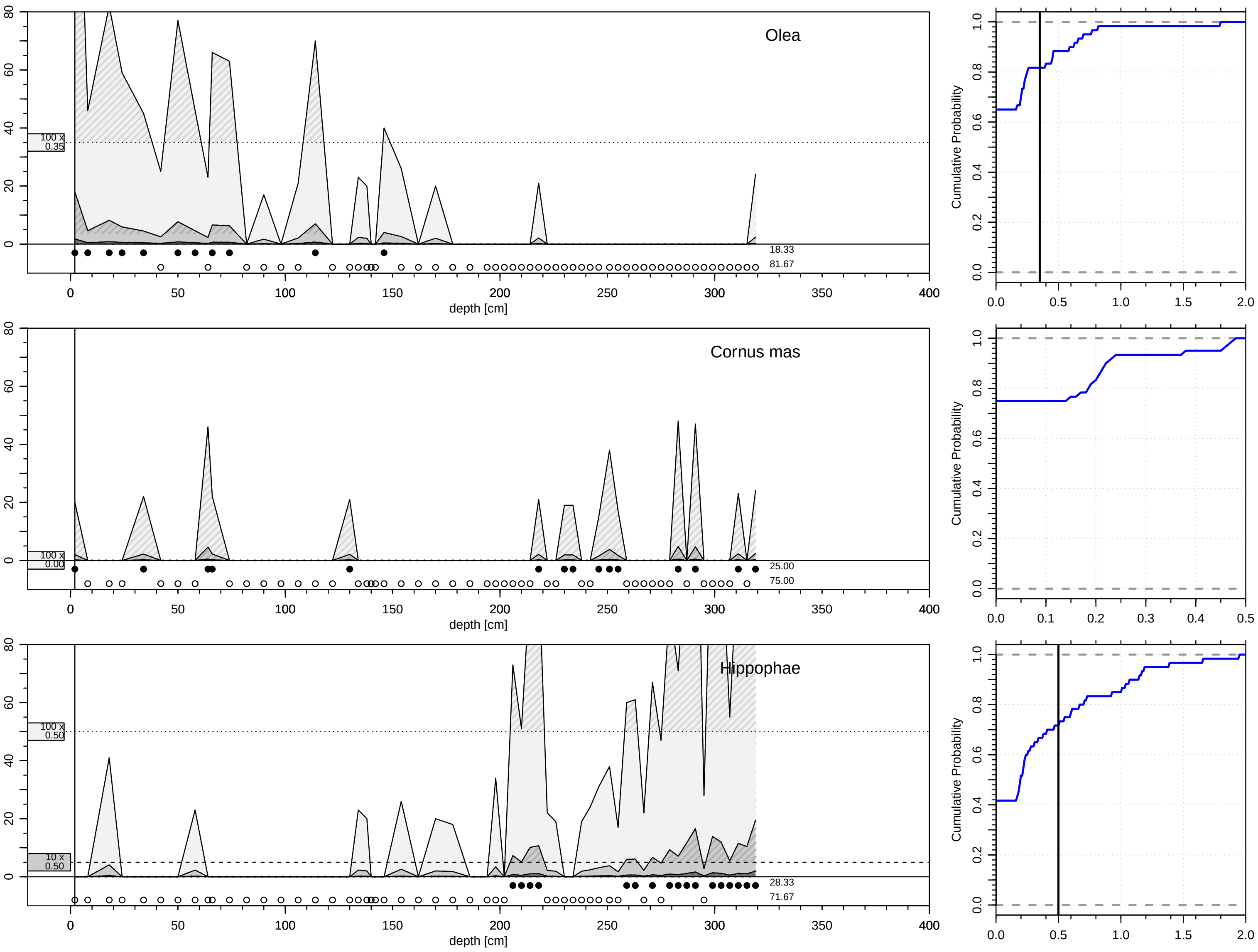


Figure A.5.: Lake Prespa, definitions of the thresholds (part 5).

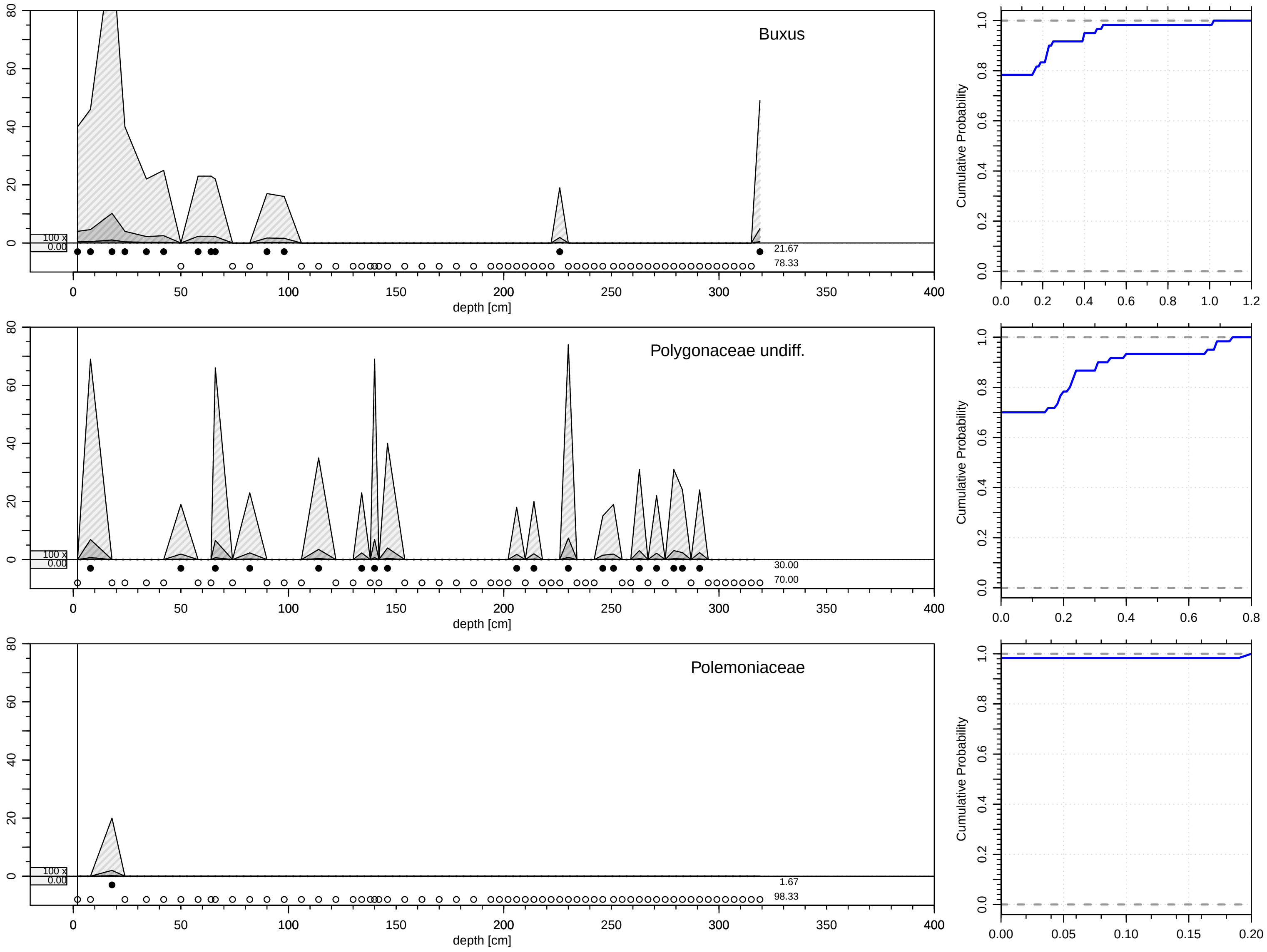


Figure A.6.: Lake Prespa, definitions of the thresholds (part 6).

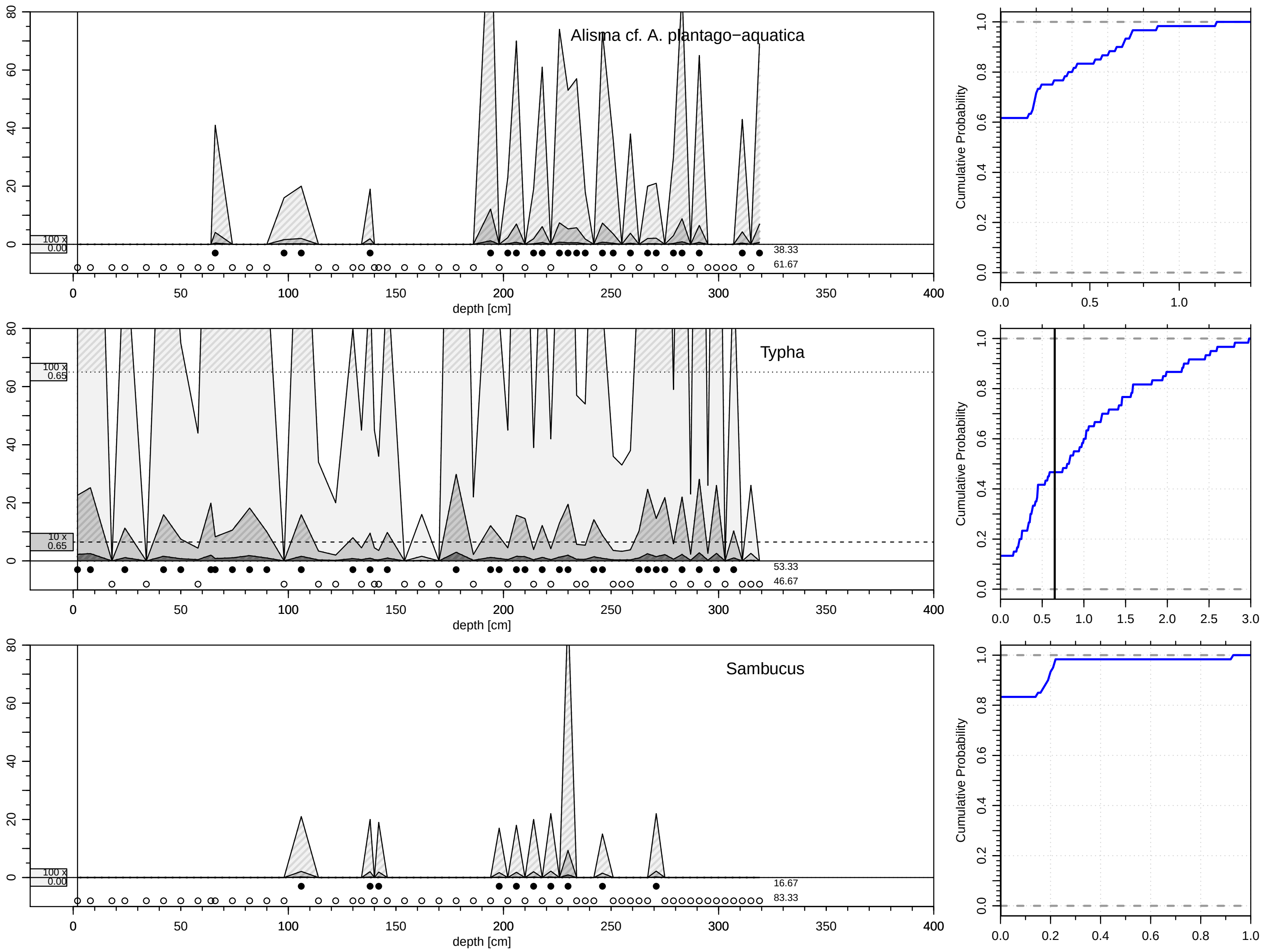


Figure A.7.: Lake Prespa, definitions of the thresholds (part 7).

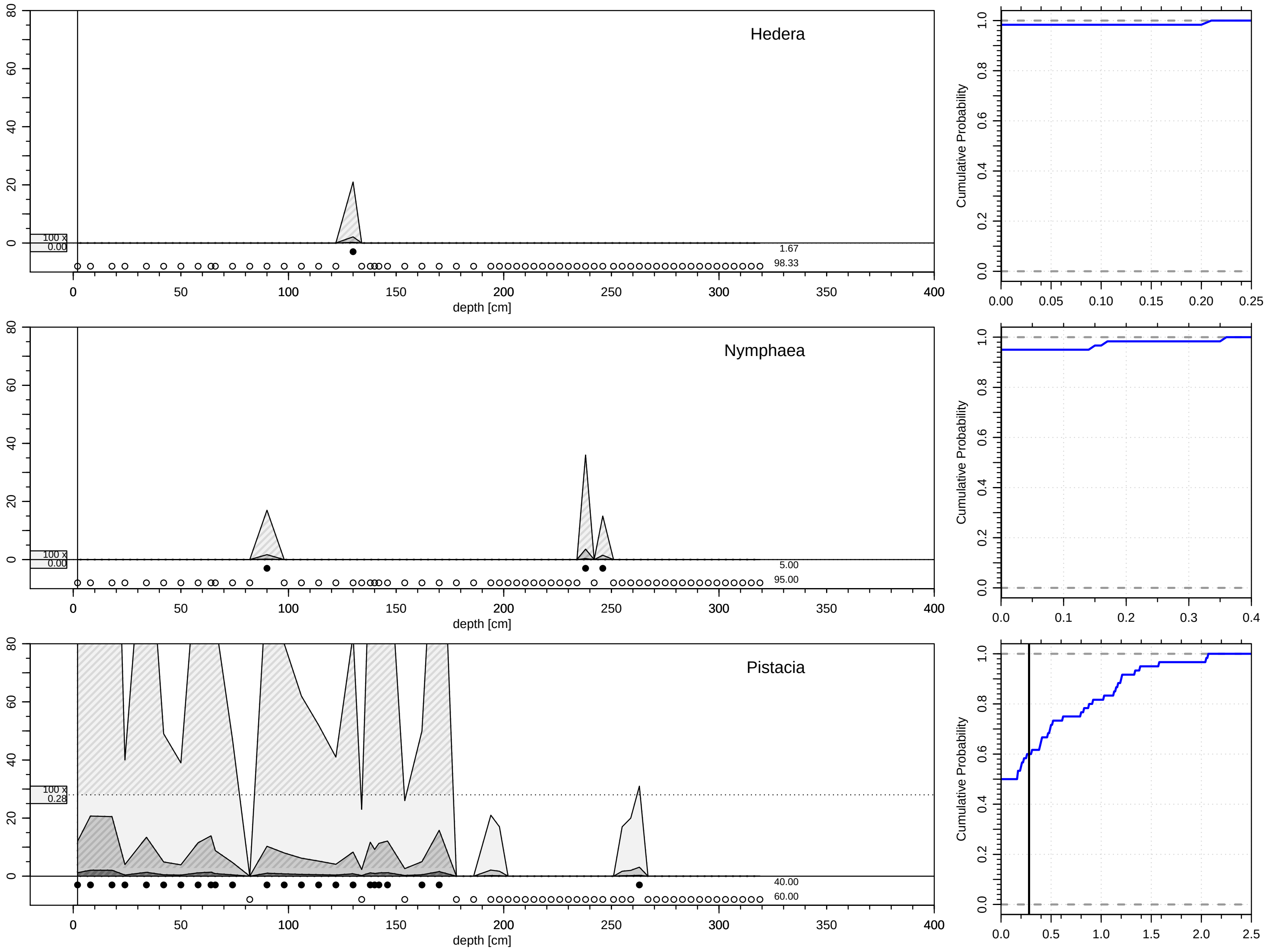


Figure A.8.: Lake Prespa, definitions of the thresholds (part 8).

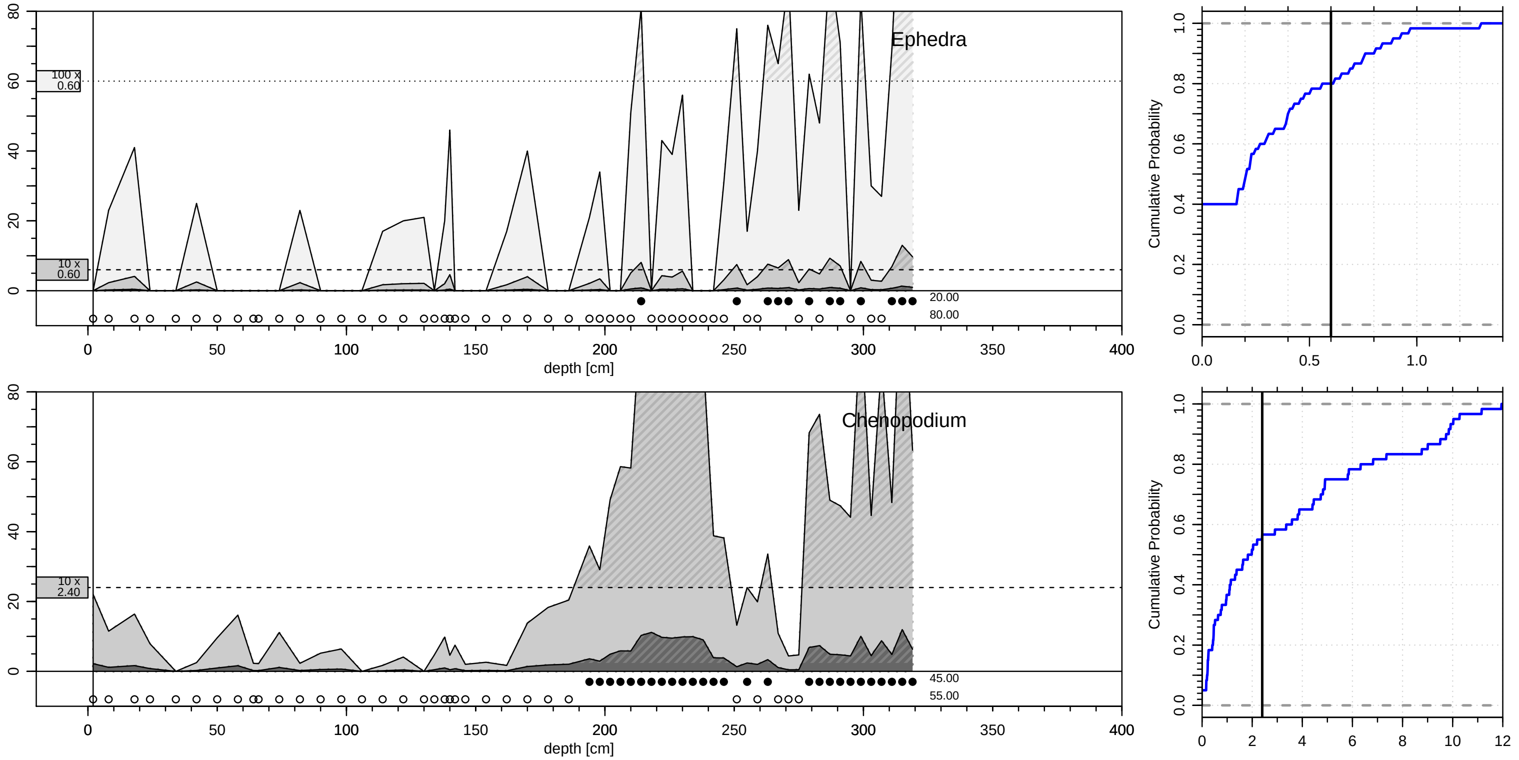


Figure A.9.: Lake Prespa, definitions of the thresholds (part 9).

A.2. Palaeoclimate Transfer Functions: Additional Figures

Figure A.10 shows the missing two dimensional marginal distributions for Figure 3.4 (s. p. 37). The estimation is described in detail in chapter 3.2.2. The mapping of Figure A.10 is identical to Figure 3.4:

Each cross represents one realisation $\vec{c}_{\lambda\varphi}$ over land surface at all grid points. Their colour marks the presence (yellow) or absence (grey) of the taxon. The contour lines and the colour scale represent the two dimensional marginal distributions of the estimated three dimensional GLM. Dark blue areas represent high values and light blue ones low values. The horizontal and vertical black line mark the maximum value of the estimated GLM.

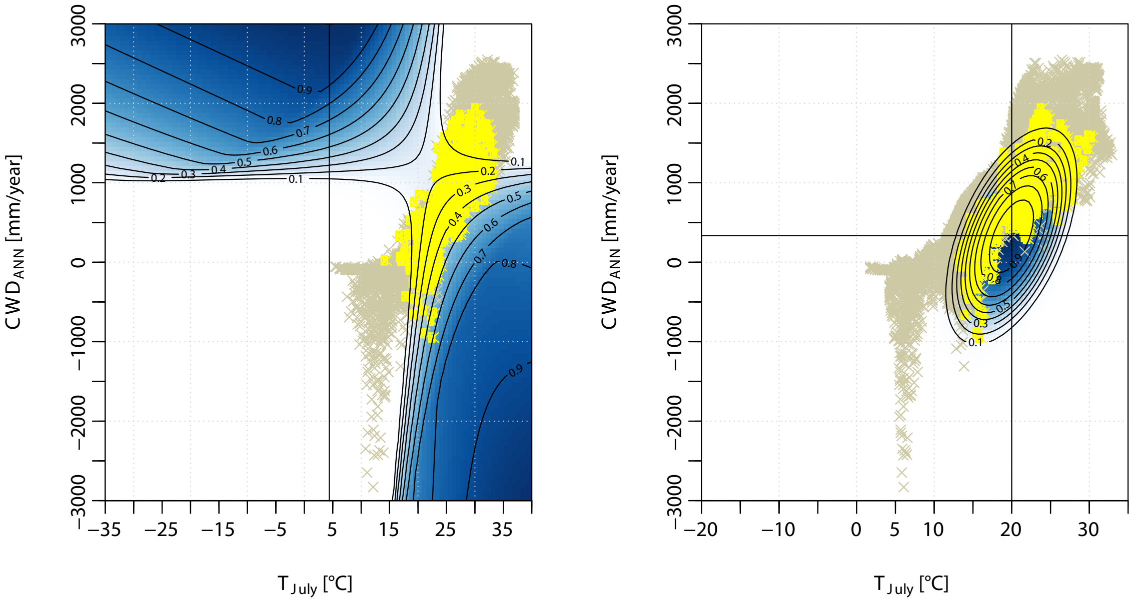


Figure A.10.: Remaining marginal distribution of the transfer function shown in Figure 3.4 for *Olea* and CRU TS 3.1 (left) resp. ERA (right).

A.3. Palaeoclimate Reconstruction Result: Additional Figures

Figure A.11 shows the filtered Indicator Taxa Matrix $\mathcal{I}^{N_\iota \times N_{k_0}}_{filtered}$ for Lake Prespa with age scale. Figure 3.3c is the associated figure with depth scale. The age-to-depth model is established by Aufgebauer *et al.* (2012) (mentioned in chapter 2.3.1). The mapping of the figure is identical to Figure 3.3c:

The absence information ($t_{\iota k_0} = 0$) for each regarded taxon k_0 is shown in two shades of light grey and the presence information ($t_{\iota k_0} = 1$) in two shades of dark grey. The pollen assemblage zones (PAZ) as defined in Panagiotopoulos *et al.* (2013) are marked in black and labelled on the top of the figure. The shortcuts on the vertical axis refer to the full taxon name in Table A.1.

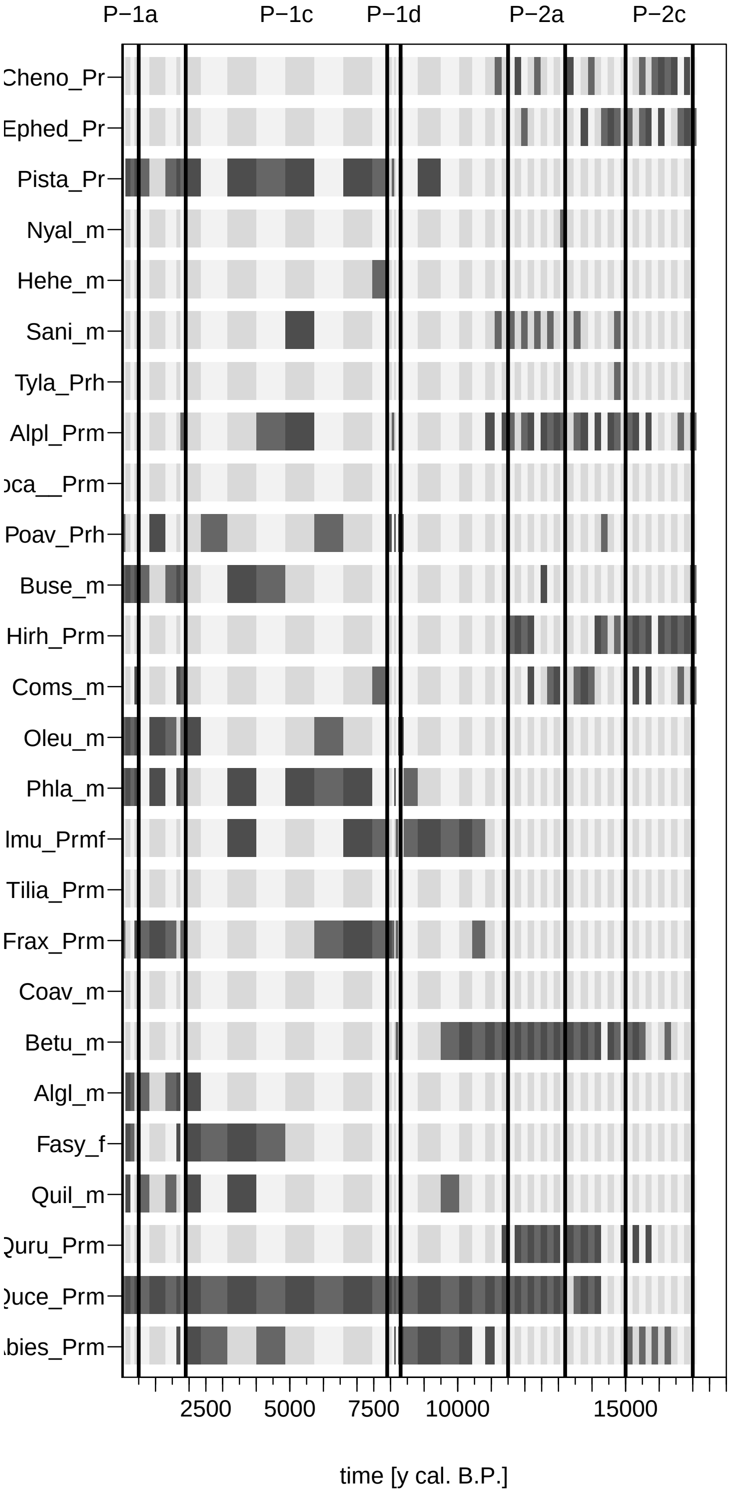


Figure A.11.: Indicator Taxa Matrix $\mathcal{I}_{filtered}^{N_t \times N_{k_0}}$ shown in Fig 3.3c for Lake Prespa with age scale.

B

Lake Kinneret

This appendix chapter presents additional information and figures for the palaeoclimate reconstruction of Lake Kinneret which bases on the BBM. The data tables for the PRM presented in chapter 4.4 can be found here, too.

B.1. Thresholds and Biome Assignment

The first section presents the table with all defined threshold values ϑ_{k_0} and entries of the biome assignment matrix $\mathcal{M}^{N_{k_0}\times N_{l_0}}$ as required for the estimation of the biome probability $W_{B_{l_0}}$ (s. chap. 4.2.2). The second part presents the depth profile figures of the relative abundance $\vec{\tilde{\omega}}_{k_0}$ and the corresponding ECDF for each regarded taxon listed in Table B.1.

B.1.1. Threshold and Taxa Biome Assignment Table

Table B.1.: Taxa used for the palaeoclimate reconstruction of Lake Kinneret, the assigned biome type and the corresponding thresholds ϑ_{k_0}.
The assigned biome type is indicated by the column vectors $\vec{m}_{l_0}$ of the biome assignment matrix $\mathcal{M}^{N_{k_0}\times N_{l_0}}$. It describes whether the regarded taxon k_0 belongs to this biome type l_0 ($m_{k_0 l_0} = 1$) or not ($m_{k_0 l_0} = 0$). In general the first column $\vec{m}_1$ represents a Mediterranean and the second column $\vec{m}_2$ a Irano-Turanian biome type.

No.	Taxon name	Shortcut	$\vec{m}_1$	$\vec{m}_2$	Threshold [%]
1	*Adonis type*	Adoni_K	0	1	0
2	*Anemone type*	Anemo_K	0	1	0
3	*Apiaceae*	Apiac_K	0	1	1.7
4	*Artemisia*	Artem_K	0	1	3.1
5	*Brassicaceae*	Brass_K	0	1	1.65
6	*Campanulaceae*	Campa_K	0	1	0
7	*Cannabis type*	Canna_K	0	1	0
8	*Caryophyllaceae*	Caryo_K	0	1	0.7

continued on next page...

. . . continuation

No.	Taxon name	Shortcut	$\vec{m}_1$	$\vec{m}_2$	Threshold [%]
9	*Centaurea*	Centa_K	0	1	0.5
10	*Cerealea type*	Cerea_K	0	1	2.2
11	*Chenopodiaceae*	Cheno_K	0	1	4.5
12	*Cistus type*	Cistu_K	1	0	0
13	*Elatine type*	Elati_K	0	1	0
14	*Ephedra dyst*	Ep_dt_K	0	1	0
15	*Ephedra frag*	Ep_fg_K	0	1	0
16	*Ericaceae*	Erica_K	1	0	0
17	*Fabaceae*	Fabac_K	0	1	1.6
18	*Fumana type*	Fuman_K	1	0	0
19	*Hypericum type*	Hyper_K	0	1	0
20	*Juglans*	Jugla_K	1	0	0
21	*Liguliflorae*	Ligul_K	0	1	7.2
22	*Linum type*	Linum_K	0	1	0
23	*Malvaceae*	Malva_K	0	1	0
24	*Olea*	Olea__K	1	0	3.8
25	*Papaver type*	Papav_K	0	1	0
26	*Phillyrea*	Phill_K	1	0	0.3
27	*Pinus*	Pinus_K	1	0	2.6
28	*Pistacia*	Pista_K	1	0	1.2
29	*Plantago*	Plant_K	0	1	1
30	*Poaceae*	Poace_K	0	1	10
31	*Quercus calliprinos type*	Qu_ca_K	1	0	5.6
32	*Quercus indet*	Queru_K	1	0	0.45
33	*Quercus ithaburensis type*	Qu_it_K	1	0	3.8
34	*Rhamnus type*	Rhamn_K	1	0	0
35	*Rubiaceae*	Rubia_K	0	1	0.3
36	*Rumex*	Rumex_K	0	1	0.45
37	*Sanguisorba type*	Sangu_K	1	0	1
38	*Scabiosa type*	Scabi_K	0	1	0
39	*Styrax type*	Styra_K	0	1	0
40	*Thalictrum type*	Thali_K	0	1	0
41	*Tubuliflorae*	Tubul_K	0	1	1.95
42	*Vitis*	Vitis_K	1	0	0.3

B.1.2. Threshold ϑ_{k_0} Definition

The mapping of all following figures in this chapter is an extension of Figure 3.2: Shown on the left side is the depth profile of the relative abundance $\vec{\widetilde{\omega}}_{k_0}$ and on the right side the corresponding ECDF. The depth profile of the relative abundance $\widetilde{\omega}_{\iota\, k_0}$ in each pollen sample ι is shown in dark grey. The light grey area is the ten times enlarged and the very light grey one the 100 times enlarged profile. The threshold value ϑ_{k_0} is visualized by the differently shaded areas in the depth profile and by the vertical bold black line in the ECDF. The dots directly below the abscissa in the depth profile indicate whether the taxon in this pollen sample layer contributes to the biome probability $W_{B_{l_0}}$ in this layer (=black) or not (=circle). The coloured box indicates the entry $m_{k_0 l_0} = 1$ for the regarded taxon and is the assigned biome type presented in Table B.1: shaded in green represents a Mediterranean and shaded in red a Irano-Turanian biome type.

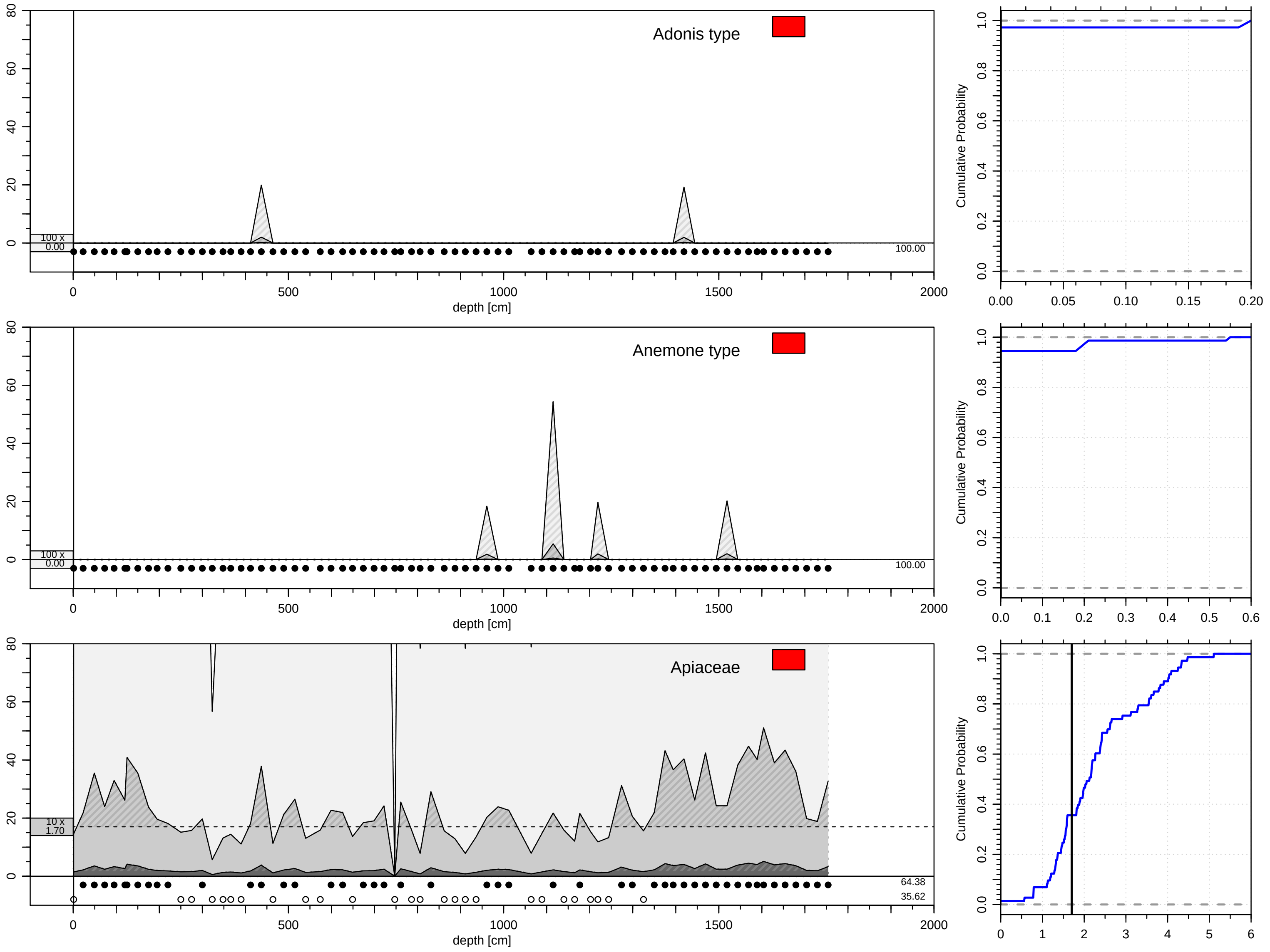


Figure B.1.: Lake Kinneret, definitions of the thresholds (part 1).

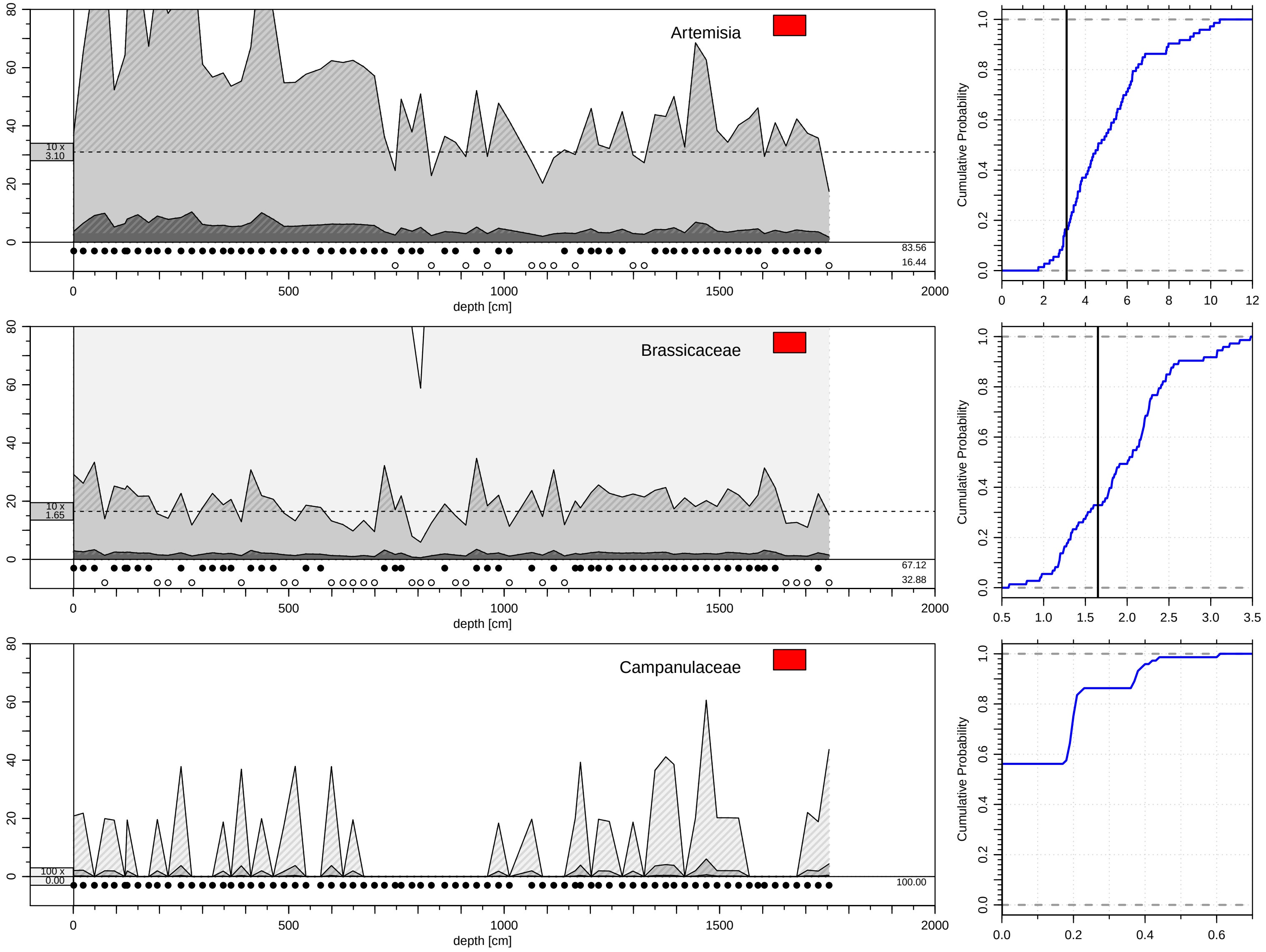


Figure B.2.: Lake Kinneret, definitions of the thresholds (part 2).

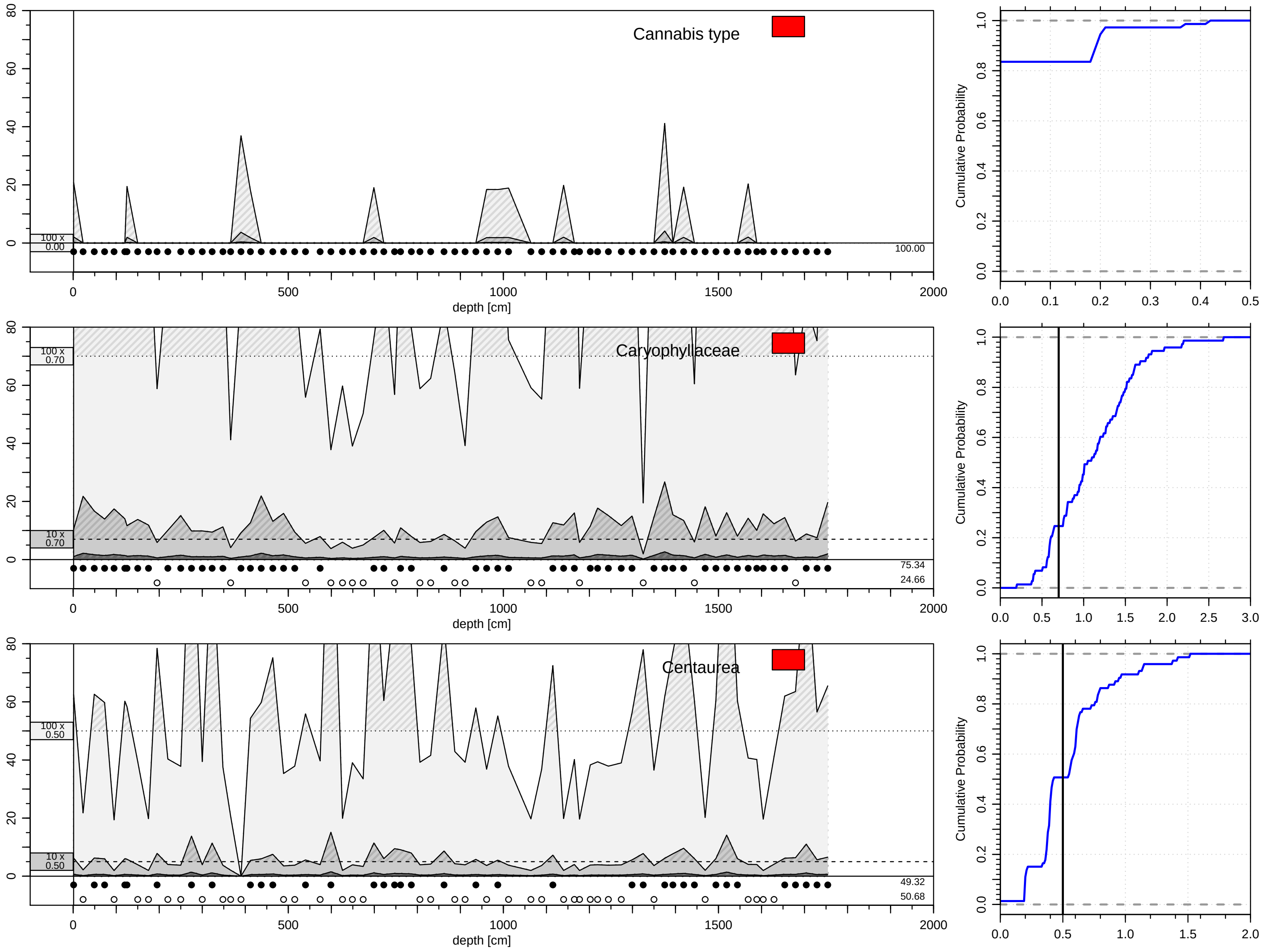


Figure B.3.: Lake Kinneret, definitions of the thresholds (part 3).

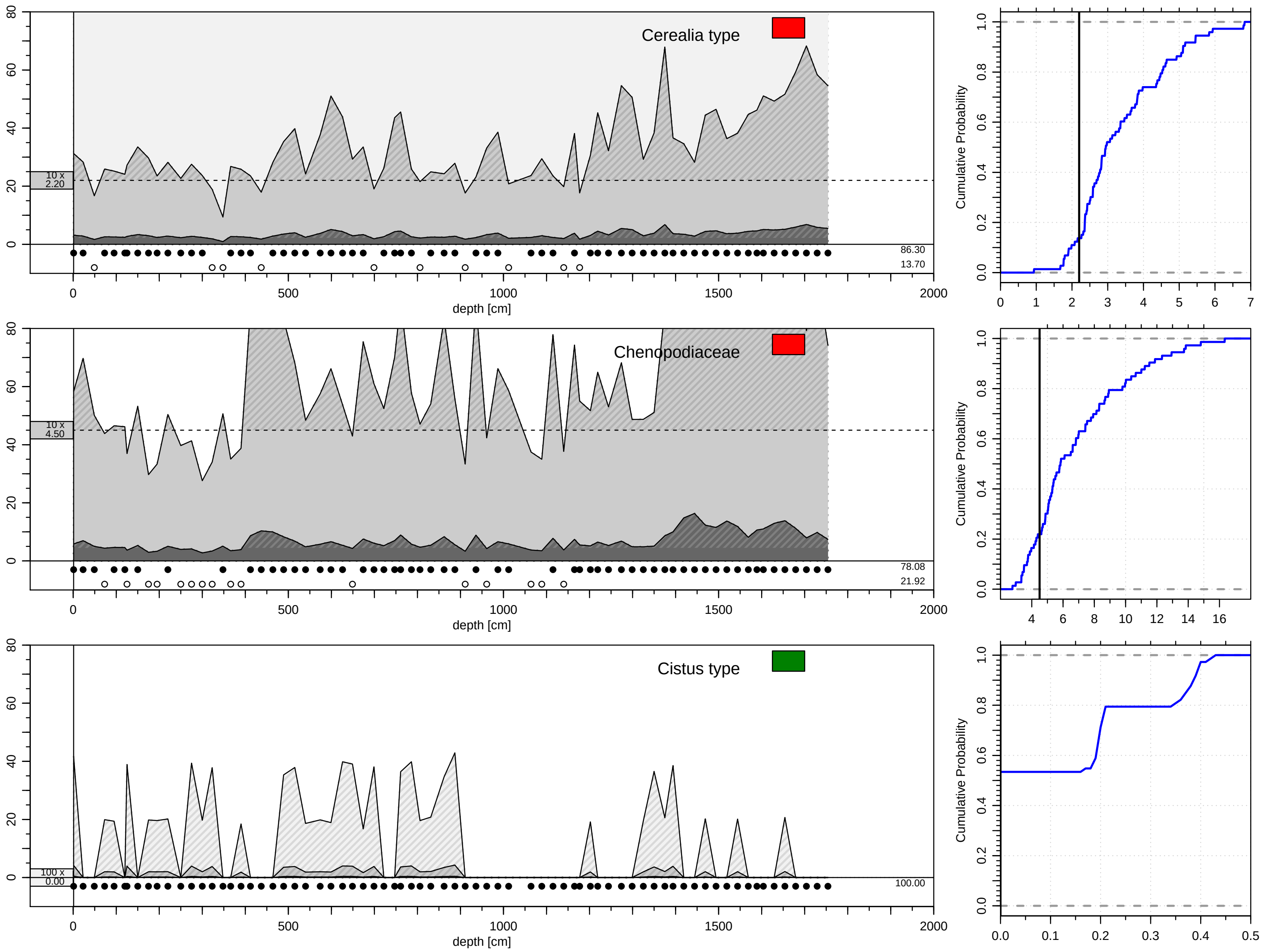


Figure B.4.: Lake Kinneret, definitions of the thresholds (part 4).

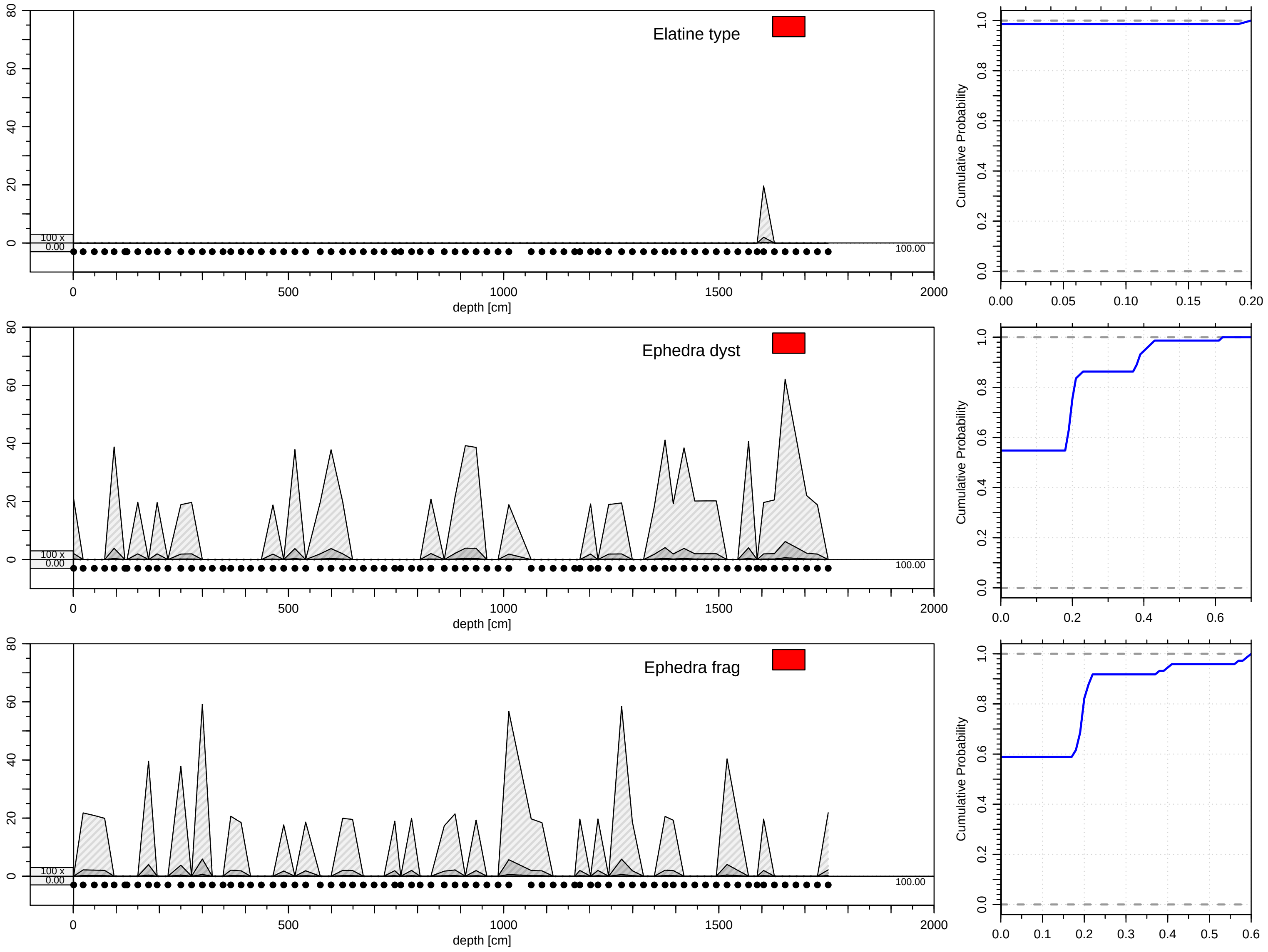


Figure B.5.: Lake Kinneret, definitions of the thresholds (part 5).

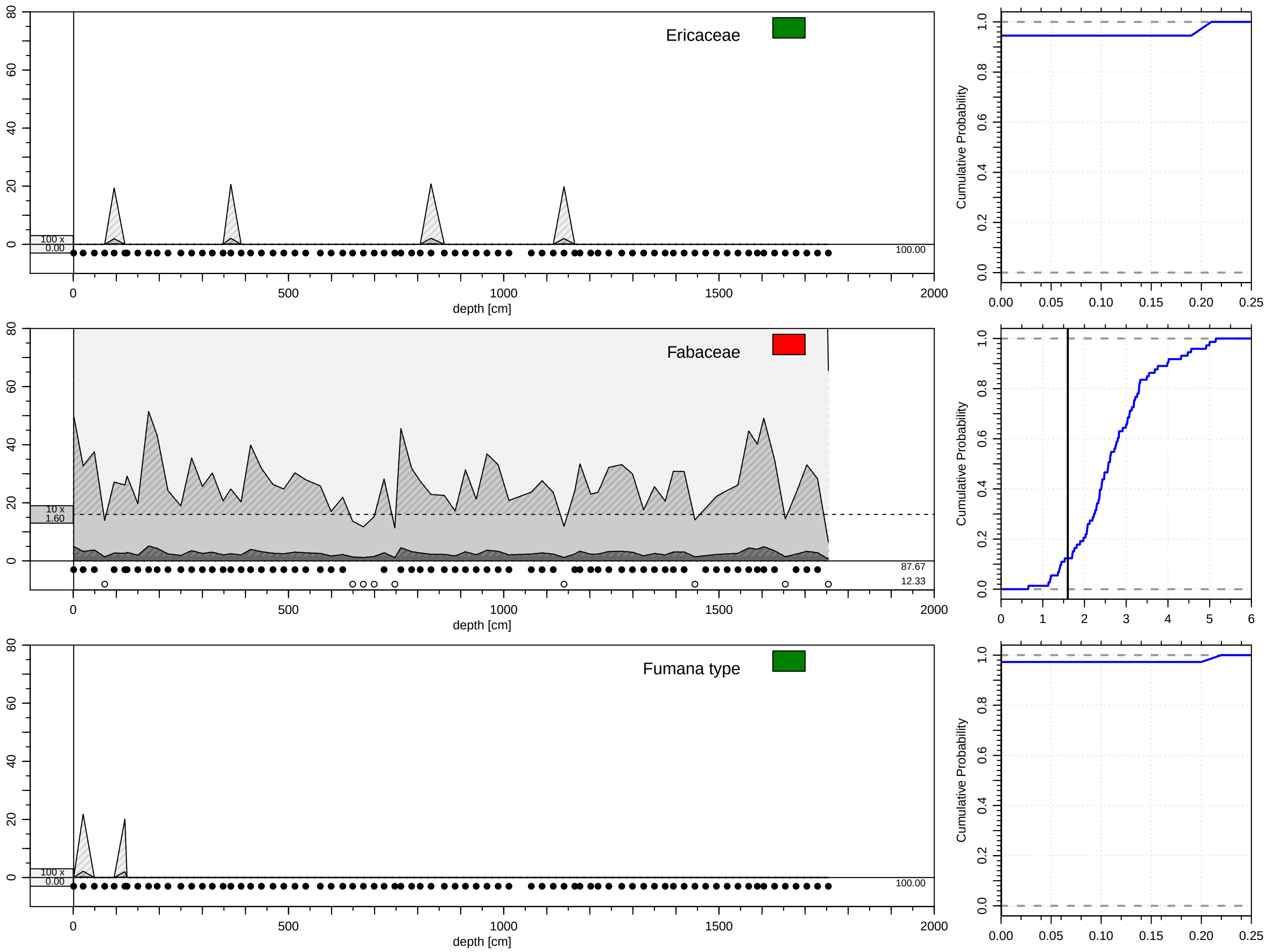


Figure B.6.: Lake Kinneret, definitions of the thresholds (part 6).

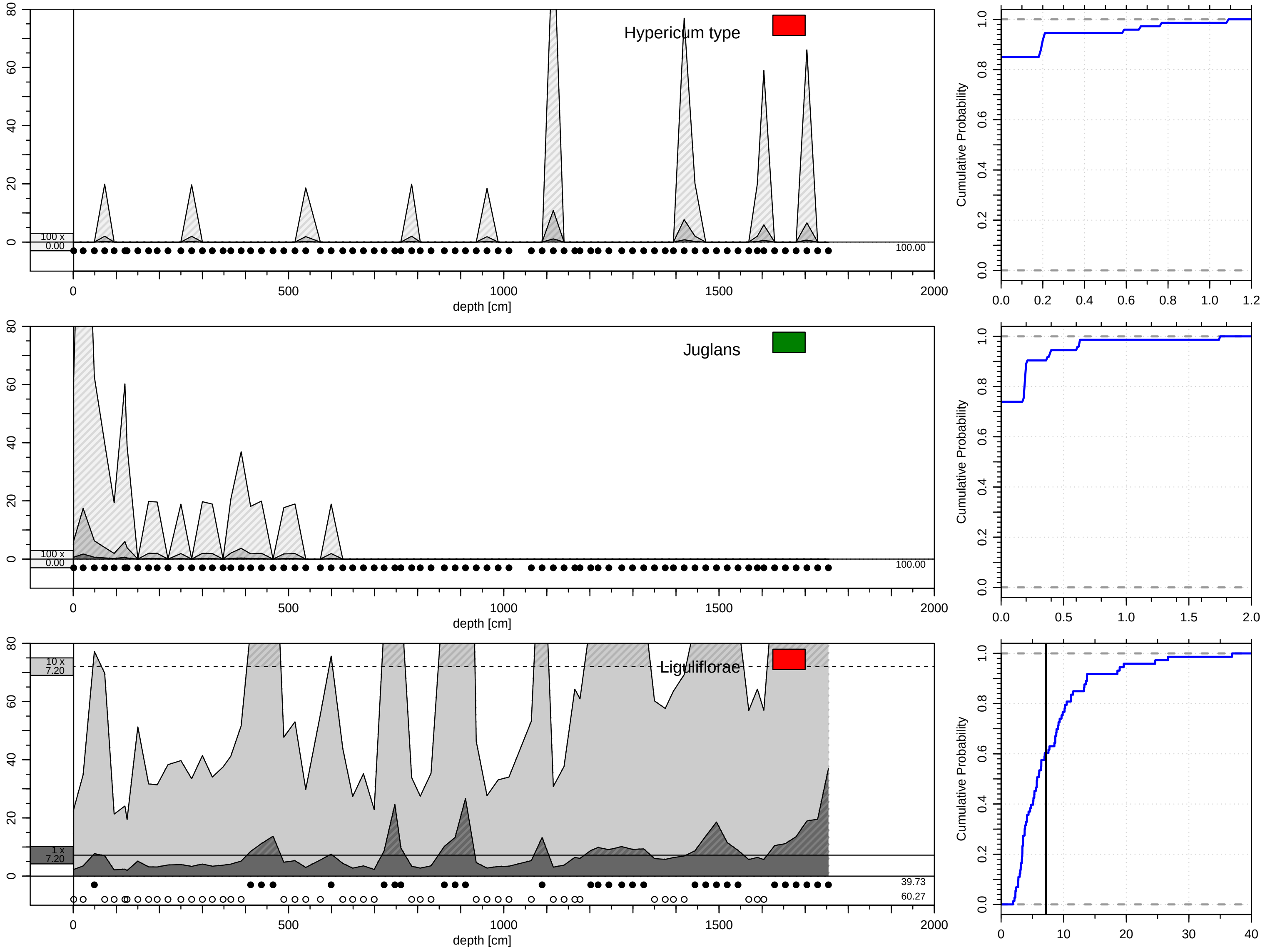


Figure B.7.: Lake Kinneret, definitions of the thresholds (part 7).

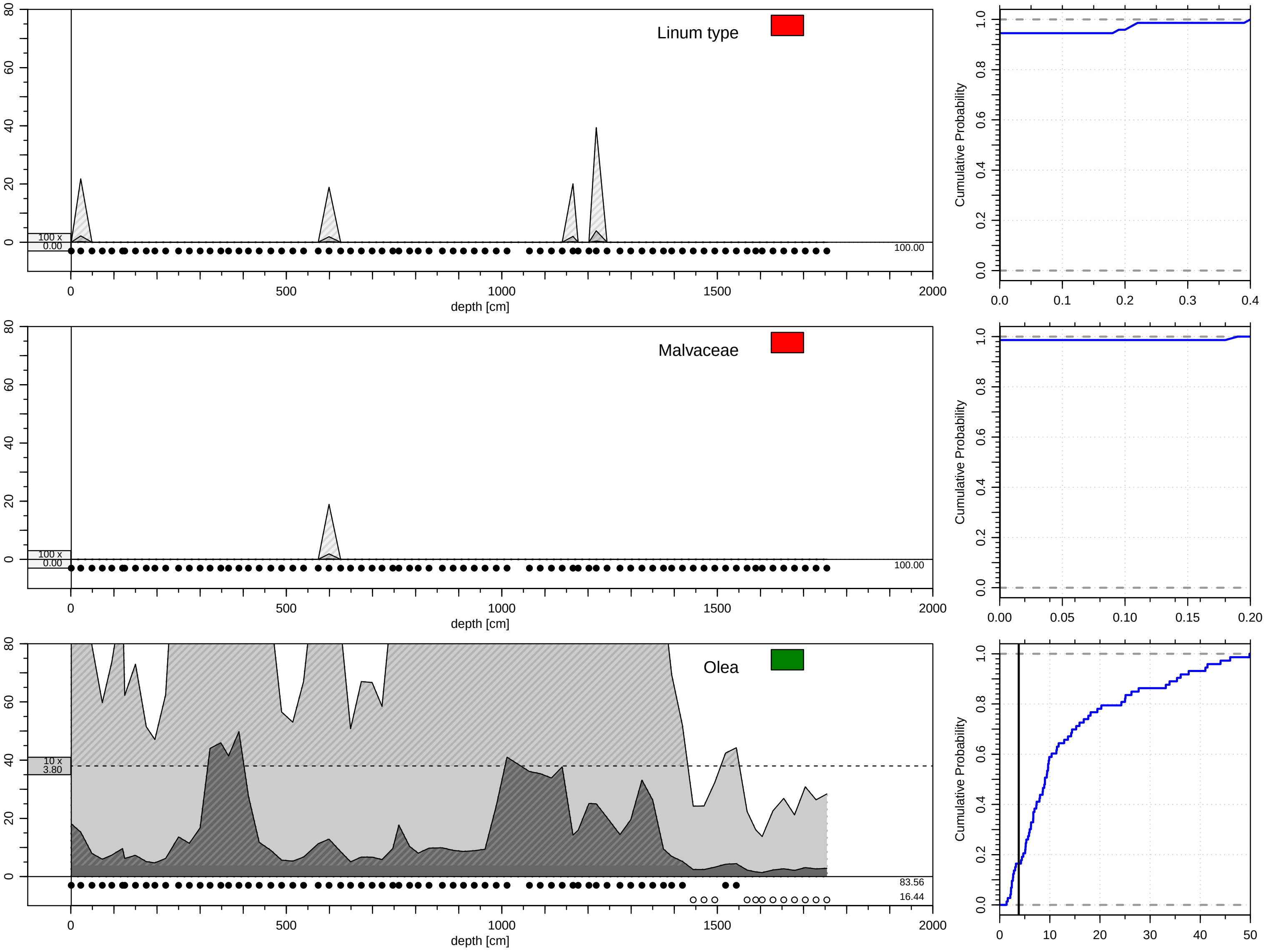


Figure B.8.: Lake Kinneret, definitions of the thresholds (part 8).

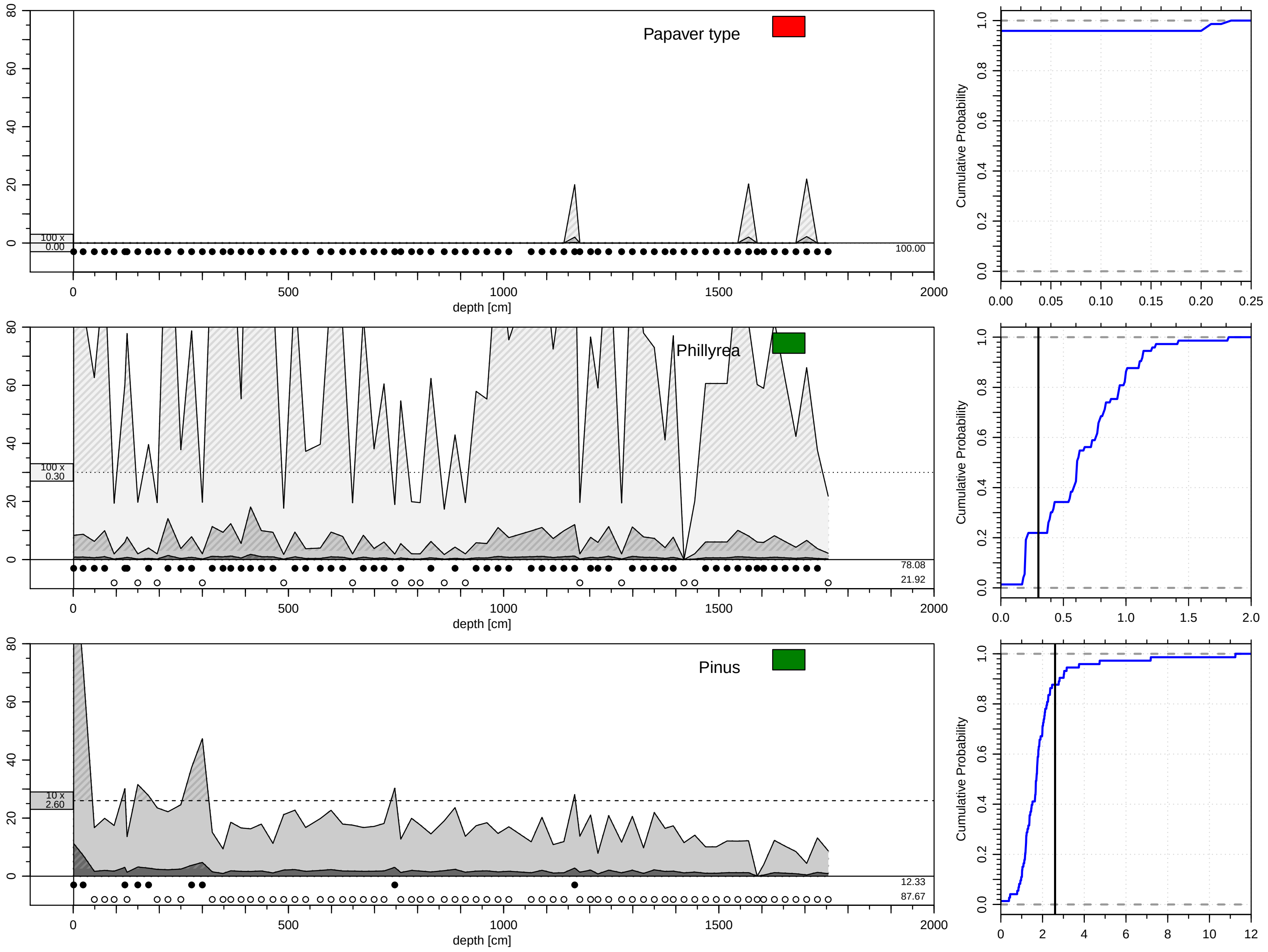


Figure B.9.: Lake Kinneret, definitions of the thresholds (part 9).

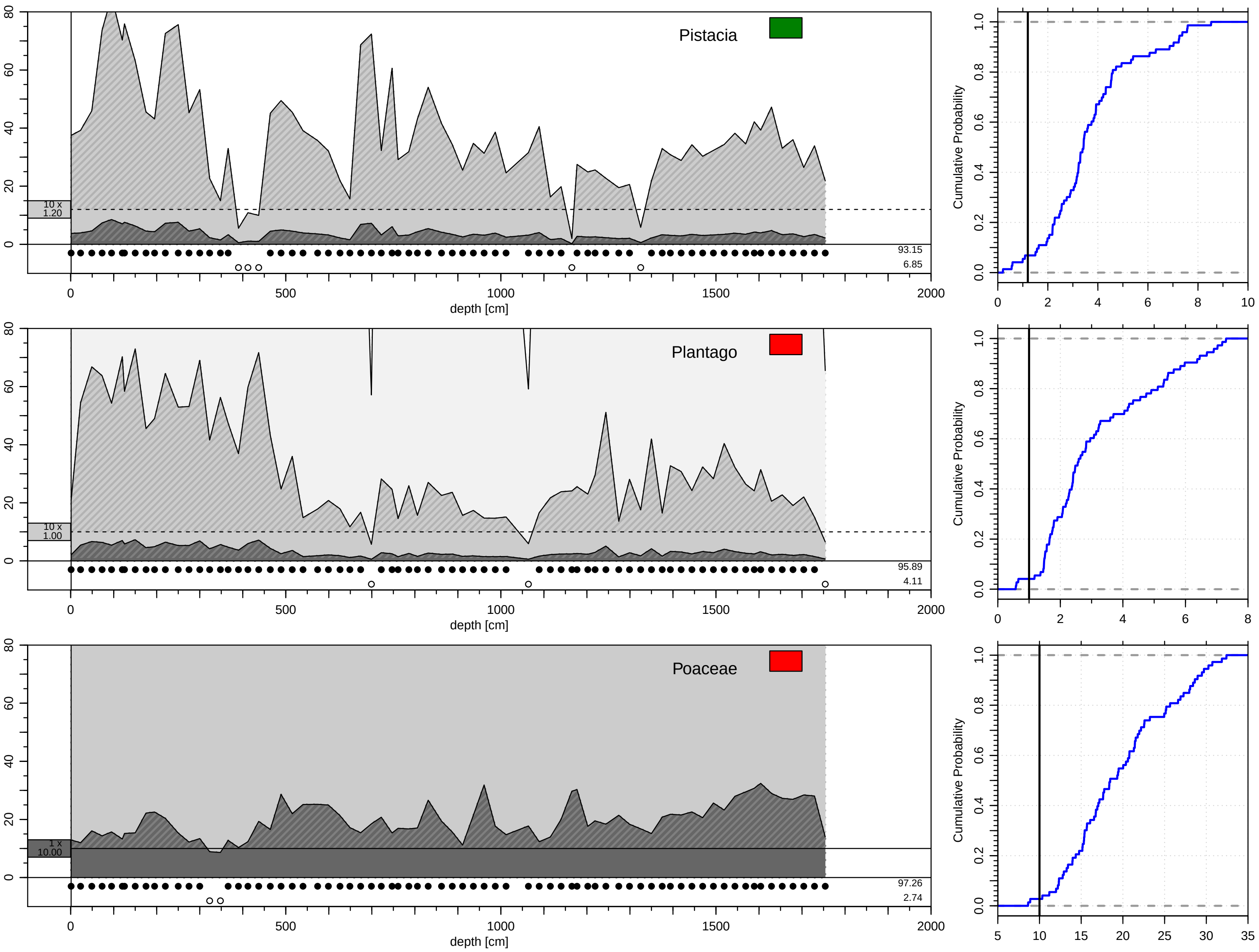


Figure B.10.: Lake Kinneret, definitions of the thresholds (part 10).

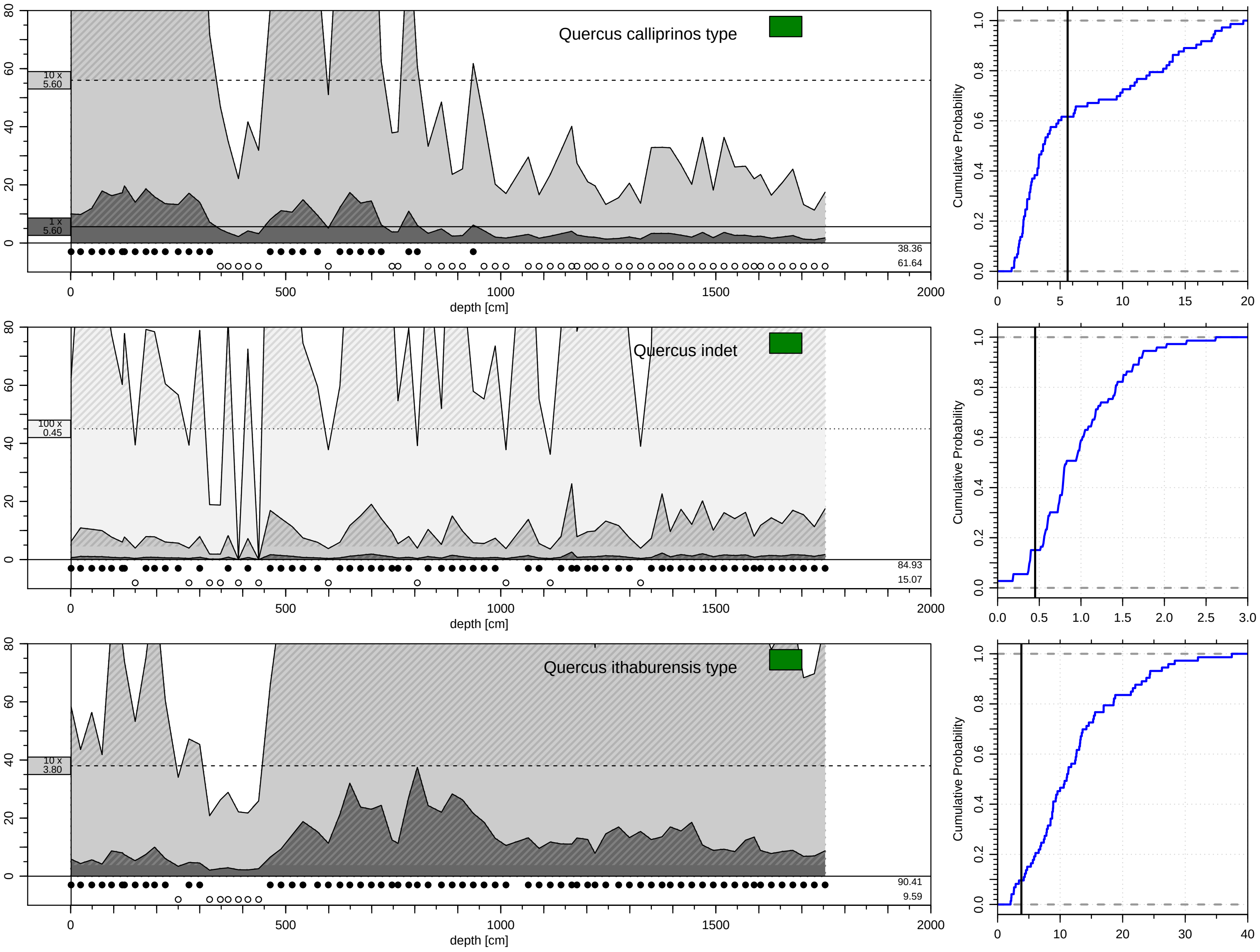


Figure B.11.: Lake Kinneret, definitions of the thresholds (part 11).

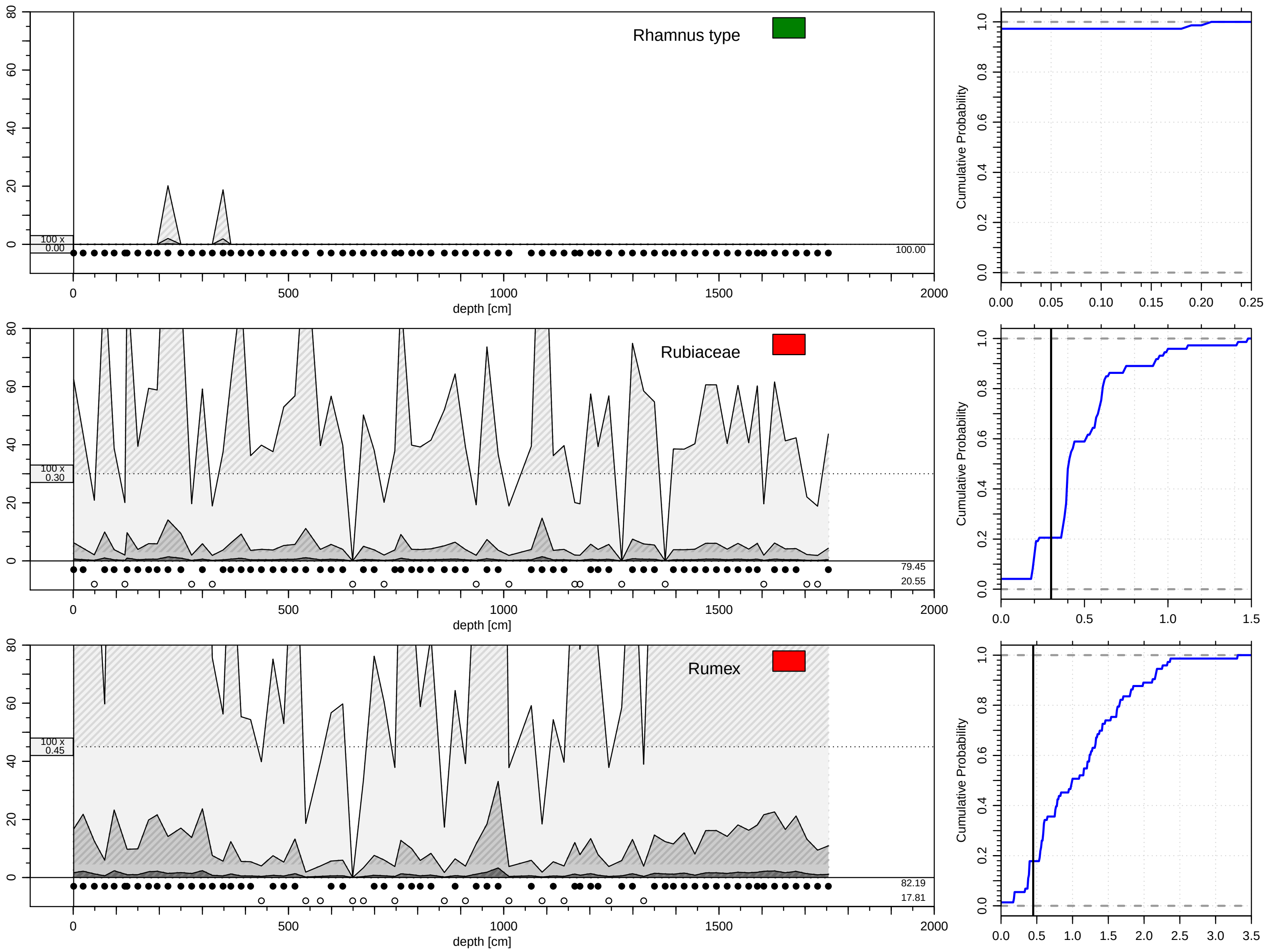


Figure B.12.: Lake Kinneret, definitions of the thresholds (part 12).

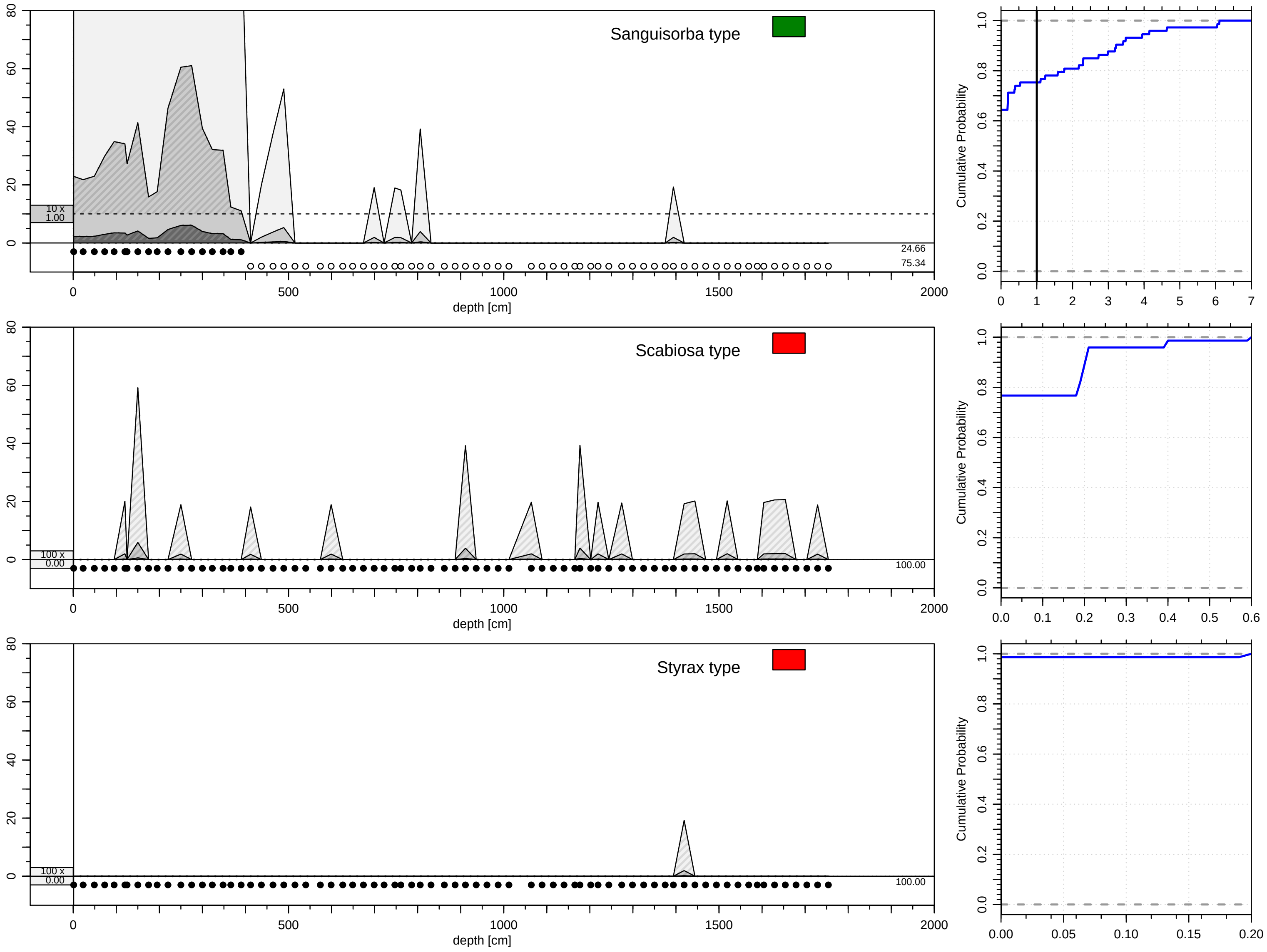


Figure B.13.: Lake Kinneret, definitions of the thresholds (part 13).

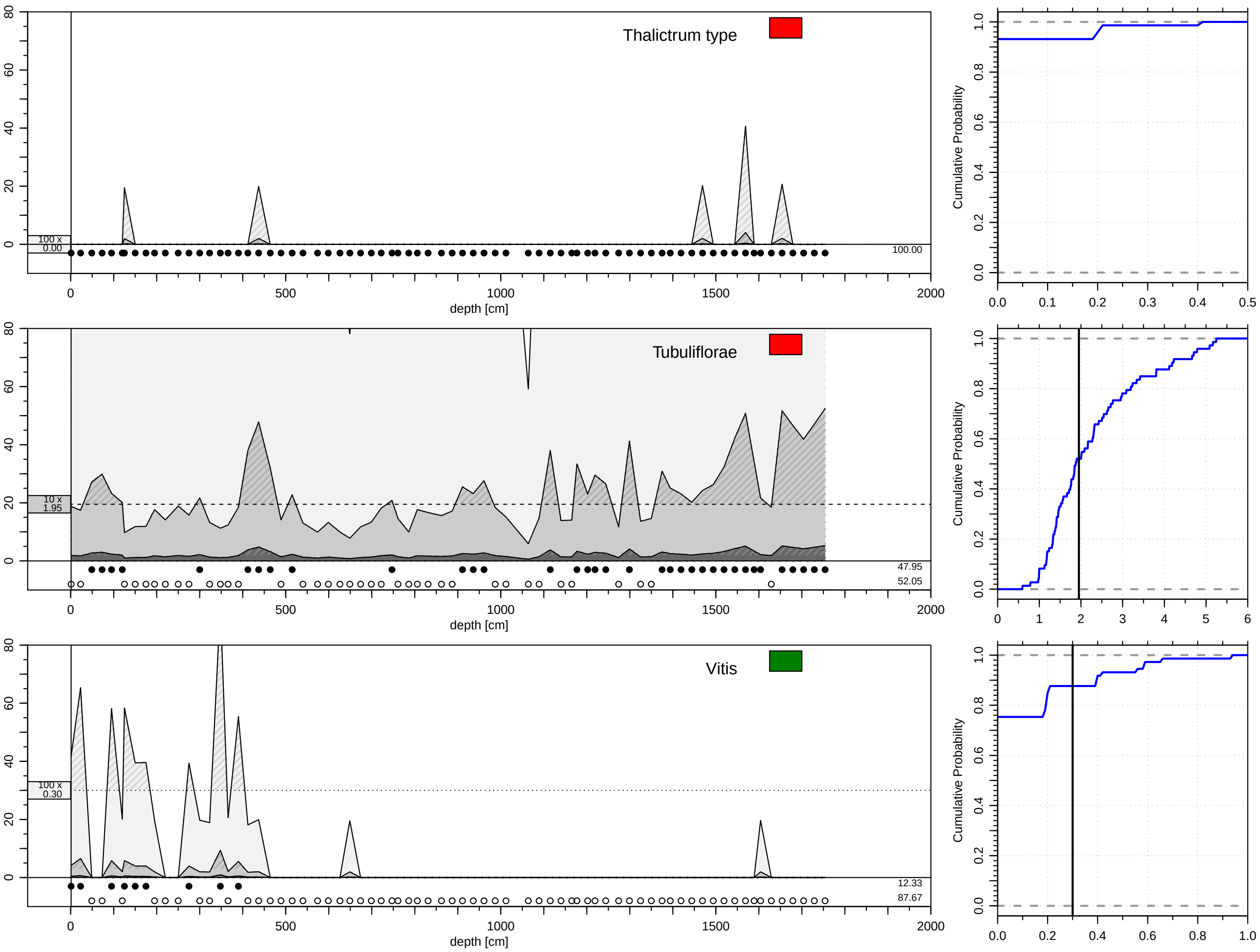


Figure B.14.: Lake Kinneret, definitions of the thresholds (part 14).

B.2. Biome Likelihood/Transfer Function $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$: Additional Figures

This chapter presents additional figures for the estimation of the biome likelihood - biome transfer function $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$ as described in chapter 4.2.3. Firstly, the biome areas for biome setup 3 and 4 are presented, directly followed by the figures for the verification of the biome likelihood for a total of three biomes ($BSS_{B_l}^{RF}$ and $BSS_{B_l}^{ZV}$). Thereby Figure B.17 is associated with Figure 4.2 and Figure B.18 with Figure 4.3.

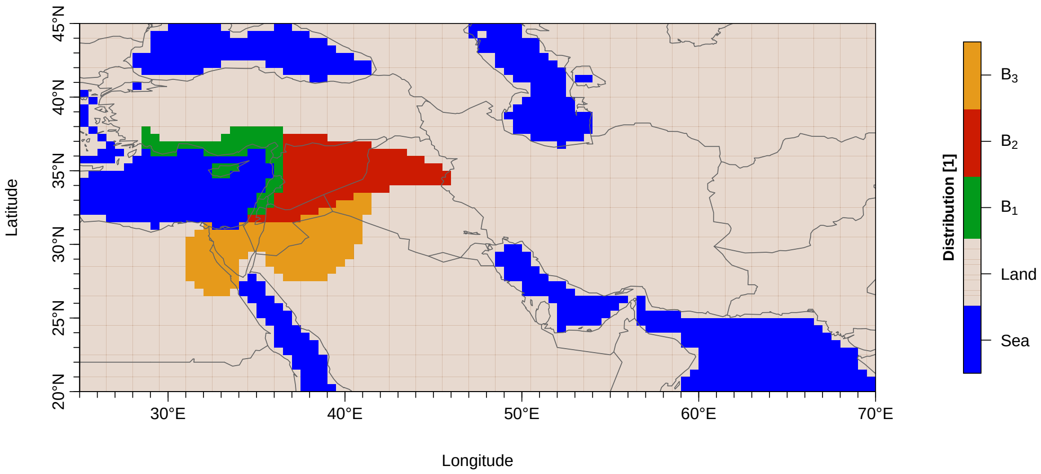


Figure B.15.: Definition of the biome areas denoted by biome setup = 3 bases on Fig. 2.11b. The Mediterranean biome area B_1 is extended. In detail it is defined as $B_1 := B_1 + B_{1-a} + B_{1-b}$.

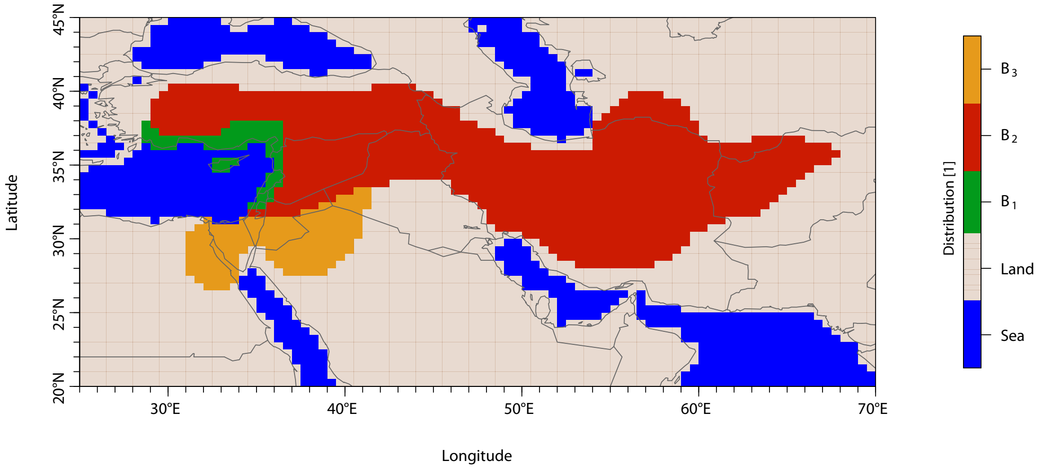
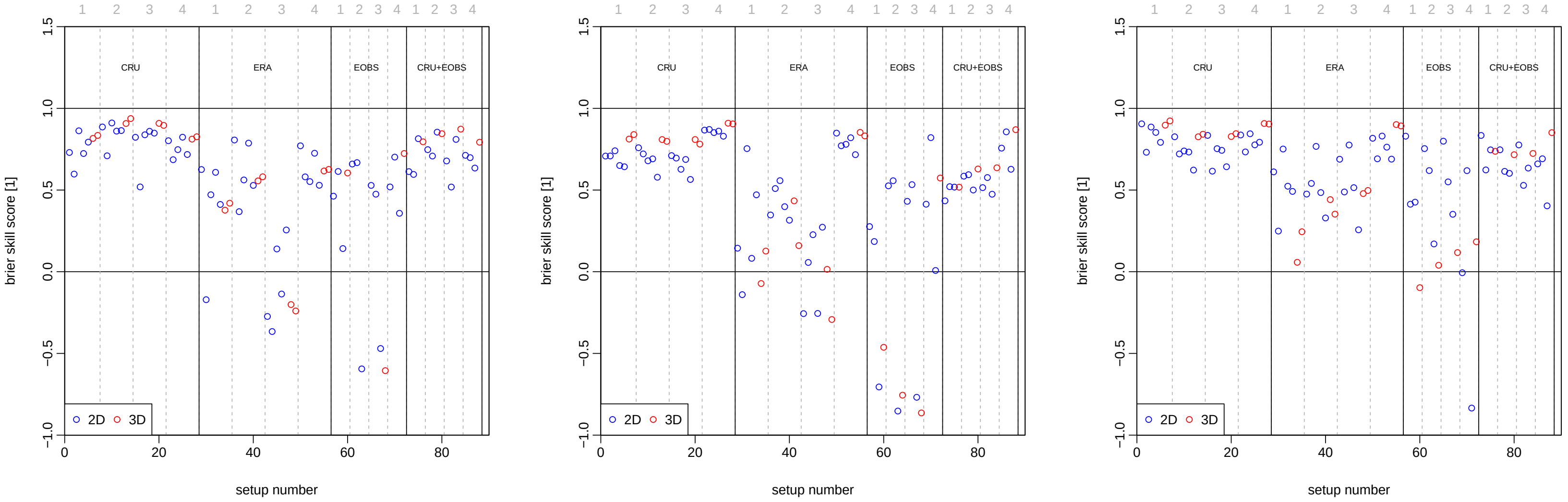


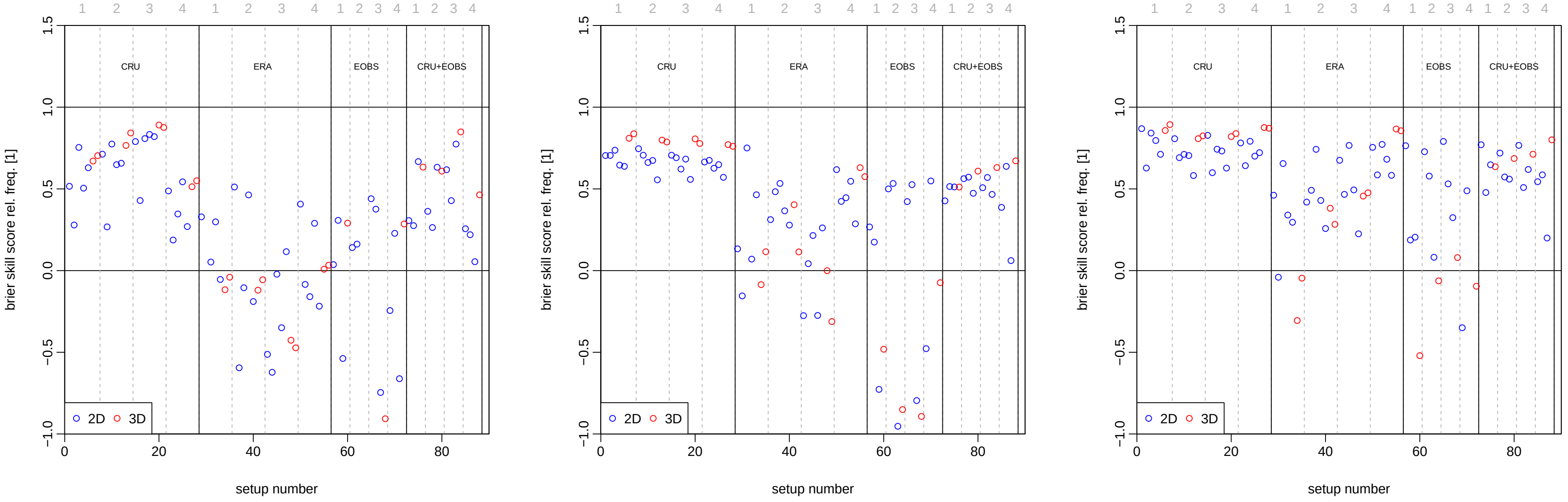
Figure B.16.: Definition of the biome areas denoted by biome setup = 4 bases on Fig. 2.11b. The Mediterranean biome area B_1 and the Irano-Turanian biome area B_2 are extended. In detail these biome areas are defined as $B_1 := B_1 + B_{1-a} + B_{1-b}$ and $B_2 := B_2 + B_{2-a}$.



(a) $BSS^{ZV}_{B_1}$ for a total of three biomes

(b) $BSS^{ZV}_{B_2}$ for a total of three biomes

(c) $BSS^{ZV}_{B_3}$ for a total of three biomes

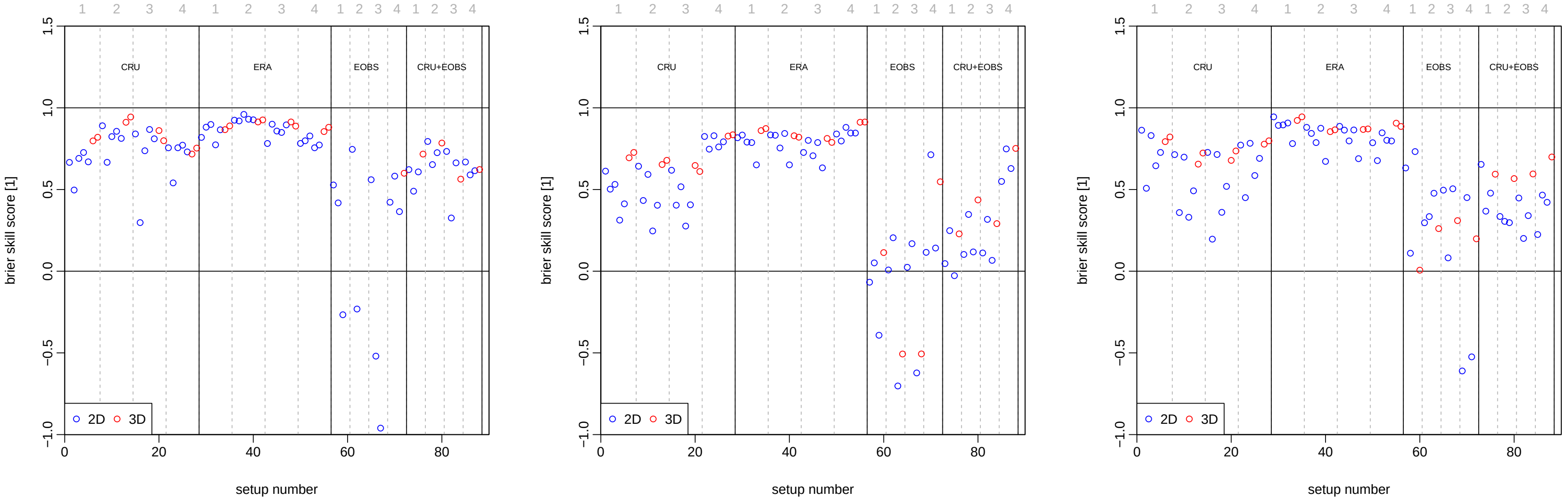


(d) $BSS^{RF}_{B_1}$ for a total of three biomes

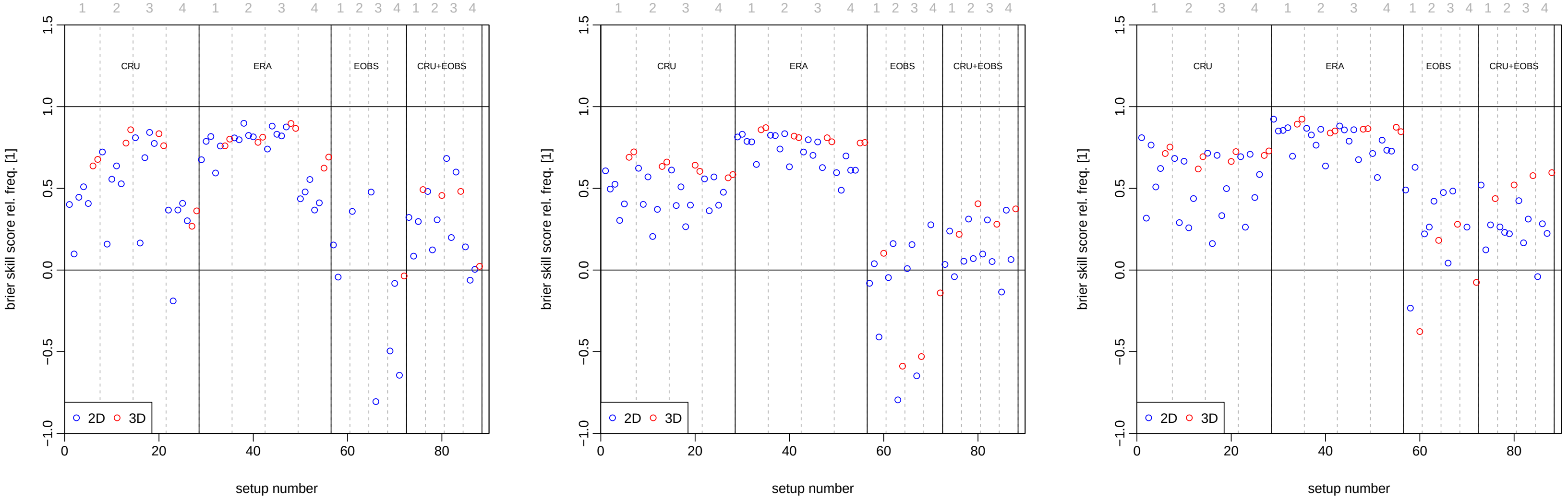
(e) $BSS^{RF}_{B_2}$ for a total of three biomes

(f) $BSS^{RF}_{B_3}$ for a total of three biomes

Figure B.17.: Verification of the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_l = l|\vec{c}\right)$ with Brier skill score $BSS^{ZV}_{B_l}$ and $BSS^{RF}_{B_l}$ for a total of three biomes. The parameters which determine the biome likelihood are estimated as given by the setup number shown on the x-axis. The setup numbers are defined in Table B.4. The climate input data for the verification is always the CRU TS 3.1 dataset. The dotted grey vertical lines separate the different "*biome setups*" as specified on top of the Figures (grey numbers). The black vertical lines separate the different climate databases as input data in the estimation process.



(a) $BSS^{ZV}_{B_1}$ for a total of three biomes **(b)** $BSS^{ZV}_{B_2}$ for a total of three biomes **(c)** $BSS^{ZV}_{B_3}$ for a total of three biomes



(d) $BSS^{RF}_{B_1}$ for a total of three biomes **(e)** $BSS^{RF}_{B_2}$ for a total of three biomes **(f)** $BSS^{RF}_{B_3}$ for a total of three biomes

Figure B.18.: Verification of the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_l = l|\vec{c}\right)$ with Brier skill score $BSS^{ZV}_{B_l}$ and $BSS^{RF}_{B_l}$ for a total of three biomes. The parameters which determine the biome likelihood are estimated as given by the setup number shown on the x-axis. The setup numbers are defined in Table B.5. The climate input data for the verification is always the ERA-Interim dataset for the 850 hPa temperatures. The dotted grey vertical lines separate the different "*biome setups*" as specified on top of the Figures (grey numbers). The black vertical lines separate the different climate databases as input data in the estimation process.

B.3. Verification Data Tables for the Biome Likelihood/Transfer Function $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$

As mentioned in chapter 4.2.3 different possible setups for the estimation of the biome likelihood were verificated. This appendix chapter presents the data tables with the results for the Brier skill scores defined in chapter 4.2.3 (p. 58). The setup numbers (s.no.) are defined in this tables by the column *predictors*, *KlimDataSet* and *biome setup*. More detailed the column

- *predictors* define the number of dimensions of the considered climate state vector as defined in equation 4.11 as well as the combinations of the climate variables (T_{DJF}, T_{JJA}, T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN} or CWD_{ANN}) for the estimation of the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}|\overrightarrow{c}_0\right)$.
- *KlimDataSet* breaks down which kind of climate data set is applied for the estimation of the biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}|\overrightarrow{c}_0\right)$. Thereby
  - 1 is equivalent to CRU TS 3.1 for T_{DJF}, T_{JJA}, P_{ANN} and CWD_{ANN},
  - 2 is equivalent to CRU TS 3.1 for P_{ANN} and CWD_{ANN} and ERA-Interim for T_{DJF}^{850hpa} and T_{JJA}^{850hpa},
  - 3 is equivalent to E-OBS for T_{DJF}, T_{JJA} and P_{ANN} and
  - 4 is equivalent to mixture of CRU TS 3.1 and E-OBS for T_{DJF}, T_{JJA} and P_{ANN}.
- *biome setup* breaks down which definition of the biome areas is applied for the estimation of the biome likelihood. Thereby
  - *biome setup 1* defines the biome areas B_1, B_2 and B_3 as presented in Figure 2.11a,
  - *biome setup 2* defines the biome areas B_1, B_2 and B_3 as presented in Figure 2.11b,
  - *biome setup 3* defines the biome areas B_1, B_2 and B_3 as presented in Figure B.15 and
  - *biome setup 4* defines the biome areas B_1, B_2 and B_3 as presented in Figure B.16.

The other columns in these tables show the result for the Brier score for each regarded biome type B_l. In detail

- RF_{B_l} is the relative frequency of each biome B_l (s. eq. 4.9),
- BS_{B_l} is the Brier score for each biome (s. eq. 4.12),
- $BSS_{B_l}^{RF}$ Brier skill score calculated with the reference score $BS_{ref}^{RF_{B_l}}$ (s. eq. 4.13a),
- $BSS_{B_l}^{ZV}$ Brier skill score calculated with the reference score BS_{ref}^{ZV} (s. eq. 4.13b),
- N_{obs} number of grid points with presence information of any biome type B_l (s. eq. 4.9).

The column $clim. - rec.$ in Table B.2 and B.3 indicates if there is available a palaeoclimate and where it can be found. If the number of the figure is not highlighted the referred number is assigned to Lake Kinneret marginal pdfs. A red highlighted number is assigned to Birkat Ram marginal probabilities and a orange highlighted number to Ein Gedi marginal pdfs.

It is important to emphasize that every setup number in Table B.2, B.3, B.4 and B.5 defines the same setup aside from the total number of biomes. Summarized Table B.2 shows the database for Figure 4.2, Table B.3 for Figure 4.3, Table B.4 for Figure B.17 and Table B.5 for Figure B.18.

Table B.2.: Table with the data for the verification of biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}\middle|\vec{c}_0\right)$ verificated with the CRU TS 3.1 climate dataset for a total of two biomes. A detailed description of the columns can be found at the beginning of chapter B.3.

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. – rec.*
1	23.15%	0.095	0.469	0.622	76.85%	0.095	0.469	0.622	108	T_{DJF}, P_{ANN}	1	1	4.11, C.1b, C.1d
2		0.074	0.583	0.704		0.074	0.583	0.704		T_{DJF}, T_{JJA}	1		
3		0.052	0.709	0.793		0.052	0.709	0.793		T_{JJA}, P_{ANN}	1		
4		0.093	0.477	0.628		0.093	0.477	0.628		T_{DJF}, CWD_{ANN}	1		
5		0.058	0.671	0.766		0.058	0.671	0.766		T_{JJA}, CWD_{ANN}	1		
6		0.050	0.716	0.798		0.050	0.716	0.798		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1		
7		0.046	0.744	0.817		0.046	0.744	0.817		$T_{DJF}, T_{JJA}, P_{ANN}$	1		B.20, 4.14a, 4.15, 4.14b
8	10.76%	0.029	0.694	0.883	89.24%	0.029	0.694	0.883	158	T_{DJF}, P_{ANN}	1	2	
9		0.060	0.372	0.759		0.060	0.372	0.759		T_{DJF}, T_{JJA}	1		
10		0.023	0.761	0.908		0.023	0.761	0.908		T_{JJA}, P_{ANN}	1		
11		0.027	0.720	0.892		0.027	0.720	0.892		T_{DJF}, CWD_{ANN}	1		
12		0.023	0.761	0.908		0.023	0.761	0.908		T_{JJA}, CWD_{ANN}	1		
13		0.023	0.763	0.909		0.023	0.763	0.909		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1		
14		0.017	0.820	0.931		0.017	0.820	0.931		$T_{DJF}, T_{JJA}, P_{ANN}$	1		B.21

continued on next page…

. . . continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. – rec.*
15	30.20%	0.052	0.753	0.792	69.80%	0.052	0.753	0.792	202	T_{DJF}, P_{ANN}	1	3	4.8a, 4.8c
16		0.073	0.652	0.706		0.073	0.652	0.706		T_{DJF}, T_{JJA}	1		
17		0.048	0.771	0.807		0.048	0.771	0.807		T_{JJA}, P_{ANN}	1		
18		0.031	0.852	0.875		0.031	0.852	0.875		T_{DJF}, CWD_{ANN}	1		4.8b, 4.8d
19		0.036	0.831	0.857		0.036	0.831	0.857		T_{JJA}, CWD_{ANN}	1		
20		0.024	0.887	0.905		0.024	0.887	0.905		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1		4.4d, 4.16d, 5.3d
21		0.029	0.862	0.884		0.029	0.862	0.884		$T_{DJF}, T_{JJA}, P_{ANN}$	1		4.4a-4.4c, B.23, B.24, 4.16a-4.16c, , 5.3a-5.3c
22	6.06%	0.029	0.483	0.882	93.94%	0.029	0.483	0.882	1007	T_{DJF}, P_{ANN}	1	4	
23		0.032	0.430	0.870		0.032	0.430	0.870		T_{DJF}, T_{JJA}	1		
24		0.038	0.341	0.850		0.038	0.341	0.850		T_{JJA}, P_{ANN}	1		
25		0.024	0.575	0.903		0.024	0.575	0.903		T_{DJF}, CWD_{ANN}	1		
26		0.040	0.298	0.840		0.040	0.298	0.840		T_{JJA}, CWD_{ANN}	1		
27		0.027	0.519	0.891		0.027	0.519	0.891		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1		
28		0.026	0.550	0.898		0.026	0.550	0.898		$T_{DJF}, T_{JJA}, P_{ANN}$	1		B.22

continued on next page. . .

...continuation

$s.no.$	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	$predictors$	$KlimDataSet$	$biome\ setup$	$clim. - rec.$
29	23.15%	0.142	0.200	0.430	76.85%	0.142	0.200	0.430	108	$T^{850hpa}_{DJF}, P_{ANN}$	2	1	
30		0.474	-1.666	-0.897		0.474	-1.666	-0.897		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}$	2		
31		0.186	-0.048	0.254		0.186	-0.048	0.254		$T^{850hpa}_{JJA}, P_{ANN}$	2		
32		0.130	0.267	0.478		0.130	0.267	0.478		$T^{850hpa}_{DJF}, CWD_{ANN}$	2		
33		0.195	-0.095	0.221		0.195	-0.095	0.221		$T^{850hpa}_{JJA}, CWD_{ANN}$	2		
34		0.237	-0.334	0.050		0.237	-0.334	0.050		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}$	2		
35		0.224	-0.261	0.103		0.224	-0.261	0.103		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}$	2		
36	10.76%	0.050	0.482	0.801	89.24%	0.050	0.482	0.801	158	$T^{850hpa}_{DJF}, P_{ANN}$	2	2	
37		0.146	-0.521	0.416		0.146	-0.521	0.416		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}$	2		
38		0.112	-0.171	0.550		0.112	-0.171	0.550		$T^{850hpa}_{JJA}, P_{ANN}$	2		
39		0.042	0.558	0.830		0.042	0.558	0.830		$T^{850hpa}_{DJF}, CWD_{ANN}$	2		
40		0.119	-0.236	0.525		0.119	-0.236	0.525		$T^{850hpa}_{JJA}, CWD_{ANN}$	2		
41		0.108	-0.120	0.570		0.108	-0.120	0.570		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}$	2		
42		0.108	-0.120	0.570		0.108	-0.120	0.570		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}$	2		

continued on next page...

…continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. − rec.*
43	30.20%	0.400	-0.897	-0.600	69.80%	0.400	-0.897	-0.600	202	T^{850hpa}_{DJF}, P_{ANN}	2	3	
44		0.398	-0.890	-0.593		0.398	-0.890	-0.593		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2		
45		0.256	-0.216	-0.025		0.256	-0.216	-0.025		T^{850hpa}_{JJA}, P_{ANN}	2		
46		0.329	-0.563	-0.318		0.329	-0.563	-0.318		T^{850hpa}_{DJF}, CWD_{ANN}	2		
47		0.219	-0.038	0.125		0.219	-0.038	0.125		T^{850hpa}_{JJA}, CWD_{ANN}	2		
48		0.352	-0.671	-0.409		0.352	-0.671	-0.409		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2		
49		0.382	-0.814	-0.529		0.382	-0.814	-0.529		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2		
50	6.06%	0.034	0.402	0.864	93.94%	0.034	0.402	0.864	1007	T^{850hpa}_{DJF}, P_{ANN}	2	4	
51		0.061	-0.072	0.756		0.061	-0.072	0.756		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2		
52		0.067	-0.170	0.734		0.067	-0.170	0.734		T^{850hpa}_{JJA}, P_{ANN}	2		
53		0.039	0.318	0.845		0.039	0.318	0.845		T^{850hpa}_{DJF}, CWD_{ANN}	2		
54		0.069	-0.219	0.723		0.069	-0.219	0.723		T^{850hpa}_{JJA}, CWD_{ANN}	2		
55		0.056	0.017	0.776		0.056	0.017	0.776		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2		5.4d
56		0.056	0.025	0.778		0.056	0.025	0.778		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2		4.6a, 4.7, 4.6b, B.19, 5.4a - 5.4c
57	23.15%	0.115	0.351	0.538	76.85%	0.115	0.351	0.538	108	T_{DJF}, P_{ANN}	3	1	
58		0.090	0.493	0.639		0.090	0.493	0.639		T_{DJF}, T_{JJA}	3		
59		0.402	-1.262	-0.610		0.402	-1.262	-0.610		T_{JJA}, P_{ANN}	3		
60		0.137	0.228	0.451		0.137	0.228	0.451		T_{DJF}, T_{JJA}, P_{ANN}	3		

continued on next page…

...continuation

$s.no.$	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	$predictors$	$KlimDataSet$	$biome\ setup$	$clim. - rec.$
61	10.76%	0.089	0.075	0.645	89.24%	0.089	0.075	0.645	158	T_{DJF}, P_{ANN}	3	2	
62		0.073	0.243	0.709		0.073	0.243	0.709		T_{DJF}, T_{JJA}	3		
63		0.541	-4.639	-1.166		0.541	-4.639	-1.166		T_{JJA}, P_{ANN}	3		
64		0.422	-3.398	-0.689		0.422	-3.398	-0.689		$T_{DJF}, T_{JJA}, P_{ANN}$	3		
65	30.20%	0.125	0.407	0.500	69.80%	0.125	0.407	0.500	202	T_{DJF}, P_{ANN}	3	3	
66		0.111	0.475	0.557		0.111	0.475	0.557		T_{DJF}, T_{JJA}	3		
67		0.441	-1.092	-0.764		0.441	-1.092	-0.764		T_{JJA}, P_{ANN}	3		
68		0.344	-0.631	-0.375		0.344	-0.631	-0.375		$T_{DJF}, T_{JJA}, P_{ANN}$	3		
69	6.06%	0.083	-0.451	0.670	93.94%	0.083	-0.451	0.670	1007	T_{DJF}, P_{ANN}	3	4	
70		0.034	0.411	0.866		0.034	0.411	0.866		T_{DJF}, T_{JJA}	3		
71		0.097	-0.697	0.614		0.097	-0.697	0.614		T_{JJA}, P_{ANN}	3		
72		0.054	0.058	0.786		0.054	0.058	0.786		$T_{DJF}, T_{JJA}, P_{ANN}$	3		
73	23.15%	0.122	0.314	0.512	76.85%	0.122	0.314	0.512	108	T_{DJF}, P_{ANN}	4	1	
74		0.086	0.518	0.657		0.086	0.518	0.657		T_{DJF}, T_{JJA}	4		
75		0.069	0.613	0.724		0.069	0.613	0.724		T_{JJA}, P_{ANN}	4		
76		0.086	0.519	0.657		0.086	0.519	0.657		$T_{DJF}, T_{JJA}, P_{ANN}$	4		

continued on next page...

…continuation

$s.no.$	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	$predictors$	$KlimDataSet$	$biome\ setup$	$clim.-rec.$
77	10.76%	0.064	0.334	0.744	89.24%	0.064	0.334	0.744	158	T_{DJF}, P_{ANN}	4	2	
78		0.057	0.402	0.771		0.057	0.402	0.771		T_{DJF}, T_{JJA}	4		
79		0.035	0.630	0.858		0.035	0.630	0.858		T_{JJA}, P_{ANN}	4		
80		0.046	0.526	0.818		0.046	0.526	0.818		$T_{DJF}, T_{JJA}, P_{ANN}$	4		
81	30.20%	0.084	0.603	0.665	69.80%	0.084	0.603	0.665	202	T_{DJF}, P_{ANN}	4	3	
82		0.083	0.605	0.667		0.083	0.605	0.667		T_{DJF}, T_{JJA}	4		
83		0.061	0.710	0.755		0.061	0.710	0.755		T_{JJA}, P_{ANN}	4		
84		0.040	0.812	0.841		0.040	0.812	0.841		$T_{DJF}, T_{JJA}, P_{ANN}$	4		6.3a, 6.3b, E.2, E.3, E.4
85	6.06%	0.053	0.066	0.787	93.94%	0.053	0.066	0.787	1007	T_{DJF}, P_{ANN}	4	4	
86		0.031	0.449	0.874		0.031	0.449	0.874		T_{DJF}, T_{JJA}	4		
87		0.062	-0.090	0.752		0.062	-0.090	0.752		T_{JJA}, P_{ANN}	4		
88		0.040	0.298	0.840		0.040	0.298	0.840		$T_{DJF}, T_{JJA}, P_{ANN}$	4		

Table B.3.: Table with the data for the verification of biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$ verificated with the ERA-Interim climate dataset for a total of two biomes. A detailed description of the columns can be found at the beginning of chapter B.3.

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. − rec.*
1	23.15%	0.116	0.350	0.538	76.85%	0.116	0.350	0.538	108	T_{DJF}, P_{ANN}	1	1	4.11, C.1b, C.1d
2		0.071	0.601	0.716		0.071	0.601	0.716		T_{DJF}, T_{JJA}	1		
3		0.118	0.339	0.529		0.118	0.339	0.529		T_{JJA}, P_{ANN}	1		
4		0.119	0.329	0.522		0.119	0.329	0.522		T_{DJF}, CWD_{ANN}	1		
5		0.117	0.343	0.533		0.117	0.343	0.533		T_{JJA}, CWD_{ANN}	1		
6		0.039	0.782	0.845		0.039	0.782	0.845		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1		
7		0.038	0.788	0.849		0.038	0.788	0.849		$T_{DJF}, T_{JJA}, P_{ANN}$	1		B.20, 4.14a, 4.15, 4.14b
8	10.76%	0.028	0.706	0.887	89.24%	0.028	0.706	0.887	158	T_{DJF}, P_{ANN}	1	2	
9		0.063	0.346	0.749		0.063	0.346	0.749		T_{DJF}, T_{JJA}	1		
10		0.045	0.527	0.819		0.045	0.527	0.819		T_{JJA}, P_{ANN}	1		
11		0.036	0.625	0.856		0.036	0.625	0.856		T_{DJF}, CWD_{ANN}	1		
12		0.041	0.577	0.837		0.041	0.577	0.837		T_{JJA}, CWD_{ANN}	1		
13		0.016	0.833	0.936		0.016	0.833	0.936		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1		
14		0.014	0.849	0.942		0.014	0.849	0.942		$T_{DJF}, T_{JJA}, P_{ANN}$	1		B.21

continued on next page…

. . . continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. – rec.*
15	30.20%	0.044	0.789	0.822	69.80%	0.044	0.789	0.822	202	T_{DJF}, P_{ANN}	1	3	4.8a, 4.8c
16		0.154	0.268	0.383		0.154	0.268	0.383		T_{DJF}, T_{JJA}	1		
17		0.078	0.629	0.687		0.078	0.629	0.687		T_{JJA}, P_{ANN}	1		
18		0.033	0.842	0.867		0.033	0.842	0.867		T_{DJF}, CWD_{ANN}	1		4.8b, 4.8d
19		0.047	0.778	0.813		0.047	0.778	0.813		T_{JJA}, CWD_{ANN}	1		
20		0.034	0.837	0.863		0.034	0.837	0.863		T_{DJF}, T_{JJA}, CWD_{ANN}	1		4.4d, 4.16d, 5.3d
21		0.055	0.740	0.781		0.055	0.740	0.781		T_{DJF}, T_{JJA}, P_{ANN}	1		4.4a-4.4c, B.23, B.24, 4.16a-4.16c, , 5.3a-5.3c
22	6.06%	0.036	0.361	0.855	93.94%	0.036	0.361	0.855	1007	T_{DJF}, P_{ANN}	1	4	
23		0.061	-0.075	0.755		0.061	-0.075	0.755		T_{DJF}, T_{JJA}	1		
24		0.036	0.362	0.855		0.036	0.362	0.855		T_{JJA}, P_{ANN}	1		
25		0.033	0.425	0.869		0.033	0.425	0.869		T_{DJF}, CWD_{ANN}	1		
26		0.038	0.331	0.848		0.038	0.331	0.848		T_{JJA}, CWD_{ANN}	1		
27		0.040	0.290	0.838		0.040	0.290	0.838		T_{DJF}, T_{JJA}, CWD_{ANN}	1		
28		0.037	0.356	0.853		0.037	0.356	0.853		T_{DJF}, T_{JJA}, P_{ANN}	1		B.22

continued on next page. . .

…continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. − rec.*
29	23.15%	0.065	0.637	0.742	76.85%	0.065	0.637	0.742	108	T^{850hpa}_{DJF}, P_{ANN}	2	1	
30		0.032	0.821	0.873		0.032	0.821	0.873		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2		
31		0.036	0.797	0.856		0.036	0.797	0.856		T^{850hpa}_{JJA}, P_{ANN}	2		
32		0.069	0.614	0.726		0.069	0.614	0.726		T^{850hpa}_{DJF}, CWD_{ANN}	2		
33		0.033	0.817	0.870		0.033	0.817	0.870		T^{850hpa}_{JJA}, CWD_{ANN}	2		
34		0.033	0.815	0.868		0.033	0.815	0.868		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2		
35		0.033	0.816	0.869		0.033	0.816	0.869		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2		
36	10.76%	0.020	0.791	0.920	89.24%	0.020	0.791	0.920	158	T^{850hpa}_{DJF}, P_{ANN}	2	2	
37		0.015	0.840	0.939		0.015	0.840	0.939		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2		
38		0.010	0.892	0.958		0.010	0.892	0.958		T^{850hpa}_{JJA}, P_{ANN}	2		
39		0.020	0.792	0.920		0.020	0.792	0.920		T^{850hpa}_{DJF}, CWD_{ANN}	2		
40		0.012	0.877	0.953		0.012	0.877	0.953		T^{850hpa}_{JJA}, CWD_{ANN}	2		
41		0.018	0.812	0.928		0.018	0.812	0.928		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2		
42		0.019	0.801	0.923		0.019	0.801	0.923		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2		

continued on next page…

…continuation

$s.no.$	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	$predictors$	$KlimDataSet$	$biome\ setup$	$clim.-rec.$
43	30.20%	0.066	0.687	0.736	69.80%	0.066	0.687	0.736	202	$T^{850hpa}_{DJF}, P_{ANN}$	2	3	
44		0.031	0.854	0.877		0.031	0.854	0.877		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}$	2		
45		0.040	0.811	0.841		0.040	0.811	0.841		$T^{850hpa}_{JJA}, P_{ANN}$	2		
46		0.042	0.801	0.833		0.042	0.801	0.833		$T^{850hpa}_{DJF}, CWD_{ANN}$	2		
47		0.026	0.878	0.897		0.026	0.878	0.897		$T^{850hpa}_{JJA}, CWD_{ANN}$	2		
48		0.026	0.879	0.898		0.026	0.879	0.898		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}$	2		
49		0.033	0.843	0.867		0.033	0.843	0.867		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}$	2		
50	6.06%	0.032	0.431	0.870	93.94%	0.032	0.431	0.870	1007	$T^{850hpa}_{DJF}, P_{ANN}$	2	4	
51		0.033	0.415	0.867		0.033	0.415	0.867		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}$	2		
52		0.026	0.551	0.898		0.026	0.551	0.898		$T^{850hpa}_{JJA}, P_{ANN}$	2		
53		0.034	0.395	0.862		0.034	0.395	0.862		$T^{850hpa}_{DJF}, CWD_{ANN}$	2		
54		0.033	0.422	0.868		0.033	0.422	0.868		$T^{850hpa}_{JJA}, CWD_{ANN}$	2		
55		0.022	0.607	0.910		0.022	0.607	0.910		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}$	2		5.4d
56		0.018	0.688	0.929		0.018	0.688	0.929		$T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}$	2		4.6a, 4.7, 4.6b, B.19, 5.4a - 5.4c
57	23.15%	0.131	0.266	0.477	76.85%	0.131	0.266	0.477	108	T_{DJF}, P_{ANN}	3	1	
58		0.092	0.483	0.632		0.092	0.483	0.632		T_{DJF}, T_{JJA}	3		
59		0.497	-1.793	-0.988		0.497	-1.793	-0.988		T_{JJA}, P_{ANN}	3		
60		0.308	-0.731	-0.232		0.308	-0.731	-0.232		$T_{DJF}, T_{JJA}, P_{ANN}$	3		

continued on next page…

...continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS_{B_1}^{RF}$	$BSS_{B_1}^{ZV}$	RF_{B_2}	BS_{B_2}	$BSS_{B_2}^{RF}$	$BSS_{B_2}^{ZV}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*	*clim. − rec.*
61	10.76%	0.112	-0.163	0.553	89.24%	0.112	-0.163	0.553	158	T_{DJF}, P_{ANN}	3	2	
62		0.241	-1.514	0.034		0.241	-1.514	0.034		T_{DJF}, T_{JJA}	3		
63		0.629	-5.547	-1.514		0.629	-5.547	-1.514		T_{JJA}, P_{ANN}	3		
64		0.608	-5.329	-1.431		0.608	-5.329	-1.431		$T_{DJF}, T_{JJA}, P_{ANN}$	3		
65	30.20%	0.163	0.228	0.349	69.80%	0.163	0.228	0.349	202	T_{DJF}, P_{ANN}	3	3	
66		0.267	-0.266	-0.068		0.267	-0.266	-0.068		T_{DJF}, T_{JJA}	3		
67		0.508	-1.409	-1.031		0.508	-1.409	-1.031		T_{JJA}, P_{ANN}	3		
68		0.470	-1.229	-0.879		0.470	-1.229	-0.879		$T_{DJF}, T_{JJA}, P_{ANN}$	3		
69	6.06%	0.100	-0.751	0.601	93.94%	0.100	-0.751	0.601	1007	T_{DJF}, P_{ANN}	3	4	
70		0.057	-0.004	0.771		0.057	-0.004	0.771		T_{DJF}, T_{JJA}	3		
71		0.092	-0.612	0.633		0.092	-0.612	0.633		T_{JJA}, P_{ANN}	3		
72		0.067	-0.183	0.731		0.067	-0.183	0.731		$T_{DJF}, T_{JJA}, P_{ANN}$	3		
73	23.15%	0.152	0.147	0.393	76.85%	0.152	0.147	0.393	108	T_{DJF}, P_{ANN}	4	1	
74		0.098	0.451	0.610		0.098	0.451	0.610		T_{DJF}, T_{JJA}	4		
75		0.158	0.114	0.370		0.158	0.114	0.370		T_{JJA}, P_{ANN}	4		
76		0.109	0.387	0.564		0.109	0.387	0.564		$T_{DJF}, T_{JJA}, P_{ANN}$	4		

continued on next page...

…continuation

$s.no.$	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	N_{obs}	$predictors$	$KlimDataSet$	$biome\ setup$	$clim.-rec.$
77	10.76%	0.092	0.044	0.633	89.24%	0.092	0.044	0.633	158	T_{DJF}, P_{ANN}	4	2	
78		0.066	0.313	0.736		0.066	0.313	0.736		T_{DJF}, T_{JJA}	4		
79		0.072	0.249	0.712		0.072	0.249	0.712		T_{JJA}, P_{ANN}	4		
80		0.063	0.341	0.747		0.063	0.341	0.747		$T_{DJF}, T_{JJA}, P_{ANN}$	4		
81	30.20%	0.104	0.508	0.585	69.80%	0.104	0.508	0.585	202	T_{DJF}, P_{ANN}	4	3	
82		0.157	0.254	0.371		0.157	0.254	0.371		T_{DJF}, T_{JJA}	4		
83		0.118	0.440	0.528		0.118	0.440	0.528		T_{JJA}, P_{ANN}	4		
84		0.133	0.371	0.470		0.133	0.371	0.470		$T_{DJF}, T_{JJA}, P_{ANN}$	4		6.3a, 6.3b, E.2, E.3, E.4
85	6.06%	0.062	-0.097	0.750	93.94%	0.062	-0.097	0.750	1007	T_{DJF}, P_{ANN}	4	4	
86		0.055	0.029	0.779		0.055	0.029	0.779		T_{DJF}, T_{JJA}	4		
87		0.063	-0.099	0.750		0.063	-0.099	0.750		T_{JJA}, P_{ANN}	4		
88		0.062	-0.094	0.751		0.062	-0.094	0.751		$T_{DJF}, T_{JJA}, P_{ANN}$	4		

Table B.4.: Table with the data for the verification of biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(b_{l_0}\middle|\vec{c}_0\right)$ verificated with the CRU TS 3.1 climate dataset for a total of three biomes. A detailed description of the columns can be found at the beginning of chapter B.3.

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
1	8.50%	0.038	0.516	0.730	28.23%	0.060	0.704	0.708	63.27%	0.031	0.869	0.905	294	T_{DJF}, P_{ANN}	1	1
2		0.056	0.279	0.598		0.060	0.705	0.709		0.087	0.628	0.731		T_{DJF}, T_{JJA}	1	
3		0.019	0.754	0.863		0.053	0.737	0.740		0.037	0.842	0.886		T_{JJA}, P_{ANN}	1	
4		0.039	0.505	0.724		0.072	0.646	0.650		0.047	0.796	0.853		T_{DJF}, CWD_{ANN}	1	
5		0.029	0.629	0.793		0.073	0.638	0.643		0.067	0.712	0.792		T_{JJA}, CWD_{ANN}	1	
6		0.026	0.670	0.816		0.039	0.810	0.812		0.033	0.857	0.897		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1	
7		0.023	0.704	0.835		0.033	0.838	0.840		0.025	0.894	0.923		$T_{DJF}, T_{JJA}, P_{ANN}$	1	
8	5.41%	0.015	0.712	0.886	44.90%	0.063	0.746	0.759	49.68%	0.048	0.807	0.826	314	T_{DJF}, P_{ANN}	1	2
9		0.038	0.268	0.710		0.073	0.706	0.721		0.077	0.691	0.721		T_{DJF}, T_{JJA}	1	
10		0.012	0.775	0.911		0.084	0.662	0.679		0.072	0.711	0.739		T_{JJA}, P_{ANN}	1	
11		0.018	0.648	0.861		0.081	0.674	0.691		0.074	0.704	0.733		T_{DJF}, CWD_{ANN}	1	
12		0.018	0.657	0.864		0.110	0.556	0.578		0.105	0.582	0.622		T_{JJA}, CWD_{ANN}	1	
13		0.012	0.766	0.907		0.050	0.799	0.809		0.048	0.807	0.826		$T_{DJF}, T_{JJA}, CWD_{ANN}$	1	
14		0.008	0.842	0.937		0.053	0.787	0.798		0.044	0.825	0.842		$T_{DJF}, T_{JJA}, P_{ANN}$	1	

continued on next page…

. . . continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
15	17.04%	0.030	0.790	0.823	39.39%	0.070	0.706	0.711	43.58%	0.042	0.828	0.835	358	T_{DJF}, P_{ANN}	1	3
16		0.081	0.429	0.519		0.074	0.691	0.696		0.099	0.599	0.616		T_{DJF}, T_{JJA}	1	
17		0.027	0.808	0.838		0.090	0.621	0.627		0.063	0.743	0.753		T_{JJA}, P_{ANN}	1	
18		0.024	0.833	0.860		0.076	0.682	0.687		0.066	0.732	0.743		T_{DJF}, CWD_{ANN}	1	
19		0.025	0.820	0.848		0.105	0.558	0.565		0.092	0.627	0.643		T_{JJA}, CWD_{ANN}	1	
10		0.015	0.891	0.908		0.046	0.806	0.809		0.044	0.820	0.828		T_{DJF}, T_{JJA}, CWD_{ANN}	1	
21		0.018	0.876	0.896		0.053	0.778	0.781		0.040	0.838	0.845		T_{DJF}, T_{JJA}, P_{ANN}	1	
22	5.25%	0.025	0.487	0.802	81.34%	0.051	0.665	0.867	13.41%	0.025	0.781	0.837	1163	T_{DJF}, P_{ANN}	1	4
23		0.040	0.187	0.686		0.049	0.675	0.871		0.042	0.642	0.733		T_{DJF}, T_{JJA}	1	
24		0.032	0.347	0.747		0.057	0.626	0.852		0.024	0.791	0.845		T_{JJA}, P_{ANN}	1	
25		0.023	0.543	0.823		0.053	0.649	0.860		0.035	0.700	0.777		T_{DJF}, CWD_{ANN}	1	
26		0.036	0.269	0.718		0.065	0.571	0.830		0.032	0.722	0.793		T_{JJA}, CWD_{ANN}	1	
27		0.024	0.513	0.812		0.035	0.771	0.909		0.014	0.876	0.907		T_{DJF}, T_{JJA}, CWD_{ANN}	1	
28		0.022	0.550	0.826		0.036	0.761	0.905		0.015	0.871	0.904		T_{DJF}, T_{JJA}, P_{ANN}	1	

continued on next page. . .

…continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS_{B_1}^{RF}$	$BSS_{B_1}^{ZV}$	RF_{B_2}	BS_{B_2}	$BSS_{B_2}^{RF}$	$BSS_{B_2}^{ZV}$	RF_{B_3}	BS_{B_3}	$BSS_{B_3}^{RF}$	$BSS_{B_3}^{ZV}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
29	8.50%	0.052	0.329	0.626	28.23%	0.176	0.133	0.144	63.27%	0.125	0.461	0.611	294	T_{DJF}^{850hpa}, P_{ANN}	2	1
30		0.163	-1.099	-0.171		0.234	-0.155	-0.140		0.242	-0.041	0.249		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}	2	
31		0.074	0.052	0.471		0.051	0.751	0.754		0.080	0.655	0.751		T_{JJA}^{850hpa}, P_{ANN}	2	
32		0.055	0.298	0.609		0.188	0.070	0.082		0.153	0.340	0.524		T_{DJF}^{850hpa}, CWD_{ANN}	2	
33		0.082	-0.054	0.412		0.109	0.464	0.471		0.164	0.296	0.492		T_{JJA}^{850hpa}, CWD_{ANN}	2	
34		0.087	-0.117	0.377		0.220	-0.086	-0.072		0.303	-0.305	0.058		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, CWD_{ANN}	2	
35		0.081	-0.040	0.420		0.179	0.115	0.127		0.243	-0.046	0.245		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}	2	
36	5.41%	0.025	0.512	0.806	44.90%	0.170	0.312	0.347	49.68%	0.145	0.419	0.475	314	T_{DJF}^{850hpa}, P_{ANN}	2	2
37		0.082	-0.595	0.368		0.128	0.483	0.509		0.127	0.491	0.541		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}	2	
38		0.057	-0.105	0.562		0.115	0.534	0.557		0.064	0.742	0.767		T_{JJA}^{850hpa}, P_{ANN}	2	
39		0.028	0.463	0.787		0.157	0.366	0.398		0.143	0.429	0.485		T_{DJF}^{850hpa}, CWD_{ANN}	2	
40		0.061	-0.189	0.528		0.179	0.278	0.316		0.186	0.257	0.329		T_{JJA}^{850hpa}, CWD_{ANN}	2	
41		0.057	-0.120	0.556		0.148	0.403	0.434		0.155	0.382	0.441		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, CWD_{ANN}	2	
42		0.054	-0.056	0.581		0.219	0.115	0.160		0.179	0.283	0.353		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}	2	

continued on next page…

. . . continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
43	17.04%	0.214	-0.513	-0.274	39.39%	0.305	-0.276	-0.256	43.58%	0.080	0.675	0.689	358	T^{850hpa}_{DJF}, P_{ANN}	2	3
44		0.229	-0.622	-0.366		0.229	0.042	0.057		0.131	0.467	0.488		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2	
45		0.144	-0.022	0.140		0.187	0.215	0.227		0.058	0.766	0.776		T^{850hpa}_{JJA}, P_{ANN}	2	
46		0.191	-0.350	-0.136		0.304	-0.275	-0.255		0.124	0.494	0.515		T^{850hpa}_{DJF}, CWD_{ANN}	2	
47		0.125	0.116	0.256		0.176	0.261	0.273		0.191	0.225	0.257		T^{850hpa}_{JJA}, CWD_{ANN}	2	
48		0.202	-0.426	-0.201		0.239	-0.001	0.015		0.134	0.456	0.478		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2	
49		0.208	-0.473	-0.240		0.313	-0.312	-0.292		0.129	0.476	0.497		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2	
50	5.25%	0.029	0.407	0.771	81.34%	0.058	0.618	0.848	13.41%	0.029	0.755	0.817	1163	T^{850hpa}_{DJF}, P_{ANN}	2	4
51		0.054	-0.085	0.581		0.087	0.424	0.771		0.048	0.586	0.691		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2	
52		0.058	-0.160	0.552		0.084	0.446	0.780		0.026	0.772	0.830		T^{850hpa}_{JJA}, P_{ANN}	2	
53		0.035	0.289	0.725		0.069	0.547	0.820		0.037	0.681	0.763		T^{850hpa}_{DJF}, CWD_{ANN}	2	
54		0.061	-0.218	0.529		0.108	0.286	0.717		0.048	0.583	0.689		T^{850hpa}_{JJA}, CWD_{ANN}	2	
55		0.049	0.008	0.617		0.056	0.630	0.853		0.015	0.867	0.901		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2	
56		0.048	0.033	0.626		0.065	0.575	0.831		0.017	0.856	0.893		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2	
57	8.50%	0.075	0.037	0.463	28.23%	0.148	0.267	0.276	63.27%	0.055	0.764	0.830	294	T_{DJF}, P_{ANN}	3	1
58		0.054	0.308	0.614		0.167	0.174	0.185		0.189	0.187	0.413		T_{DJF}, T_{JJA}	3	
59		0.120	-0.538	0.142		0.350	-0.727	-0.705		0.185	0.204	0.426		T_{JJA}, P_{ANN}	3	
60		0.055	0.291	0.604		0.300	-0.481	-0.462		0.353	-0.521	-0.097		T_{DJF}, T_{JJA}, P_{ANN}	3	

continued on next page. . .

. . . continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
61	5.41%	0.044	0.141	0.659	44.90%	0.124	0.500	0.526	49.68%	0.068	0.727	0.754	314	T_{DJF}, P_{ANN}	3	2
62		0.043	0.162	0.668		0.116	0.533	0.557		0.105	0.578	0.619		T_{DJF}, T_{JJA}	3	
63		0.206	-3.022	-0.594		0.483	-0.953	-0.853		0.230	0.081	0.170		T_{JJA}, P_{ANN}	3	
64		0.297	-4.792	-1.296		0.458	-0.851	-0.756		0.266	-0.063	0.040		T_{DJF}, T_{JJA}, P_{ANN}	3	
65	17.04%	0.079	0.440	0.528	39.39%	0.138	0.422	0.431	43.58%	0.052	0.790	0.799	358	T_{DJF}, P_{ANN}	3	3
66		0.088	0.376	0.475		0.113	0.526	0.533		0.115	0.531	0.550		T_{DJF}, T_{JJA}	3	
67		0.247	-0.746	-0.470		0.429	-0.795	-0.768		0.166	0.324	0.352		T_{JJA}, P_{ANN}	3	
68		0.270	-0.907	-0.605		0.452	-0.893	-0.864		0.226	0.080	0.117		T_{DJF}, T_{JJA}, P_{ANN}	3	
69	5.25%	0.062	-0.245	0.519	81.34%	0.224	-0.478	0.413	13.41%	0.157	-0.350	-0.006	1163	T_{DJF}, P_{ANN}	3	4
70		0.038	0.228	0.702		0.068	0.549	0.821		0.059	0.489	0.619		T_{DJF}, T_{JJA}	3	
71		0.083	-0.662	0.358		0.379	-1.499	0.008		0.286	-1.461	-0.834		T_{JJA}, P_{ANN}	3	
72		0.036	0.285	0.724		0.163	-0.074	0.574		0.127	-0.096	0.183		T_{DJF}, T_{JJA}, P_{ANN}	3	
73	8.50%	0.054	0.306	0.613	28.23%	0.116	0.427	0.434	63.27%	0.053	0.770	0.834	294	T_{DJF}, P_{ANN}	4	1
74		0.056	0.275	0.596		0.098	0.515	0.521		0.121	0.478	0.623		T_{DJF}, T_{JJA}	4	
75		0.026	0.667	0.814		0.099	0.512	0.518		0.082	0.648	0.746		T_{JJA}, P_{ANN}	4	
76		0.029	0.633	0.795		0.099	0.512	0.518		0.085	0.636	0.737		T_{DJF}, T_{JJA}, P_{ANN}	4	

continued on next page. . .

. . . continuation

s.no.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
77	5.41%	0.033	0.363	0.747	44.90%	0.108	0.563	0.585	49.68%	0.070	0.719	0.746	314	T_{DJF}, P_{ANN}	4	2
78		0.038	0.264	0.708		0.106	0.572	0.594		0.107	0.573	0.614		T_{DJF}, T_{JJA}	4	
79		0.019	0.632	0.854		0.130	0.473	0.500		0.110	0.560	0.602		T_{JJA}, P_{ANN}	4	
80		0.020	0.609	0.845		0.097	0.609	0.629		0.078	0.686	0.716		$T_{DJF}, T_{JJA}, P_{ANN}$	4	
81	17.04%	0.054	0.618	0.678	39.39%	0.118	0.507	0.515	43.58%	0.057	0.767	0.776	358	T_{DJF}, P_{ANN}	4	3
82		0.081	0.428	0.519		0.103	0.570	0.576		0.121	0.509	0.529		T_{DJF}, T_{JJA}	4	
83		0.032	0.774	0.810		0.127	0.466	0.475		0.094	0.619	0.634		T_{JJA}, P_{ANN}	4	
84		0.021	0.849	0.873		0.088	0.631	0.636		0.071	0.712	0.724		$T_{DJF}, T_{JJA}, P_{ANN}$	4	
85	5.25%	0.037	0.256	0.713	81.34%	0.093	0.387	0.757	13.41%	0.053	0.544	0.660	1163	T_{DJF}, P_{ANN}	4	4
86		0.039	0.220	0.698		0.055	0.638	0.856		0.048	0.586	0.691		T_{DJF}, T_{JJA}	4	
87		0.047	0.054	0.635		0.142	0.061	0.627		0.093	0.199	0.403		T_{JJA}, P_{ANN}	4	
88		0.027	0.464	0.793		0.050	0.672	0.870		0.023	0.801	0.851		$T_{DJF}, T_{JJA}, P_{ANN}$	4	

Table B.5.: Table with the data for the verification of biome likelihood $\mathbb{P}_{\mathcal{B}_l|\vec{C}}\left(\mathscr{b}_{l_0}\middle|\vec{c}_0\right)$ verificated with the ERA-Interim climate dataset for a total of three biomes. A detailed description of the columns can be found at the beginning of chapter B.3.

No.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
1	8.50%	0.047	0.401	0.666	28.23%	0.080	0.607	0.612	63.27%	0.044	0.810	0.863	294	T_{DJF}, P_{ANN}	1	1
2		0.070	0.098	0.497		0.102	0.496	0.502		0.159	0.318	0.508		T_{DJF}, T_{JJA}	1	
3		0.043	0.446	0.691		0.096	0.525	0.531		0.055	0.765	0.830		T_{JJA}, P_{ANN}	1	
4		0.038	0.510	0.726		0.141	0.304	0.313		0.114	0.509	0.645		T_{DJF}, CWD_{ANN}	1	
5		0.046	0.407	0.669		0.121	0.405	0.412		0.088	0.622	0.727		T_{JJA}, CWD_{ANN}	1	
6		0.028	0.637	0.798		0.063	0.690	0.694		0.067	0.713	0.793		T_{DJF}, T_{JJA}, CWD_{ANN}	1	
7		0.025	0.677	0.820		0.056	0.723	0.727		0.057	0.753	0.821		T_{DJF}, T_{JJA}, P_{ANN}	1	
8	5.41%	0.014	0.723	0.890	44.90%	0.093	0.623	0.642	49.68%	0.079	0.683	0.714	314	T_{DJF}, P_{ANN}	1	2
9		0.043	0.159	0.666		0.148	0.402	0.433		0.177	0.290	0.359		T_{DJF}, T_{JJA}	1	
10		0.023	0.556	0.824		0.106	0.570	0.592		0.084	0.666	0.698		T_{JJA}, P_{ANN}	1	
11		0.019	0.637	0.856		0.197	0.206	0.246		0.185	0.259	0.330		T_{DJF}, CWD_{ANN}	1	
12		0.024	0.528	0.813		0.156	0.371	0.404		0.141	0.438	0.492		T_{JJA}, CWD_{ANN}	1	
13		0.011	0.777	0.912		0.091	0.634	0.653		0.095	0.619	0.655		T_{DJF}, T_{JJA}, CWD_{ANN}	1	
14		0.007	0.859	0.944		0.084	0.661	0.679		0.077	0.694	0.723		T_{DJF}, T_{JJA}, P_{ANN}	1	

continued on next page...

...continuation

No.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
15	17.04%	0.027	0.810	0.840	39.39%	0.093	0.612	0.618	43.58%	0.070	0.715	0.727	358	T_{DJF}, P_{ANN}	1	3
16		0.118	0.166	0.297		0.144	0.395	0.404		0.206	0.162	0.196		T_{DJF}, T_{JJA}	1	
17		0.044	0.688	0.737		0.117	0.509	0.517		0.073	0.702	0.714		T_{JJA}, P_{ANN}	1	
18		0.022	0.843	0.867		0.175	0.265	0.276		0.164	0.333	0.360		T_{DJF}, CWD_{ANN}	1	
19		0.032	0.775	0.810		0.144	0.397	0.406		0.123	0.499	0.519		T_{JJA}, CWD_{ANN}	1	
20		0.023	0.834	0.861		0.086	0.642	0.647		0.082	0.665	0.678		T_{DJF}, T_{JJA}, CWD_{ANN}	1	
21		0.034	0.761	0.799		0.094	0.604	0.610		0.068	0.725	0.736		T_{DJF}, T_{JJA}, P_{ANN}	1	
22	5.25%	0.031	0.367	0.755	81.34%	0.067	0.558	0.824	13.41%	0.036	0.694	0.772	1163	T_{DJF}, P_{ANN}	1	4
23		0.059	-0.189	0.540		0.097	0.363	0.747		0.086	0.262	0.450		T_{DJF}, T_{JJA}	1	
24		0.031	0.368	0.756		0.065	0.570	0.829		0.034	0.709	0.783		T_{JJA}, P_{ANN}	1	
25		0.029	0.408	0.771		0.092	0.397	0.761		0.065	0.444	0.586		T_{DJF}, CWD_{ANN}	1	
26		0.035	0.302	0.730		0.079	0.476	0.792		0.048	0.585	0.691		T_{JJA}, CWD_{ANN}	1	
27		0.036	0.268	0.717		0.066	0.564	0.827		0.035	0.702	0.778		T_{DJF}, T_{JJA}, CWD_{ANN}	1	
28		0.032	0.362	0.753		0.063	0.585	0.835		0.031	0.729	0.798		T_{DJF}, T_{JJA}, P_{ANN}	1	

continued on next page...

. . . continuation

No.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
29	8.50%	0.025	0.675	0.819	28.23%	0.038	0.815	0.817	63.27%	0.018	0.923	0.945	294	T_{DJF}^{850hpa}, P_{ANN}	2	1
30		0.017	0.788	0.882		0.034	0.831	0.833		0.035	0.851	0.892		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}	2	
31		0.014	0.817	0.898		0.043	0.787	0.790		0.034	0.855	0.895		T_{JJA}^{850hpa}, P_{ANN}	2	
32		0.032	0.594	0.774		0.044	0.785	0.787		0.030	0.871	0.907		T_{DJF}^{850hpa}, CWD_{ANN}	2	
33		0.019	0.759	0.866		0.072	0.647	0.651		0.070	0.697	0.781		T_{JJA}^{850hpa}, CWD_{ANN}	2	
34		0.019	0.760	0.866		0.029	0.859	0.861		0.025	0.893	0.923		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, CWD_{ANN}	2	
35		0.015	0.801	0.889		0.026	0.871	0.873		0.018	0.924	0.945		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}	2	
36	5.41%	0.010	0.809	0.924	44.90%	0.043	0.825	0.834	49.68%	0.033	0.868	0.880	314	T_{DJF}^{850hpa}, P_{ANN}	2	2
37		0.010	0.797	0.920		0.044	0.823	0.832		0.043	0.827	0.843		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}	2	
38		0.005	0.898	0.960		0.064	0.741	0.754		0.059	0.764	0.787		T_{JJA}^{850hpa}, P_{ANN}	2	
39		0.009	0.824	0.930		0.041	0.834	0.843		0.035	0.861	0.875		T_{DJF}^{850hpa}, CWD_{ANN}	2	
40		0.009	0.815	0.927		0.091	0.632	0.651		0.091	0.637	0.672		T_{JJA}^{850hpa}, CWD_{ANN}	2	
41		0.011	0.782	0.913		0.045	0.820	0.829		0.040	0.839	0.854		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, CWD_{ANN}	2	
42		0.010	0.813	0.926		0.047	0.810	0.820		0.037	0.851	0.865		T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}	2	

continued on next page. . .

…continuation

No.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
43	17.04%	0.037	0.741	0.782	39.39%	0.066	0.723	0.727	43.58%	0.029	0.882	0.887	358	T^{850hpa}_{DJF}, P_{ANN}	2	3
44		0.017	0.881	0.900		0.048	0.798	0.801		0.035	0.858	0.864		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2	
45		0.024	0.831	0.858		0.071	0.702	0.707		0.052	0.789	0.797		T^{850hpa}_{JJA}, P_{ANN}	2	
46		0.025	0.821	0.849		0.052	0.784	0.787		0.035	0.859	0.865		T^{850hpa}_{DJF}, CWD_{ANN}	2	
47		0.018	0.876	0.895		0.089	0.627	0.633		0.080	0.676	0.689		T^{850hpa}_{JJA}, CWD_{ANN}	2	
48		0.015	0.897	0.913		0.045	0.810	0.813		0.034	0.862	0.868		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2	
49		0.019	0.867	0.888		0.051	0.785	0.788		0.033	0.865	0.871		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2	
50	5.25%	0.028	0.436	0.782	81.34%	0.061	0.596	0.840	13.41%	0.033	0.713	0.786	1163	T^{850hpa}_{DJF}, P_{ANN}	2	4
51		0.026	0.478	0.798		0.078	0.489	0.797		0.050	0.566	0.677		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}	2	
52		0.022	0.554	0.828		0.046	0.698	0.880		0.024	0.795	0.847		T^{850hpa}_{JJA}, P_{ANN}	2	
53		0.031	0.367	0.755		0.059	0.611	0.846		0.031	0.733	0.801		T^{850hpa}_{DJF}, CWD_{ANN}	2	
54		0.029	0.411	0.773		0.059	0.611	0.846		0.032	0.728	0.797		T^{850hpa}_{JJA}, CWD_{ANN}	2	
55		0.019	0.624	0.855		0.034	0.778	0.912		0.015	0.874	0.906		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, CWD_{ANN}	2	
56		0.015	0.691	0.881		0.033	0.781	0.913		0.018	0.847	0.886		T^{850hpa}_{DJF}, T^{850hpa}_{JJA}, P_{ANN}	2	
57	8.50%	0.066	0.154	0.528	28.23%	0.219	-0.081	-0.067	63.27%	0.119	0.490	0.632	294	T_{DJF}, P_{ANN}	3	1
58		0.081	-0.043	0.418		0.195	0.039	0.051		0.287	-0.233	0.110		T_{DJF}, T_{JJA}	3	
59		0.177	-1.270	-0.267		0.286	-0.410	-0.392		0.086	0.629	0.732		T_{JJA}, P_{ANN}	3	
60		0.308	-2.964	-1.212		0.182	0.103	0.114		0.320	-0.377	0.006		T_{DJF}, T_{JJA}, P_{ANN}	3	

continued on next page…

. . . continuation

No.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
61	5.41%	0.033	0.359	0.746	44.90%	0.259	-0.046	0.008	49.68%	0.195	0.221	0.297	314	T_{DJF}, P_{ANN}	3	2
62		0.159	-2.106	-0.231		0.207	0.162	0.205		0.184	0.263	0.334		T_{DJF}, T_{JJA}	3	
63		0.283	-4.524	-1.190		0.444	-0.795	-0.702		0.145	0.421	0.477		T_{JJA}, P_{ANN}	3	
64		0.549	-9.718	-3.249		0.393	-0.588	-0.507		0.205	0.182	0.261		T_{DJF}, T_{JJA}, P_{ANN}	3	
65	17.04%	0.074	0.477	0.560	39.39%	0.237	0.009	0.024	43.58%	0.129	0.475	0.496	358	T_{DJF}, P_{ANN}	3	3
66		0.255	-0.805	-0.520		0.202	0.156	0.169		0.235	0.043	0.082		T_{DJF}, T_{JJA}	3	
67		0.329	-1.329	-0.960		0.393	-0.647	-0.622		0.127	0.483	0.505		T_{JJA}, P_{ANN}	3	
68		0.457	-2.235	-1.723		0.365	-0.529	-0.506		0.177	0.280	0.310		T_{DJF}, T_{JJA}, P_{ANN}	3	
69	5.25%	0.074	-0.495	0.422	81.34%	0.338	-1.228	0.115	13.41%	0.251	-1.161	-0.610	1163	T_{DJF}, P_{ANN}	3	4
70		0.054	-0.082	0.582		0.110	0.277	0.713		0.086	0.263	0.451		T_{DJF}, T_{JJA}	3	
71		0.082	-0.644	0.365		0.328	-1.161	0.142		0.238	-1.045	-0.524		T_{JJA}, P_{ANN}	3	
72		0.051	-0.036	0.600		0.173	-0.140	0.547		0.125	-0.076	0.198		T_{DJF}, T_{JJA}, P_{ANN}	3	
73	8.50%	0.053	0.322	0.622	28.23%	0.196	0.034	0.046	63.27%	0.111	0.520	0.654	294	T_{DJF}, P_{ANN}	4	1
74		0.071	0.085	0.490		0.154	0.239	0.248		0.204	0.124	0.368		T_{DJF}, T_{JJA}	4	
75		0.055	0.297	0.608		0.211	-0.040	-0.027		0.168	0.277	0.478		T_{JJA}, P_{ANN}	4	
76		0.039	0.493	0.717		0.158	0.219	0.228		0.131	0.437	0.594		T_{DJF}, T_{JJA}, P_{ANN}	4	

continued on next page. . .

. . . continuation

No.	RF_{B_1}	BS_{B_1}	$BSS^{RF}_{B_1}$	$BSS^{ZV}_{B_1}$	RF_{B_2}	BS_{B_2}	$BSS^{RF}_{B_2}$	$BSS^{ZV}_{B_2}$	RF_{B_3}	BS_{B_3}	$BSS^{RF}_{B_3}$	$BSS^{ZV}_{B_3}$	N_{obs}	*predictors*	*KlimDataSet*	*biome setup*
77	5.41%	0.027	0.481	0.794	44.90%	0.234	0.054	0.103	49.68%	0.184	0.264	0.335	314	T_{DJF}, P_{ANN}	4	2
78		0.045	0.123	0.652		0.170	0.312	0.347		0.192	0.231	0.305		T_{DJF}, T_{JJA}	4	
79		0.035	0.308	0.725		0.230	0.070	0.118		0.194	0.222	0.297		T_{JJA}, P_{ANN}	4	
80		0.028	0.456	0.785		0.147	0.406	0.436		0.120	0.521	0.567		$T_{DJF}, T_{JJA}, P_{ANN}$	4	
81	17.04%	0.045	0.683	0.733	39.39%	0.215	0.098	0.112	43.58%	0.141	0.425	0.448	358	T_{DJF}, P_{ANN}	4	3
82		0.113	0.199	0.326		0.166	0.307	0.317		0.205	0.167	0.201		T_{DJF}, T_{JJA}	4	
83		0.057	0.600	0.663		0.226	0.052	0.066		0.169	0.312	0.340		T_{JJA}, P_{ANN}	4	
84		0.073	0.481	0.563		0.172	0.280	0.291		0.104	0.578	0.595		$T_{DJF}, T_{JJA}, P_{ANN}$	4	
85	5.25%	0.043	0.142	0.668	81.34%	0.172	-0.134	0.550	13.41%	0.121	-0.040	0.225	1163	T_{DJF}, P_{ANN}	4	4
86		0.053	-0.062	0.590		0.096	0.366	0.748		0.083	0.283	0.466		T_{DJF}, T_{JJA}	4	
87		0.049	0.004	0.615		0.142	0.064	0.628		0.090	0.224	0.422		T_{JJA}, P_{ANN}	4	
88		0.049	0.023	0.622		0.095	0.374	0.752		0.047	0.596	0.699		$T_{DJF}, T_{JJA}, P_{ANN}$	4	

B.4. Additional Figures Palaeoclimate Reconstruction Result

All figures presented in this section show results for the palaeoclimate reconstruction which bases on the sediment core of Lake Kinneret. The discussion of the results can be found in chapter 4.3. The mapping of the figures is identical to Figure 4.4:
All figures show the marginal posterior pdf $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ with a coloured scale, ranging from blue for low values to red for high values. Each solid black line marks the mode, which is the highest value of each pdf in each sample layer. The dashed black lines mark the 10 %, 25 %, 75 % and the 90 % quantile. The dotted lines represent the median, which is equal to the 50 % quantile. Each red line marks the weighted arithmetic mean for each layer. The weights are the values of the marginal pdf, which is interpreted as the numerically calculated expectation value. The bold black numbers transfer the marked ticks of the depth axis in y. cal BP (s. chap. 2.3, p. 17) according to the age-to-depth model established by Schiebel (2013) and mentioned in chapter 2.3.2.1. The local pollen assemblage zones (LPAZ) are also shown and labelled with the according age.

All presented Figures base on a three-dimensional climate state vector $\vec{C}$ with different setup dependent on the s.no. which defines also the biome setup and the estimation dataset. Summarized, B.19 is as a supplement to already presented Figures. Fig. B.20, B.21 and B.22 show the influence of different biome setups on the palaeoclimate reconstructions as discussed in chapter 4.3. And finally Fig. B.23 and B.24 show the influence of the factor ϑ_{k_0} as also discussed in chapter 4.3. In detail:

Fig. B.19 Is a supplement to Figure 4.6 (4.6a and 4.6b) in which T_{JJA} is estimated with the standard climate prior compared to Fig. 4.7. It bases on $\vartheta_{k_0} = ind.$ and s.no. 56.
($\Leftrightarrow \vec{C} = \left(T_{DJF}^{850hpa}, T_{JJA}^{850hpa}, P_{ANN}\right)^T$, biome setup 4 and ERA as database for the estimation)

Fig. B.20 Has to be compared with Figure 4.4a to 4.4c and bases on $\vartheta_{k_0} = ind.$ and s.no. 7.
($\Leftrightarrow \vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$, biome setup 1 and CRU TS 3.1 as database for the estimation)

Fig. B.21 Has to be compared with Figure 4.4a to 4.4c and bases on $\vartheta_{k_0} = ind.$ and s.no. 14.
($\Leftrightarrow \vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$, biome setup 2 and CRU TS 3.1 as database for the estimation)

Fig. B.22 Has to be compared with Figure 4.4a to 4.4c and bases on $\vartheta_{k_0} = ind.$ and s.no. 28.
($\Leftrightarrow \vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$, biome setup 4 and CRU TS 3.1 as database for the estimation)

Fig. B.23 Has to be compared with Figure 4.4a to 4.4c and bases on $\vartheta_{k_0} = 0$ and s.no. 21.
($\Leftrightarrow \vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$, biome setup 3 and CRU TS 3.1 as database for the estimation)

Fig. B.24 Has to be compared with Figure 4.4a to 4.4c and bases on $\vartheta_{k_0} = Q_{k_0}\left(\vec{\tilde{\omega}}_{k_0}\right) = 33\,\%$ and s.no. 21, $Q_{k_0}\left(\vec{\tilde{\omega}}_{k_0}\right) = 33\,\%$ is the 33 % quantile of $\vec{\tilde{\omega}}_{k_0}$.
($\Leftrightarrow \vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$, biome setup 3 and CRU TS 3.1 as database for the estimation)

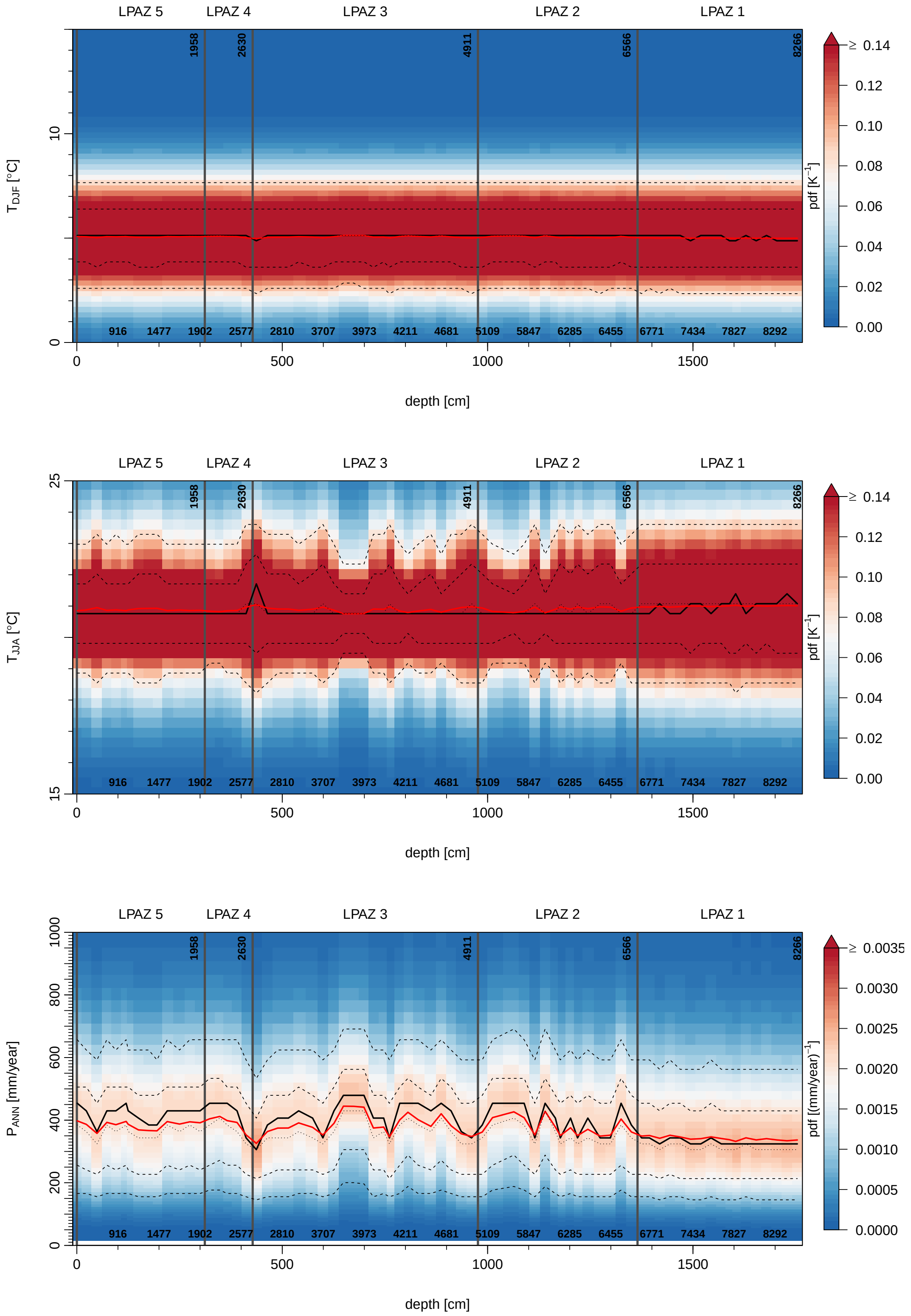


Figure B.19.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}^{850hpa}, T_{JJA}^{850hpa} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 56).

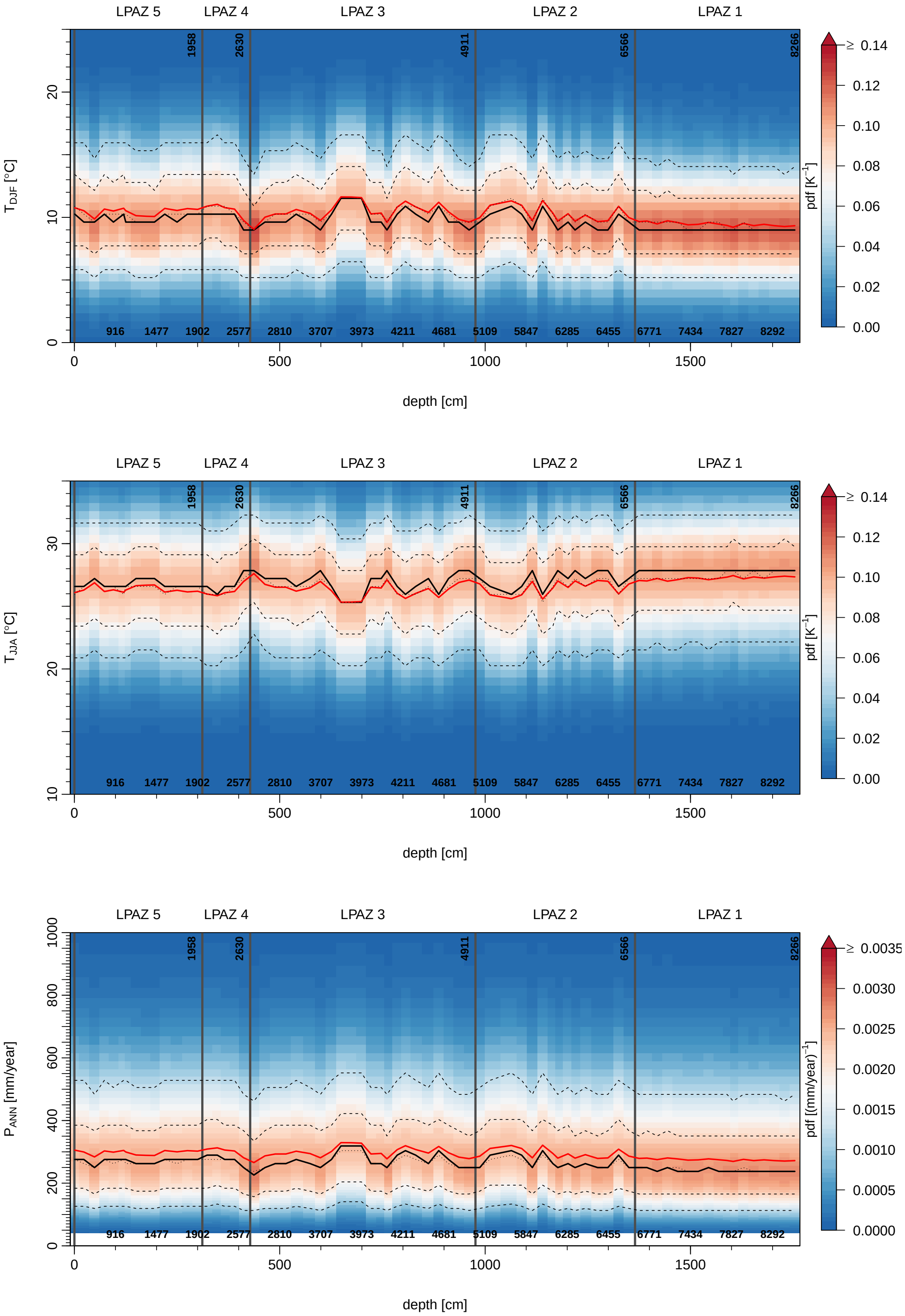


Figure B.20.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 7).

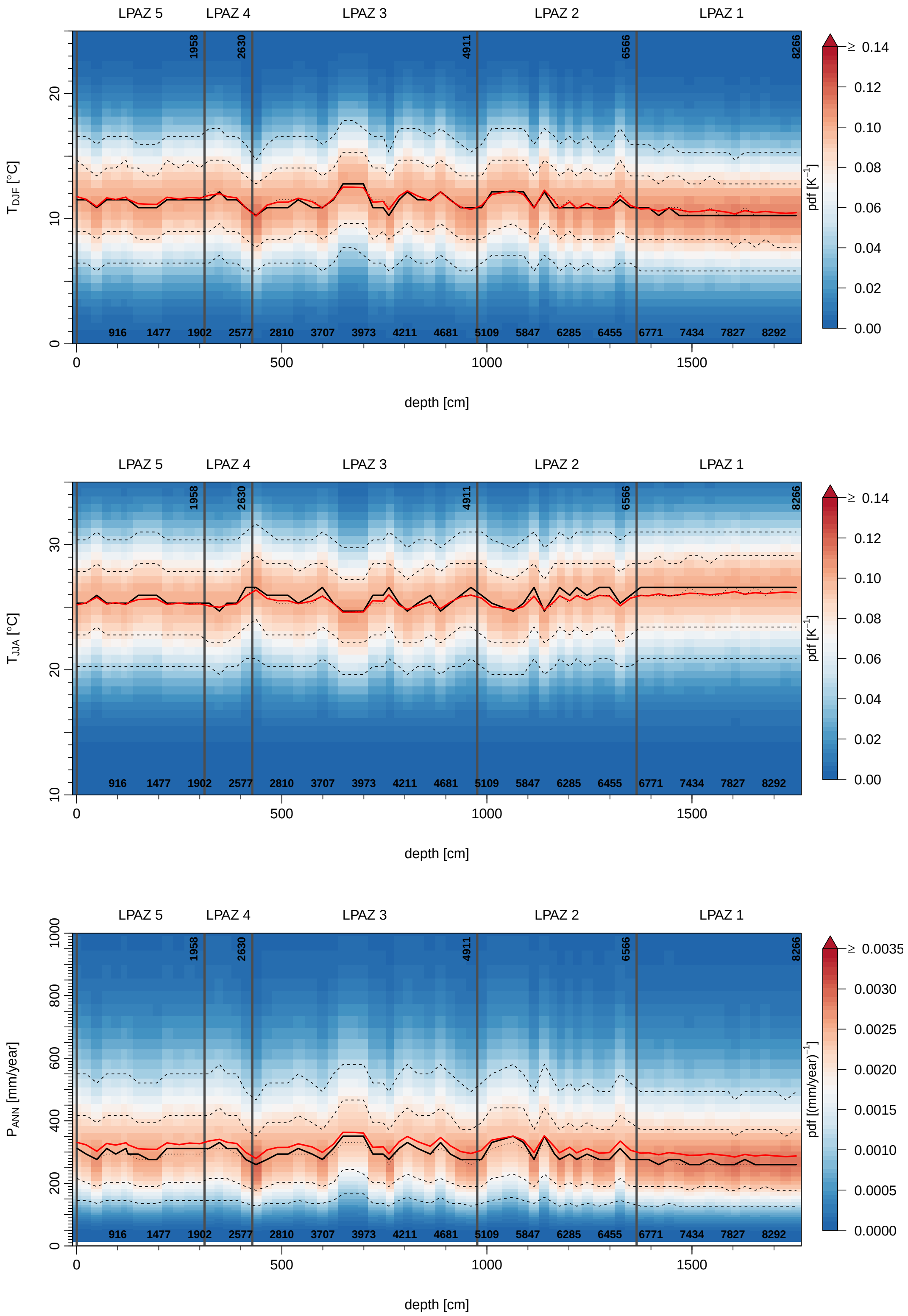


Figure B.21.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 14).

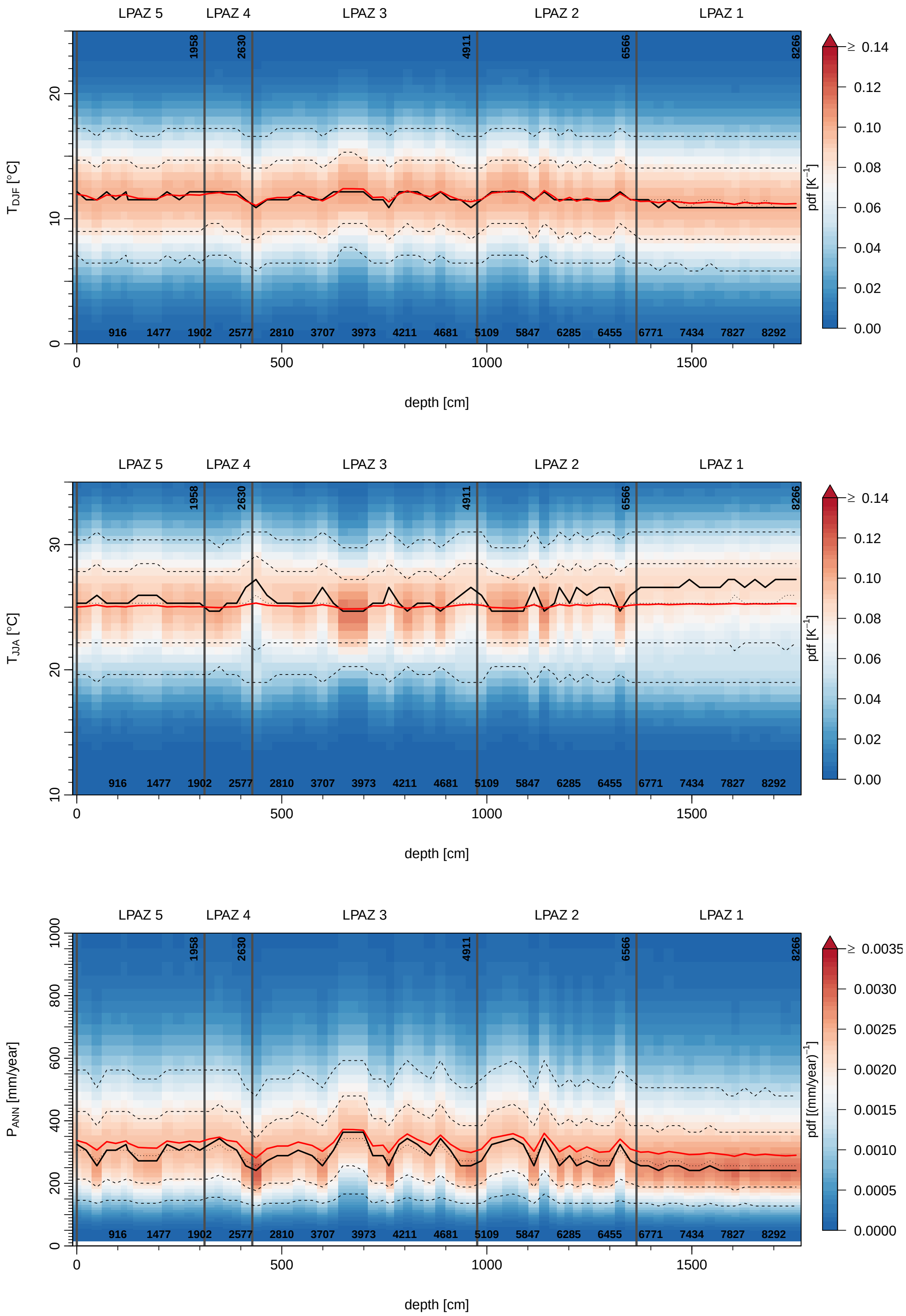


Figure B.22.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 28).

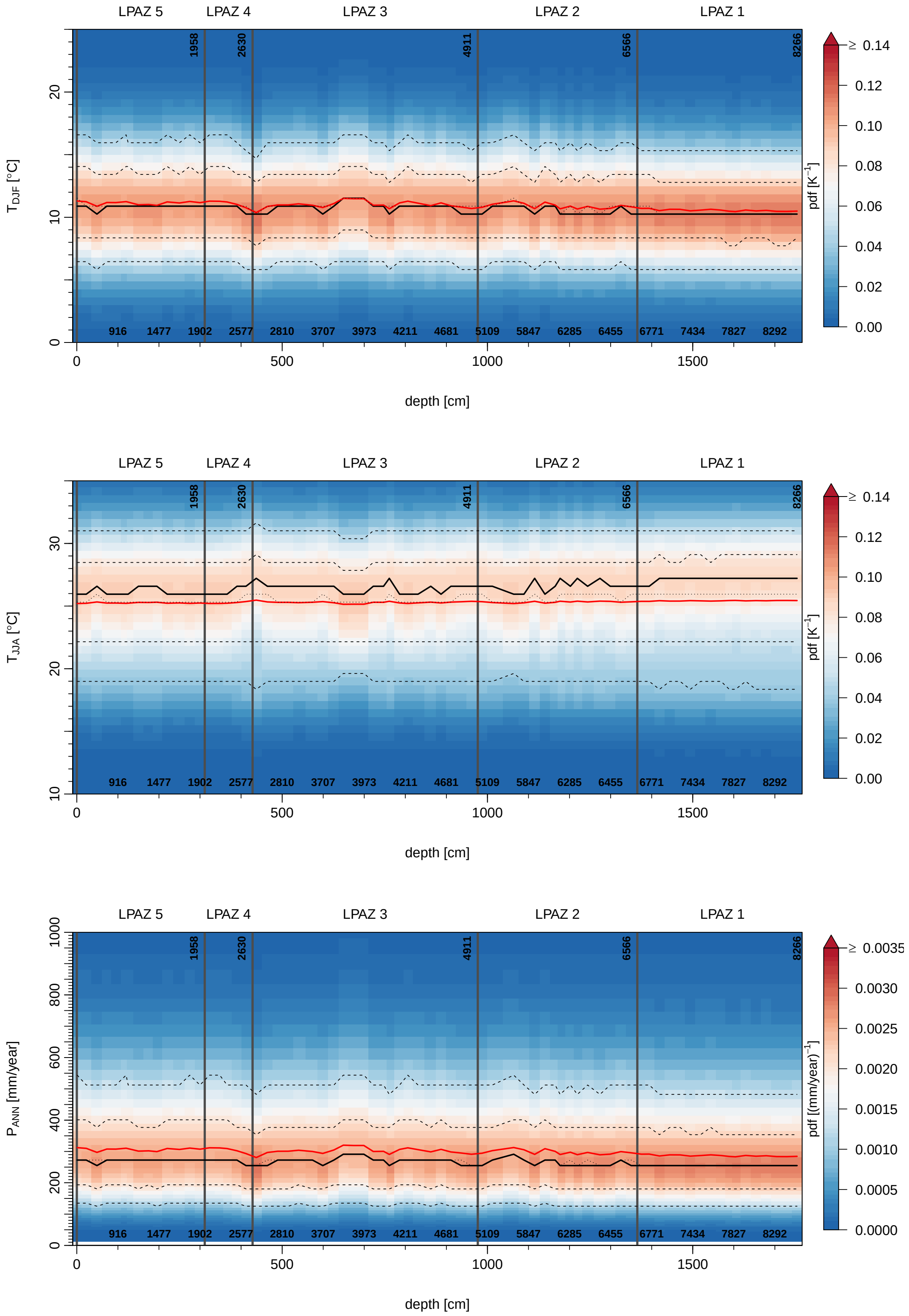


Figure B.23.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = 0$ and s.no. 21).

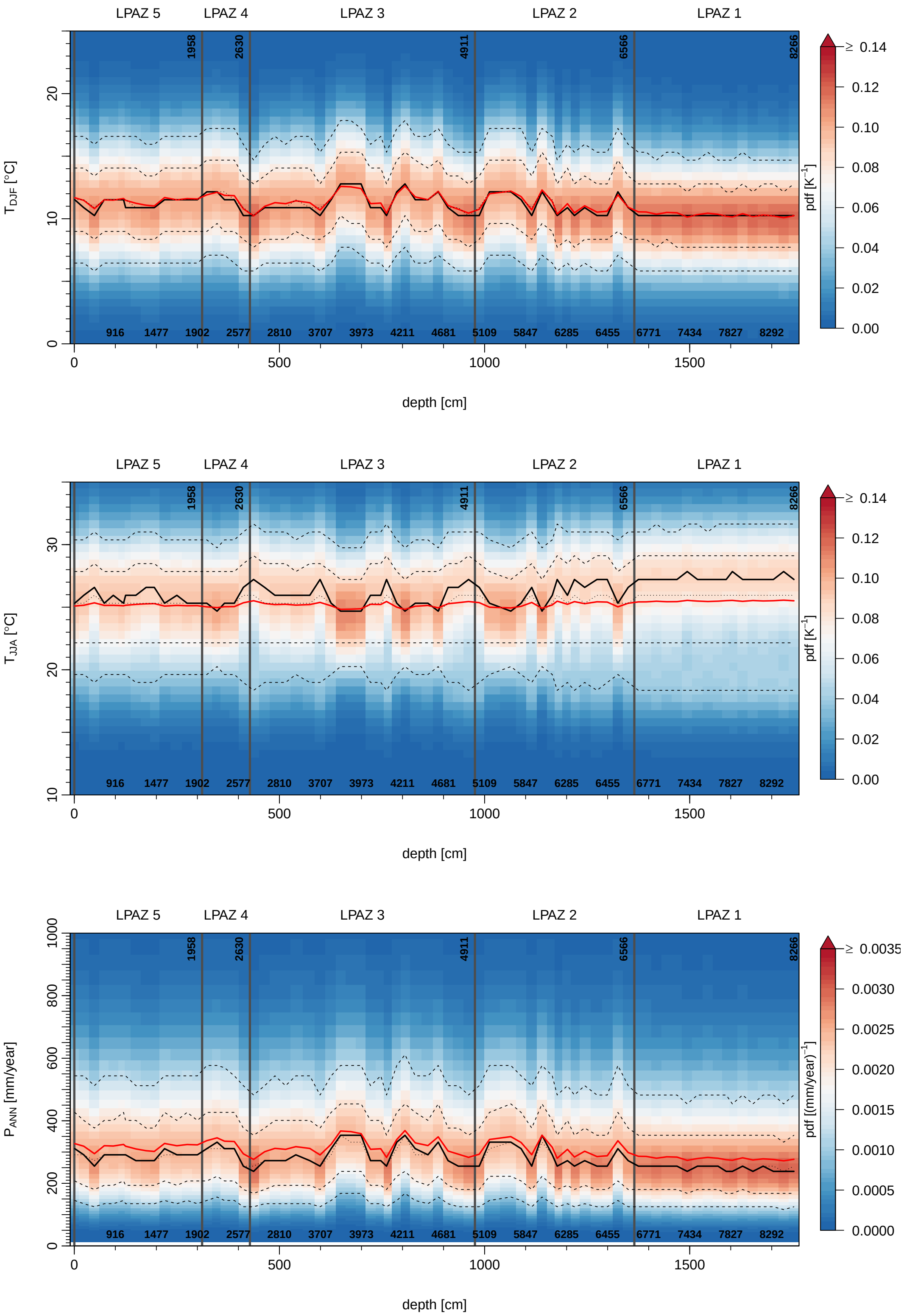


Figure B.24.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}\left(c_0|\vec{p}_0\right)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = Q_{k_0}\left(\vec{\vec{\omega}}_{k_0}\right) = 33\,\%$ and s.no. 21).

B.5. Pollen-Ratio Model

This section presents all data tables required for the PRM as applied in chapter 4.4. All tables are already described in chapter 2.2.2. Table B.7 and B.8 are the digitized versions of Table 6.1 in Weinstein (1979). The connected geographical coordinates are shown in Table B.6 and visualized in Figure 2.12. The coloured locality/site numbers in Table B.6, B.7 and B.8 indicate whether the location is situated in the Mediterranean (green), Irano-Turanian (red) or Saharo-Arabian biome area (orange) defined in Figure 2.11b.

Table B.6.: Geographical coordinates of the various localities with Airborne Pollen Spectra in Israel (alternative spelling = a. s.)

Locality		Geographical coordinates				comment
No.	Name	Google Maps		E-OBS		
		[°*N*]	[°*E*]	[°*N*]	[°*E*]	
1	Qiryat-Shemona	33.20794167	35.57023056	33.125	35.625	a. s.: Kirjat-Schmona
2	Zefat	32.96465	35.49601389	32.875	35.375	a. s.: Safed
3	Tiberias	32.79586111	35.530975	32.875	35.625	
4	Nazareth	32.69962778	35.30355833	NA	NA	
5	Nahariyya	33.01133889	35.09466389	NA	NA	
6	Haifa	32.83029722	34.97433889	32.875	35.125	critical, great shift
7	Bet-She'an	32.49710278	35.49733611	32.375	35.375	
8	Afula	32.61048056	35.28794444	32.625	35.375	
9	Pardes-Hannah	32.47632778	34.97606944	NA	NA	
10	Netanya	32.32151111	34.85323889	NA	NA	
11	Tel-Aviv	32.06615833	34.77782222	32.125	34.875	
12	Jerusalem	31.76831667	35.21371111	31.875	35.125	
13	Jericho	31.85698056	35.46056667	31.875	35.375	
14	Hebron	31.53256667	35.09983333	31.625	35.125	
15	Rehovot	31.89276944	34.811275	31.875	34.875	
16	Qiryat-Gat	31.61115	34.76846389	31.625	34.875	a. s.: Kirjat-Gat
17	Ashqelon	31.66594444	34.55946667	31.625	34.625	
18	Be'er-Sheva	31.25296944	34.79148056	31.375	34.875	critical, great shift

Table B.7.: Tab. 6.1 after Weinstein (1979) Part 1 - Annual averages of airborne pollen spectra in various localities in Israel (totally counted pollen = 100%). The coloured taxon name indicates the assigned biome type defined in Table B.1 for Lake Kinneret (Mediterranean or Irano-Turanian biome type).

Locality-No.	1	2	3	4	5	6	7	8	9
Taxon / Locality	Qiryat-Shemona	Zefat	Tiberias	Nazareth	Nahariyya	Haifa	Bet-She'an	Afula	Pardes-Hannah
Quercus	6.4	4.3	2.4	4.9	3.5	5.6	8.5	2.7	6.1
Pinus	2.6	4.2	2.8	20.2	6.4	7.5	1.6	9.5	8.5
Olea	8	17	2.9	6.9	6.2	5.4	15	3.9	9.8
Pistacia	0.4	0.3	0.2	0.8	0.8	1	0.2	0.3	0.6
Cupressus	6.3	12.7	1.7	4.1	5	9.3	1.1	8.5	3.9
Eucalyptus	9	9	34.5	14.9	20.5	20	11.6	19.7	7.1
Casuarina	4.6	0.7	4.6	4.2	11.4	8	2.4	4.6	6.6
Rosaceous trees	NA	0.6	0.5	0.5	1.9	0.2	1	0.7	0.2
Other arboreal pollen	1.9	0.7	2.1	1.4	4.6	2.5	2.1	3.9	2.1
Total arboreal pollen	39.2	49.5	51.7	56.9	60.3	59.5	43.5	53.8	44.9
Cereals	2.5	1.5	2.4	0.7	1.6	0.7	3.3	1.4	0.9
Gramineae	13.9	12.9	11.1	8.5	10.7	13.3	17.1	10.5	13.4
Compositae	7.4	1.8	4.9	2.9	4.6	2.7	3.8	1.2	1.9
Cruciferae	2.3	1.5	0.5	2	0.9	0.7	0.7	2.1	9.1
Chenopodiaceae	6.2	3.5	5.5	7.8	2.7	8	15.4	6.2	12.1
Umbelliferae	2	2.9	1.8	0.6	2.7	0.9	7.5	5.3	0.5
Papilionaceae	2	0.9	3.9	1.9	1.7	1	0.9	0.9	1.8
Artemisia	14.3	14.5	7.2	5.9	3.9	6.8	0.6	3.8	8
Ephedra	0.1	0.2	0.1	0.8	0.3	0.1	0.2	0.3	0.1
Centaurea	0.9	1.4	0.3	0.9	0.9	0.2	1.6	0.6	1.1
Poterium	0.2	2.5	0.3	3.8	0.1	1	0.1	0.1	0.1
Other nonarboreal pollen	9	6.9	8.7	5.7	6.3	4.4	7.1	13.5	6.1
Total nonarboreal pollen	60.8	50.5	48.3	42.1	39.7	40.5	56.5	46.2	55.1

Table B.8.: Tab. 6.1 after Weinstein (1979) Part 2 - Annual averages of airborne pollen spectra in various localities in Israel (Total counted pollen = 100%). The coloured taxon name indicates the assigned biome type defined in Table B.1 for Lake Kinneret (Mediterranean or Irano-Turanian biome type).

Locality No.	10	11	12	13	14	15	16	17	18
Locality / Taxon	Netanya	Tel-Aviv	Jerusalem	Jericho	Hebron	Rehovot	Qiryat-Gat	Ashqelon	Be'er-Sheva
Quercus	4.2	2.2	3.8	1.7	8	1.3	0.6	1.9	1.1
Pinus	7.8	5.6	8.7	1.4	3.9	25.2	3	2.8	2.6
Olea	9.2	8.3	7.4	2.6	3.3	2.1	4.6	9	1.8
Pistacia	0.1	0.1	0.1	0.1	0.2	0.4	NA	0.1	0.3
Cupressus	2	11	1.3	9.7	0.1	9.2	0.5	0.8	0.2
Eucalyptus	24.9	12.8	5	0.8	3.6	13	12.8	14.1	23.5
Casuarina	4	4.9	5.3	9.6	6.9	9.5	19.5	13	17.2
Rosaceous trees	0.7	0.5	1	0.6	0.8	0.4	1.3	1.1	2.3
Other arboreal pollen	1.6	2.4	2.3	11.4	1.6	1.8	5.4	2.4	10.6
Total arboreal pollen	54.5	47.8	35	37.9	28.4	58.8	46.7	45.2	59.6
Cereals	0.7	0.9	2.9	0.7	3.5	1.7	2.9	0.6	0.8
Gramineae	10.1	11.2	15.2	7.6	16.2	7.7	15.9	12.9	8.2
Compositae	5.1	3.4	5.4	1.7	4.1	3.1	2.2	4.5	2.3
Cruciferae	1	1.4	3.6	2.5	2.6	1.4	2.9	0.4	1.4
Chenopodiaceae	12.3	10.3	11.6	25.7	8.7	6.7	7.9	10.1	11.1
Umbelliferae	0.4	1.8	3	0.6	0.1	0.8	2	3.1	0.3
Papilionaceae	1	1.9	2	2.2	1.8	1.3	1.5	1.3	0.5
Artemisia	8.6	8.5	9.7	12.9	16.9	11.5	9.3	13.4	2.9
Ephedra	0.6	0.1	0.4	0.1	0.4	1.5	0.1	5	0.1
Centaurea	1.1	2.9	0.9	0.2	0.1	0.6	1.1	1.8	1
Poterium	0.2	0.8	3.6	0.9	7.9	0.5	0.2	0.4	NA
Other nonarboreal pollen	4.3	7.6	4.8	6.2	5.9	5.4	7.1	1.3	11.8
Total nonarboreal pollen	45.5	52.2	65	62.1	71.6	41.2	53.3	54.8	40.4

C

Ein Gedi

This appendix chapter presents additional information and figures for the palaeoclimate reconstruction of Ein Gedi which bases on the BBM.

C.1. Threshold and Taxa Biome Assignment Table

This section presents the table with all defined threshold values ϑ_{k_0} and entries of the biome assignment matrix $\mathcal{M}^{N_{k_0} \times N_{l_0}}$ as required for the estimation of the biome probability $W_{B_{l_0}}$ (s. chap. 4.2.2) and defined in Litt *et al.* (2012).

Table C.1.: Taxa used for the palaeoclimate reconstruction of Ein Gedi, the assigned biome type and the corresponding thresholds ϑ_{k_0} after Litt *et al.* (2012).
The assigned biome type is indicated by the column vectors $\vec{m}_{l_0}$ of the biome assignment matrix $\mathcal{M}^{N_{k_0} \times N_{l_0}}$. It describes whether the regarded taxon k_0 belongs to this biome type l_0 ($m_{k_0 l_0} = 1$) or not ($m_{k_0 l_0} = 0$). In general the first column $\vec{m}_1$ represents a Mediterranean, the second column $\vec{m}_2$ a Irano-Turanian and the third column $\vec{m}_3$ a Saharo-Arabian biome type.

No.	Taxon name	Shortcut	$\vec{m}_1$	$\vec{m}_2$	$\vec{m}_3$	Threshold [%]
1	*Acacia*		0	0	1	0.10
2	*Arbutus*		1	0	0	0.10
3	*Artemisia*	Artem_K	0	1	0	2.00
4	*Cedrus*		1	0	0	0.05
5	*Centaurea*	Centa_K	0	1	0	0.80
6	*Ceratonia*		1	0	0	0.10
7	*Cerealea type*	Cerea_K	0	1	0	1.00
8	*Chenopodiaceae*	Cheno_K	0	0	1	40.00
9	*Cistus type*	Cistu_K	1	0	0	0.10
10	*Cupressaceae*		1	0	0	0.10
11	*Ephedra*	Ep_dt_K	0	1	0	0.50
12	*Helianthemum*		1	0	0	0.05

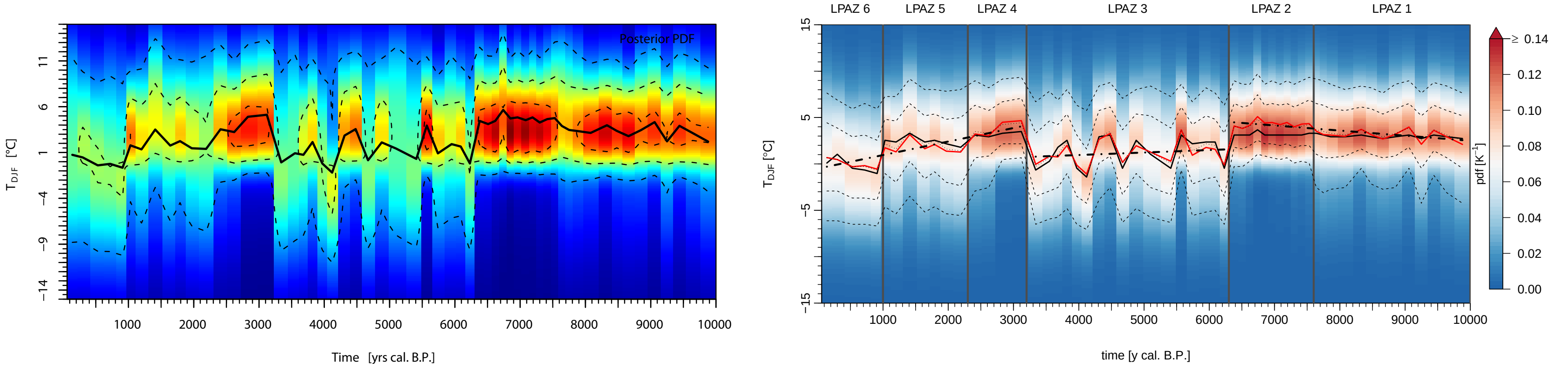
continued on next page…

. . . continuation

No.	Taxon name	Shortcut	$\vec{m}_1$	$\vec{m}_2$	$\vec{m}_3$	Threshold [%]
13	*Laurus*		1	0	0	0.10
14	*Olea*	Olea__K	1	0	0	1.50
15	*Phillyrea*	Phill_K	1	0	0	0.10
16	*Phoenix*		0	0	1	0.10
17	*Pinus*	Pinus_K	1	0	0	0.50
18	*Pistacia*	Pista_K	1	0	0	0.40
19	*Poaceae*	Poace_K	0	1	0	4.00
20	*Quercus calliprinos type*	Qu_ca_K	1	0	0	1.00
21	*Quercus ithaburensis type*	Qu_it_K	1	0	0	0.10
22	*Sarcopoterium*		1	0	0	0.30
23	*Tamarix*		0	0	1	0.20
24	*Tubuliflorae*	Tubul_K	0	1	0	10.00
25	*Zizyphus*		0	0	1	0.05
26	*Zygophylum*		0	0	1	0.10

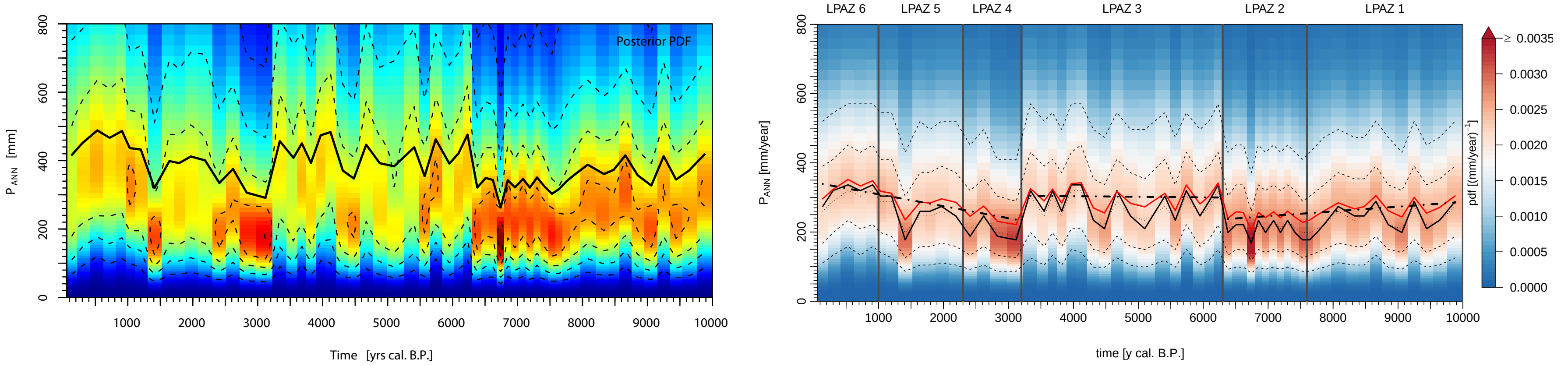
C.2. Palaeoclimate Reconstruction Result: Additional Figures

This section presents Figure C.1 which is a supplement to Figure 4.11. Figure C.1 shows the results for the palaeoclimate reconstruction which bases on the sediment core of Ein Gedi but estimated with nearly the same prior but different input climate datasets: CRU TS 1.0 (Litt *et al.* 2012) and CRU TS 3.1. The detailed discussion of the results can be found in chapter 4.5. The only difference in the prior distributions $\pi_{\vec{C}}\left(\vec{c}_0\right)$ for the climate state vector in Figure C.1 is the definition of the P_{ANN} component. As mentioned on page 70 in the discussion of Figure 4.10 the Jordan Valley values (s. Tab. 4.1) are applied in the estimation of Figure C.1b and C.1d. Also discussed there are the differences between the original Litt *et al.* (2012) prior and the Jordan Valley prior.



(a) Litt *et al.* (2012), Fig. 5b

(b) zoom out of Fig. 4.11a



(c) Litt *et al.* (2012), Fig. 5c

(d) zoom in of Fig. 4.11c

Figure C.1.: Marginal posterior pdfs $f_{C_0|\vec{P}_0}(c_0|\vec{p}_0)$ for T_{DJF} and P_{ANN} for the Ein Gedi palaeoclimate reconstruction estimated with different climate datasets: CRU TS 1.0 (Litt *et al.* (2012)) and CRU TS 3.1. All four figures are estimated with $\vartheta_{k_0} = ind.$ and biome setup 1. In detail (a) and (c) are the coloured versions of the figures presented in Litt *et al.* (2012)(Image courtesy of Dipl. Met. Dr. Christian Ohlwein). The pdfs are shown with a coloured scale, ranging from blue for low values to cyan to green to yellow and finally to red for highest values. (b) and (d) are identical to Figure 4.11a and 4.11c, only the ordinate is adjusted so that they match with (a) and (c).

D

Birkat Ram

This appendix chapter presents additional information and figures for the palaeoclimate reconstruction of Birkat Ram which bases on the BITBM.

D.1. Thresholds and Biome Assignment

First this section presents the table with all defined threshold values ϑ_{k_0} and entries of the biome assignment matrix $\mathcal{M}^{N_{k_0} \times N_{l_0}}$ as required for the estimation of the biome probability $W_{B_{l_0}}$ (s. chap. 4.2.2). The second part presents the depth profile figures of the relative abundance $\vec{\tilde{\omega}}_{k_0}$ and the corresponding ECDF for each regarded taxon listed in Table D.1.

D.1.1. Threshold and Taxa Biome Assignment Table

Table D.1.: Taxa used for the palaeoclimate reconstruction of Birkat Ram, the assigned biome type and the corresponding thresholds ϑ_{k_0}.
The assigned biome type is indicated by the column vectors $\vec{m}_{l_0}$ of the biome assignment matrix $\mathcal{M}^{N_{k_0} \times N_{l_0}}$. It describes whether the regarded taxon k_0 belongs to this biome type l_0 ($m_{k_0 l_0} = 1$) or not ($m_{k_0 l_0} = 0$). In general the first column $\vec{m}_1$ represents a Mediterranean and the second column $\vec{m}_2$ a Irano-Turanian biome type.

No.	Taxon name	Shortcut	$\vec{m}_1$	$\vec{m}_2$	Threshold [%]
1	*Artemisia*	Arte_BR	0	1	4
2	*Plantago*	PlSM_BR	0	1	0.44
3	*Pistacia*	Pieu_BR	1	0	0.8
4	*Olea*	Oleu_BR	1	0	3
5	*Quercus ithaburensis type*	Qud_BR	1	0	9
6	*Quercus calliprinos type*	Que_BR	1	0	1.2

D.1.2. Threshold ϑ_{k_0} Definition

The mapping of all following figures in this chapter is the same as for Figure 3.2: On the left side is shown the depth profile of the relative abundance $\vec{\widetilde{\omega}}_{k_0}$ and on the right side the corresponding ECDF. The depth profile of the relative abundance $\widetilde{\omega}_{\iota\,k_0}$ in each pollen sample ι is shown in dark grey. The light grey area is the ten times and the very light grey the 100 times enlarged profile. The threshold value ϑ_{k_0} is visualized by the different shaded areas in the depth profile and by the vertical bold black line in the ECDF. The resulting presence/absence information is visualized in the depth profile with black dots (=presence) and circles (=absence). These dots also indicate whether the taxon in this pollen sample layer contribute to the biome probability $W_{B_{l_0}}$ in this layer (=black) or not (=circle) (s. chap. 5.2.2).

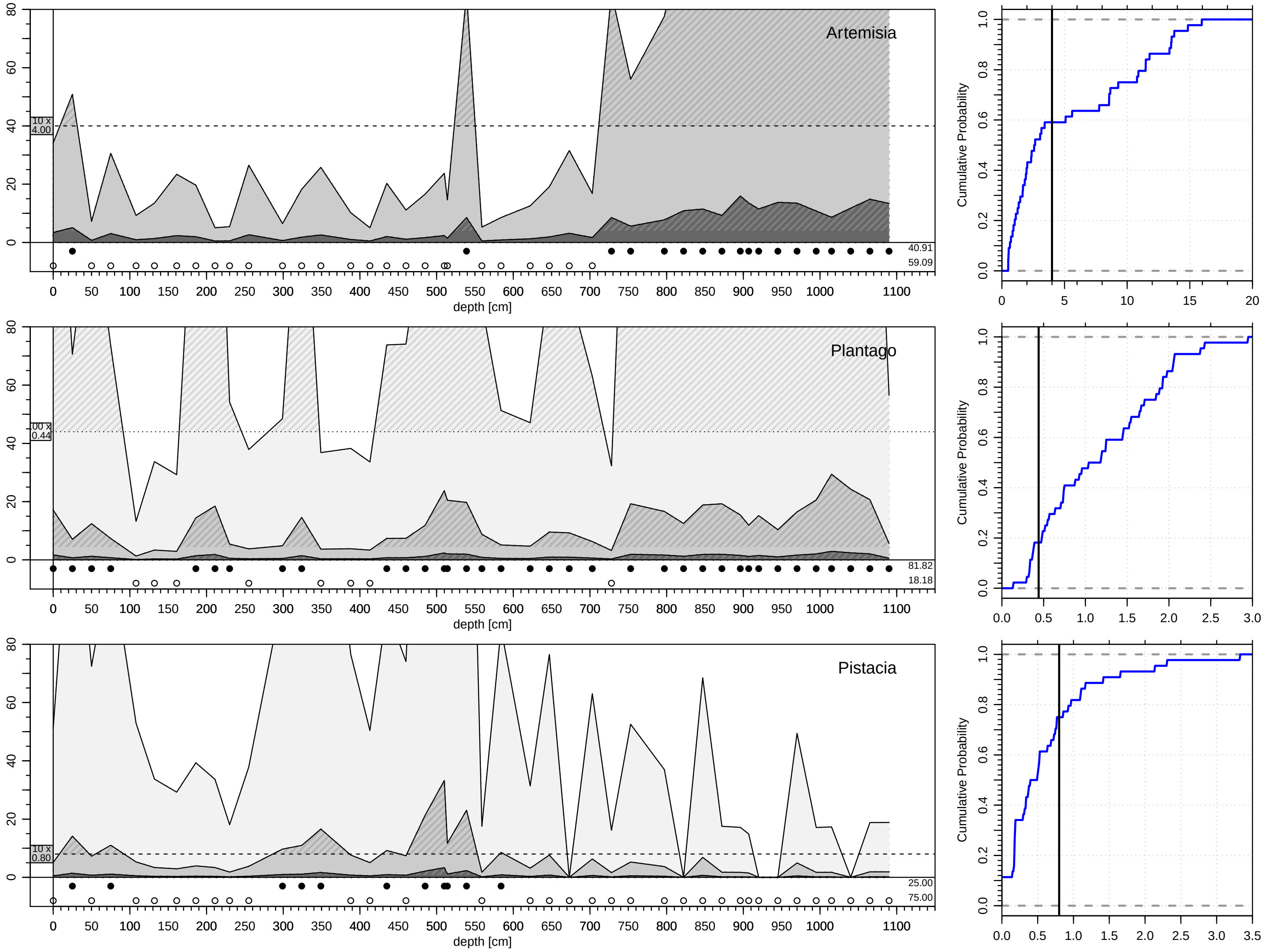


Figure D.1.: Birkat Ram, definitions of the thresholds (part 1).

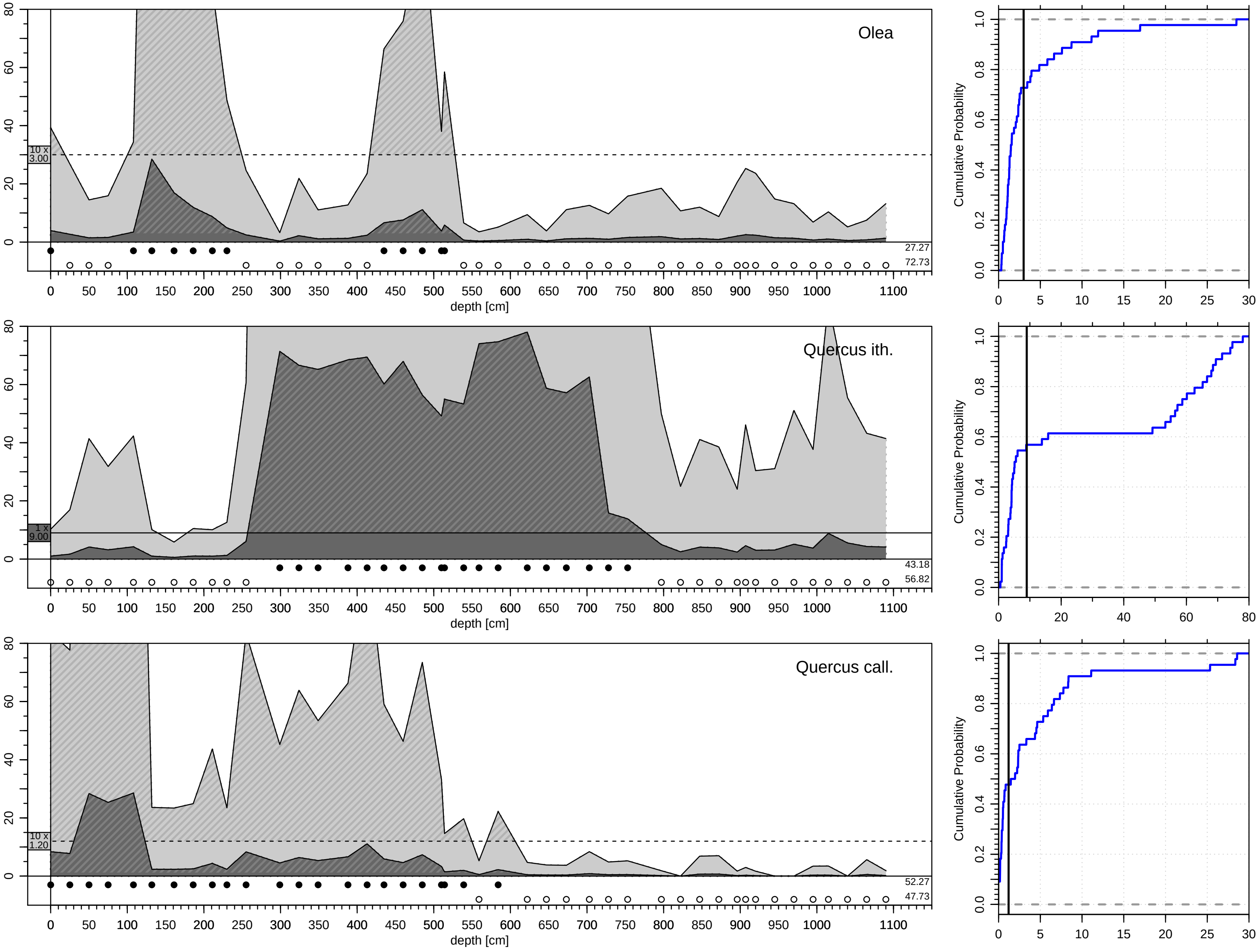


Figure D.2.: Birkat Ram, definitions of the thresholds (part 2).

E

Jordan Valley

This appendix chapter presents additional information and figures for the Jordan Valley CFR .

E.1. Additional Information on the Cost Function

This section presents the table with all values used for the weighting factor γ_M in the cost function. γ_M is a weighting factor as mentioned at the end of chapter 6.2.1 on page 99 which is defined by several minimization runs so that both parts of the cost function ($\mathcal{J}_{veg}$ and $\mathcal{J}_{Mod}$) contribute 50% of the cost function value $\mathcal{J}_R$ at the minimum or in other words $\frac{\mathcal{J}_{veg}}{\gamma_M \mathcal{J}_{Mod}}$ is in the order of one. This section presents also the mentioned figure which indicates the possible ranges in the climate phase space for the minimization with limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm for box-constrained optimization (L-BFGS-B) (at the end of chapter 6.2.3 on page 103). Equation A.1 presents $\vec{\beta}$, the values for the modern climate state which is shown with solid lines in Figure E.1.

Table E.1.: Values for the weighting factor γ_M in the cost function.

No.	**time slice**	$\gamma_M[1]$	$\frac{\mathcal{J}_{veg}}{\gamma_M \mathcal{J}_{Mod}}[1]$
1	0910 y. cal BP	8000	1.019
2	1200 y. cal BP	8125	1.027
3	1790 y. cal BP	7250	1.061
4	1970 y. cal BP	7500	1.014
5	5570 y. cal BP	7750	1.036
6	6740 y. cal BP	7375	1.069
7	7190 y. cal BP	8000	1.013

$$\vec{\beta} = \begin{pmatrix} 0.3806 & 0.4725 & -4.7846 & -0.0076 & -0.0139 & 0.0819 & -0.0064 & -0.0025 & -0.0027 \end{pmatrix}^T \tag{A.1}$$

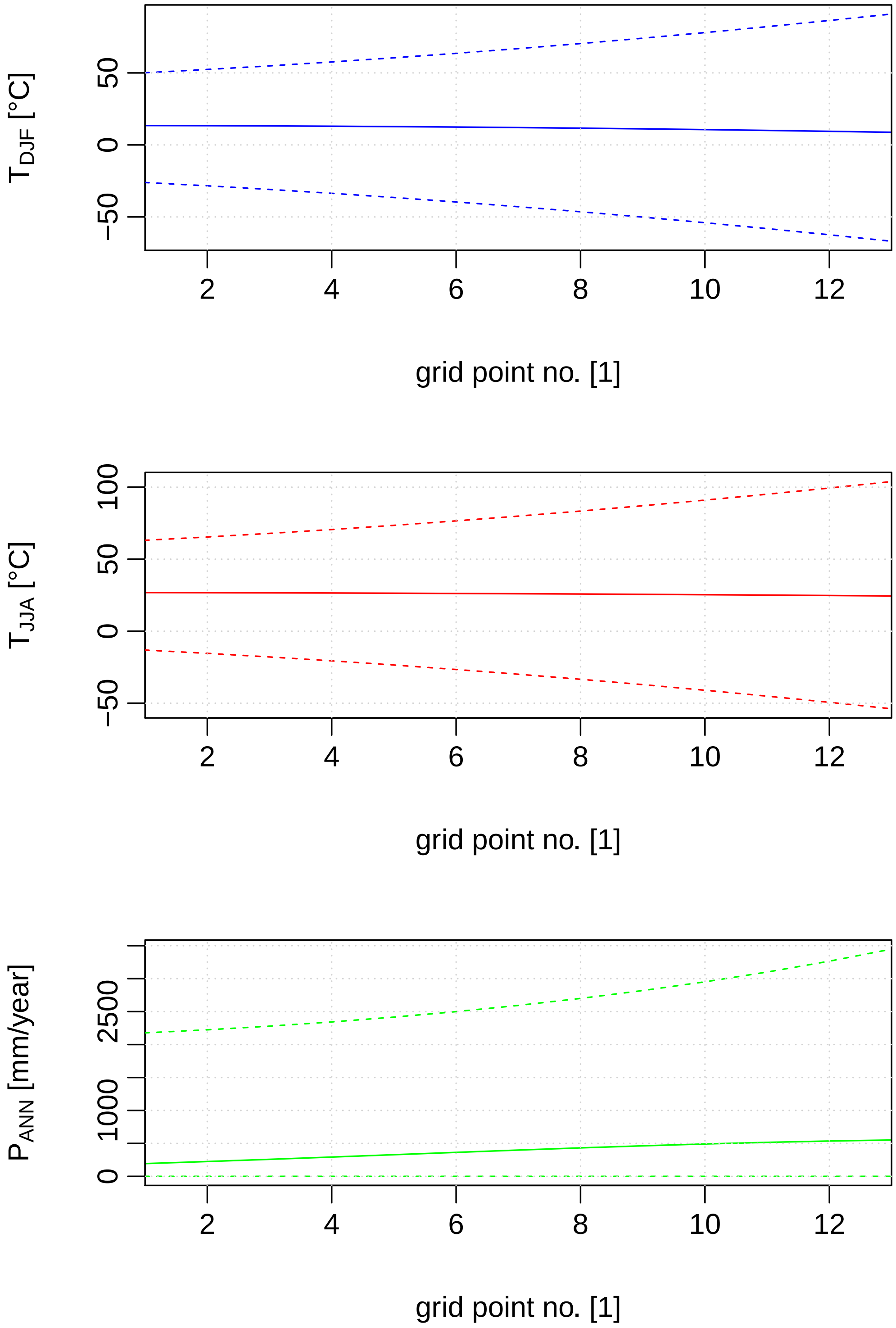


Figure E.1.: Resulting $\vec{c}_{R,0,h_{mean}}$ ranges for the estimation of the minimum of the cost function at $\vec{\beta}_{R,0}$ with the R-routine *optim* and the setup L-BFGS-B (s. chap. 6.2.3 for details). The solid lines represent the modern climate state profiles based on $\vec{\beta}$ (eq. A.1) as presented for example in Figure 6.2 for T_{DJF}. The dashed lines represent $\vec{\beta}_{upper,0}$ resp. $\vec{\beta}_{lower,0}$ as defined in equation 6.23a resp. 6.23b. For the transformation to $\vec{c}_{R,0,h_{mean}}$ equation 6.24a is applied.

E.2. Additional Figures CRU TS 3.1-E-OBS Mixture- Palaeoclimate Reconstruction Results

This section presents the local palaeoclimate reconstructions results based on a mixture of the CRU TS 3.1 and E-OBS dataset as mentioned and discussed in chapter 6.3. In detail these are the marginal posterior distributions ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ presented in Figure E.2, E.3 and E.4. All base on the biome probability $W_{B_{l_0}}$ estimated with an individual definition of $\vartheta_{k_0} = individual$ and s.no. 84 (biome setup 3, $\vec{C} = (T_{DJF}, T_{JJA}, P_{ANN})^T$ and a combination of CRU TS 3.1 and E-OBS as estimation dataset). Thereby

Ein Gedi Figure E.2 is the counterpart for 4.16a, 4.16b and 4.16c,

Lake Kinneret Figure E.3 is the counterpart for 4.4a, 4.4band 4.4c,

Birkat Ram Figure E.4 is the counterpart for 5.3a, 5.3b and 5.3c.

The mapping of the figures is identical to Figure 4.4:
All figures show the marginal distributions pdf ${}^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ with a coloured scale, ranging from blue for low values to red for high values. Each solid black line marks the mode which is the highest value of each pdf in each sample layer. The dashed black lines mark the 10 %, 25 %, 75 % and the 90 % quantile. The dotted lines are the median which is equal to the 50 % quantile. Each red line marks the weighted arithmetic mean for each layer. The weights are the values of the marginal pdf which is interpreted as the numerical calculated expectation value.

For Lake Kinneret (Fig. E.3) the bold black numbers transfer the marked ticks of the depth axis in y. cal BP according to the age-to-depth model established by Schiebel (2013). The LPAZ also defined in Schiebel (2013) are shown and labelled with the according age.

For the Ein Gedi and Birkat Ram figures the abscissa represents the age in y. cal BP. The age-to-depth model for Ein Gedi is, as mentioned in Figure 4.11, established by Migowski *et al.* (2004, 2006). For Ein Gedi the LPAZ defined in Litt *et al.* (2012) are also shown. For Birkat Ram (Fig. E.4) the age-to-depth model as well as the LPAZ are established by Schiebel (2013) as mentioned at the beginning of chapter 5.3. The hatched area between ≈ 10 000 and 17 000 y. cal BP (≡ 703 cm and 746 cm) marks again as in Figure 5.4 the range where Schiebel (2013) assumed a desiccation of Birkat Ram. As in Figure 5.4 the mode is not shown in Figure E.4.

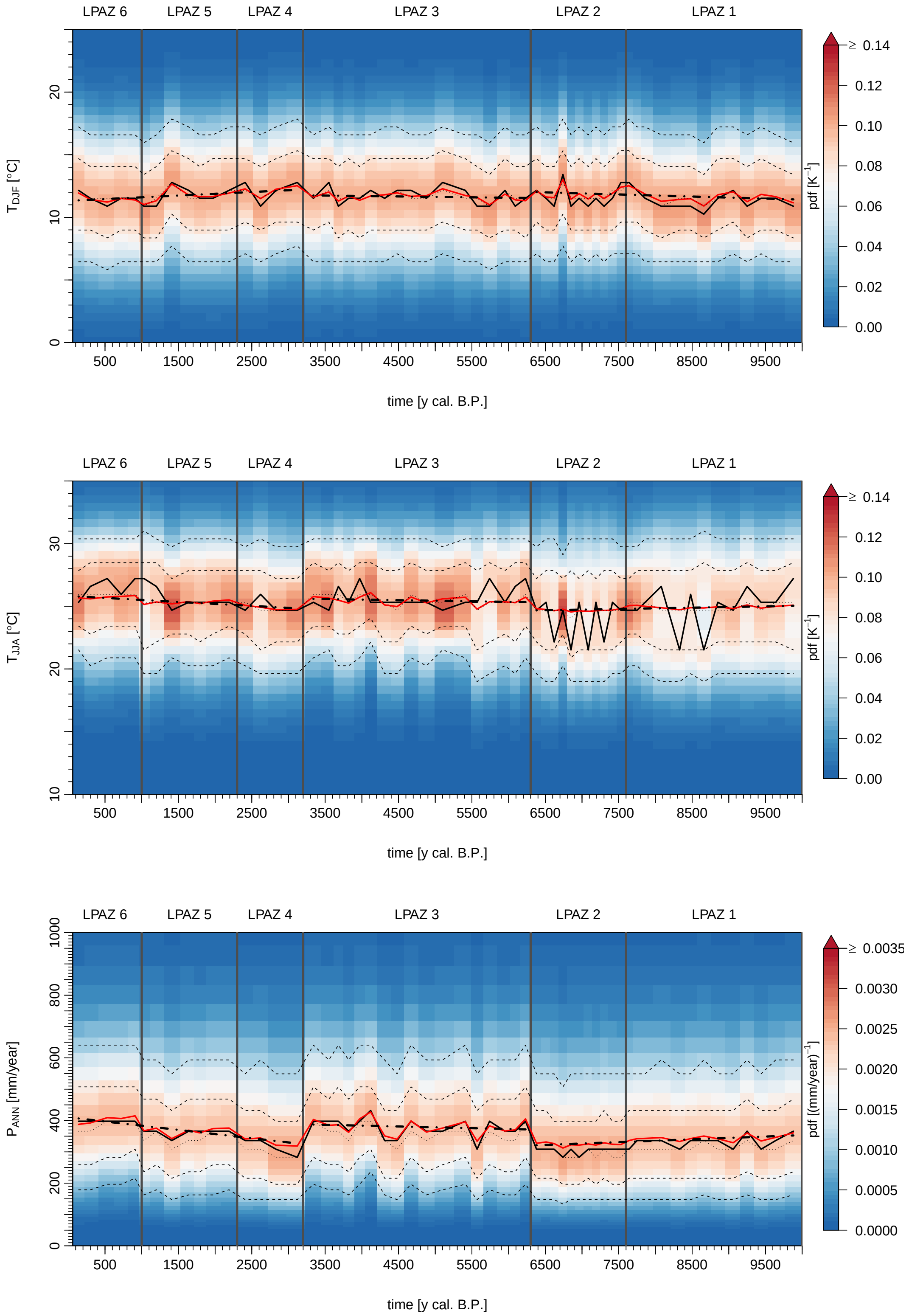


Figure E.2.: Marginal posterior pdfs $^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Ein Gedi palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 84).

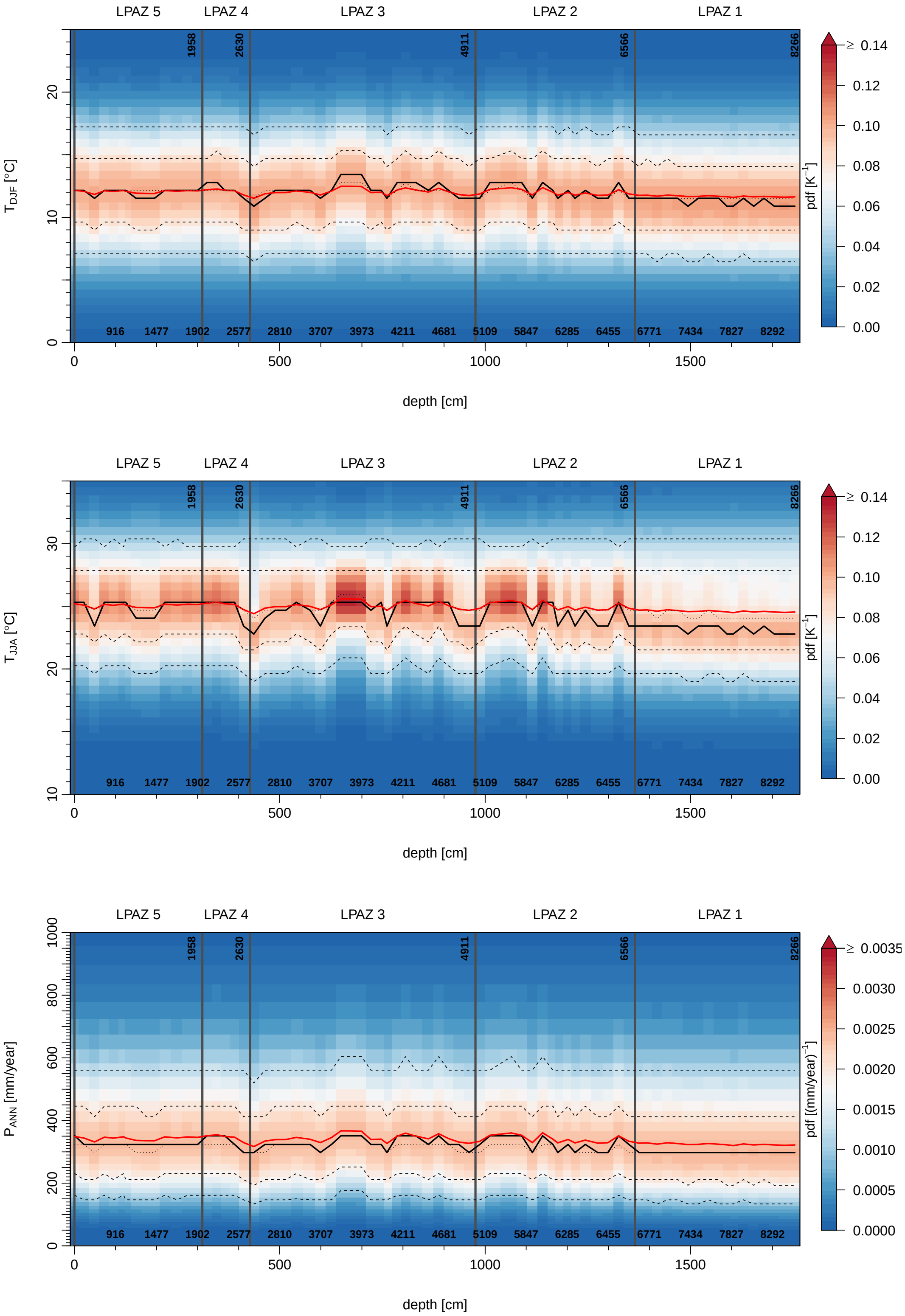


Figure E.3.: Marginal posterior pdfs $^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Lake Kinneret palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 84).

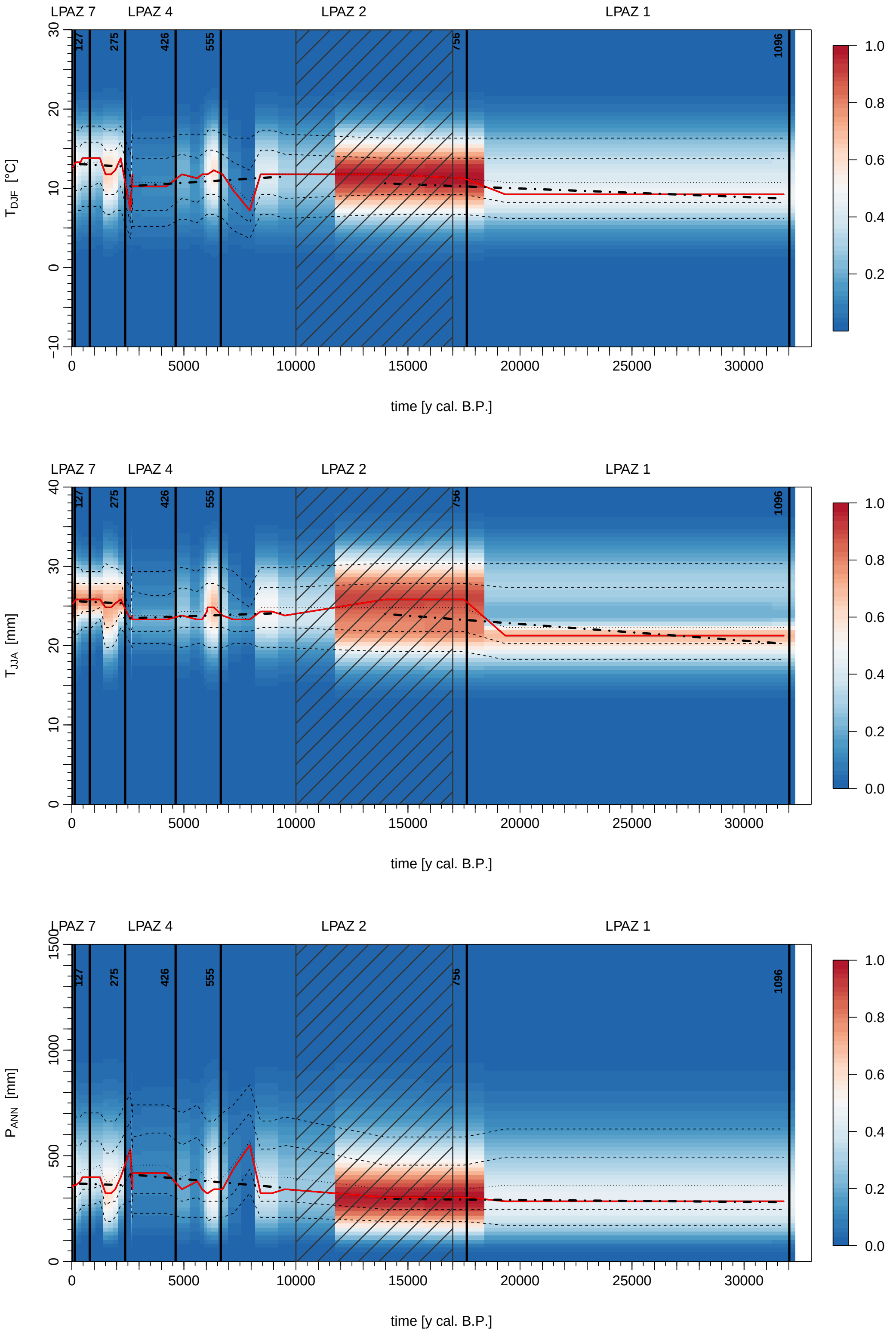


Figure E.4.: Marginal posterior probability $^{s}f_{\vec{C}|\vec{P}}\left(c_0|\vec{p}_0\right)$ for T_{DJF}, T_{JJA} and P_{ANN} for the Birkat Ram palaeoclimate reconstruction ($\vartheta_{k_0} = ind.$ and s.no. 84).

E.3. Additional Figures and Tables for the Jordan Valley Palaeoclimate Reconstruction

This section presents the remaining figures of the Jordan Valley CFR. The detailed discussion can be found in chapter 6.5.

The first part (chap. E.3.1) presents the profile reconstructions for the time slices 0910 y. cal BP, 1200 y. cal BP, 1790 y. cal BP, 1970 y. cal BP, 6740 y. cal BP and 7190 y. cal BP which are not shown in chapter 6.5. The mapping of Figure E.5, E.6, E.7, E.8, E.9 and E.10 is the same as in Figure 6.5 and 6.6: The abscissa represents the Jordan Valley grid boxes. The blue (T_{DJF}) resp. red (T_{JJA}) resp. green (P_{ANN}) data points represent the available data at each grid point (E-OBS and CRU TS 3.1). The vertical black lines mark the locations of the sediment cores Ein Gedi, Lake Kinneret and Birkat Ram. The dashed black line indicates the result of a GLM with quadratic term. The solid black lines are the corresponding 99.5 % confidence bands. Presented in orange is the palaeoclimate profile (minimum of $\mathcal{J}$) with the same mapping of the line types as the modern time slice GLM except that the confidence bands represent the 68.3% interval based on $\pm\sigma_{u_3} = \pm\sigma_{u_{P_{ANN}}}$. The values for the fossil sites are marked with an orange dot (EG = Ein Gedi), a triangle (LK = Lake Kinneret) and a square (BR = Birkat Ram).

The second part (chap. E.3.2) presents the remaining palaeoclimate CFR results for Ein Gedi and Birkat Ram summarized as a time series and already mentioned and discussed in chapter 6.5.2. The mapping of Figure E.11 and E.12 is the same as in Figure 6.7 and 6.8: The abscissa represents the age in y. cal BP. Indicated in blue colour (T_{DJF}) resp. red colour (T_{JJA}) resp. green colour (P_{ANN}) is the palaeoclimate CFR result based on an application of equation 6.24b. The values ${}^{EG}\ldots_{R,0,h_{mean}}$ resp. ${}^{BR}\ldots_{R,0,h_{mean}}$ (minimum of $\mathcal{J}$) are marked with a dot in Figure E.11 resp. a square in Figure E.12. The error bars represent the 68.3% interval based on $\pm\sigma_{u_3} = \pm\sigma_{u_{\ldots}}$ (s. eq.6.25). The mapping for the modern time slice GLM value for Ein Gedi resp. Birkat Ram is the same as in Figure 6.2 except that here it represents only the value at grid point no. 3 resp. 11.

Table E.2 presents the database for all figures presented in chapters 6.5.2 and E.3.2.

E.3.1. Profile Results

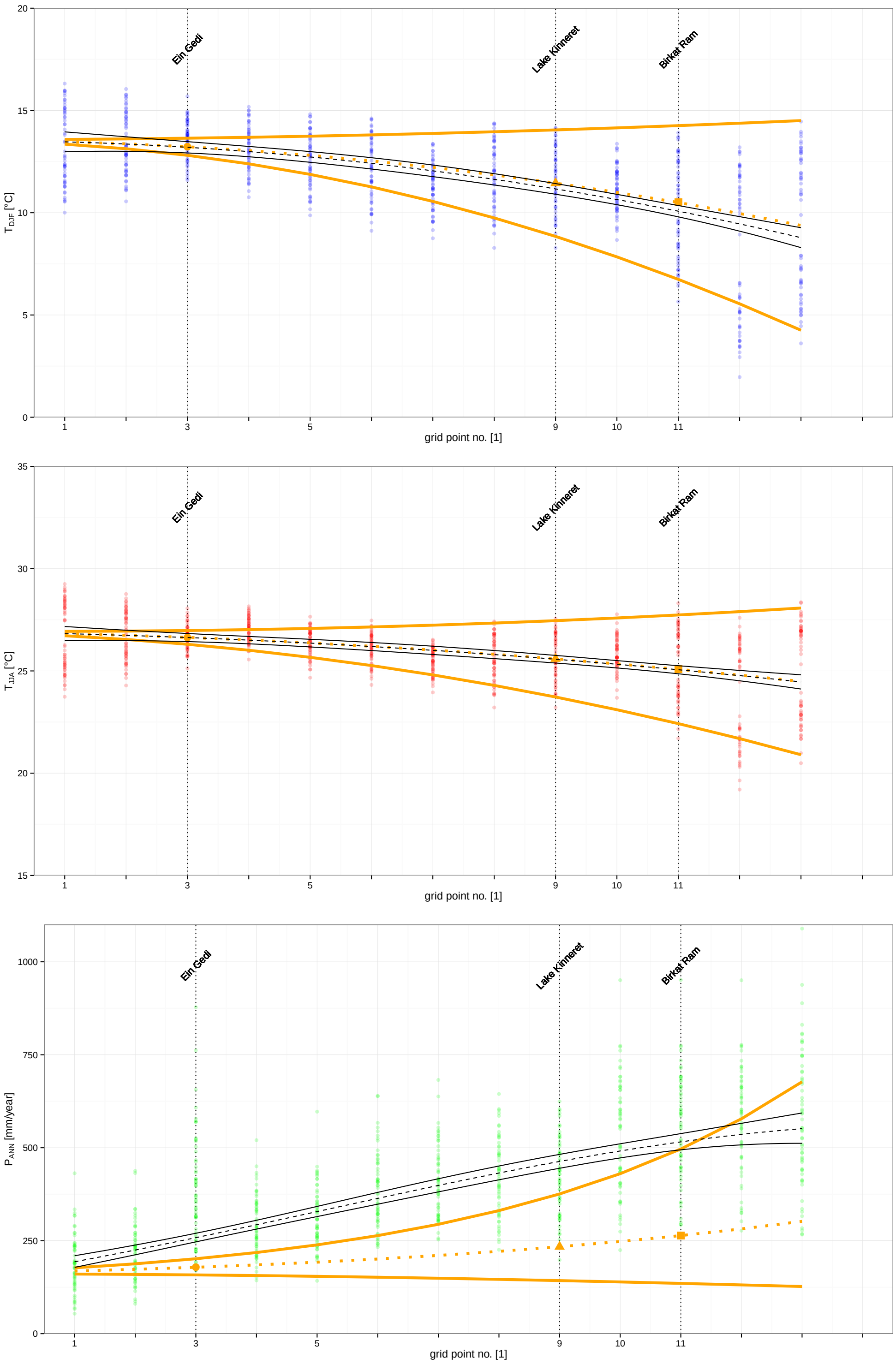


Figure E.5.: The Jordan Valley profile reconstruction result for time slice 910 y. cal BP and the climate variables $T_{DJF\,h_{mean}}$, $T_{JJA\,h_{mean}}$ and $P_{ANN\,h_{mean}}$. The mapping of the figures is defined analogously to Fig. 6.5 and in detail repeated in the beginning of chapter E.3.

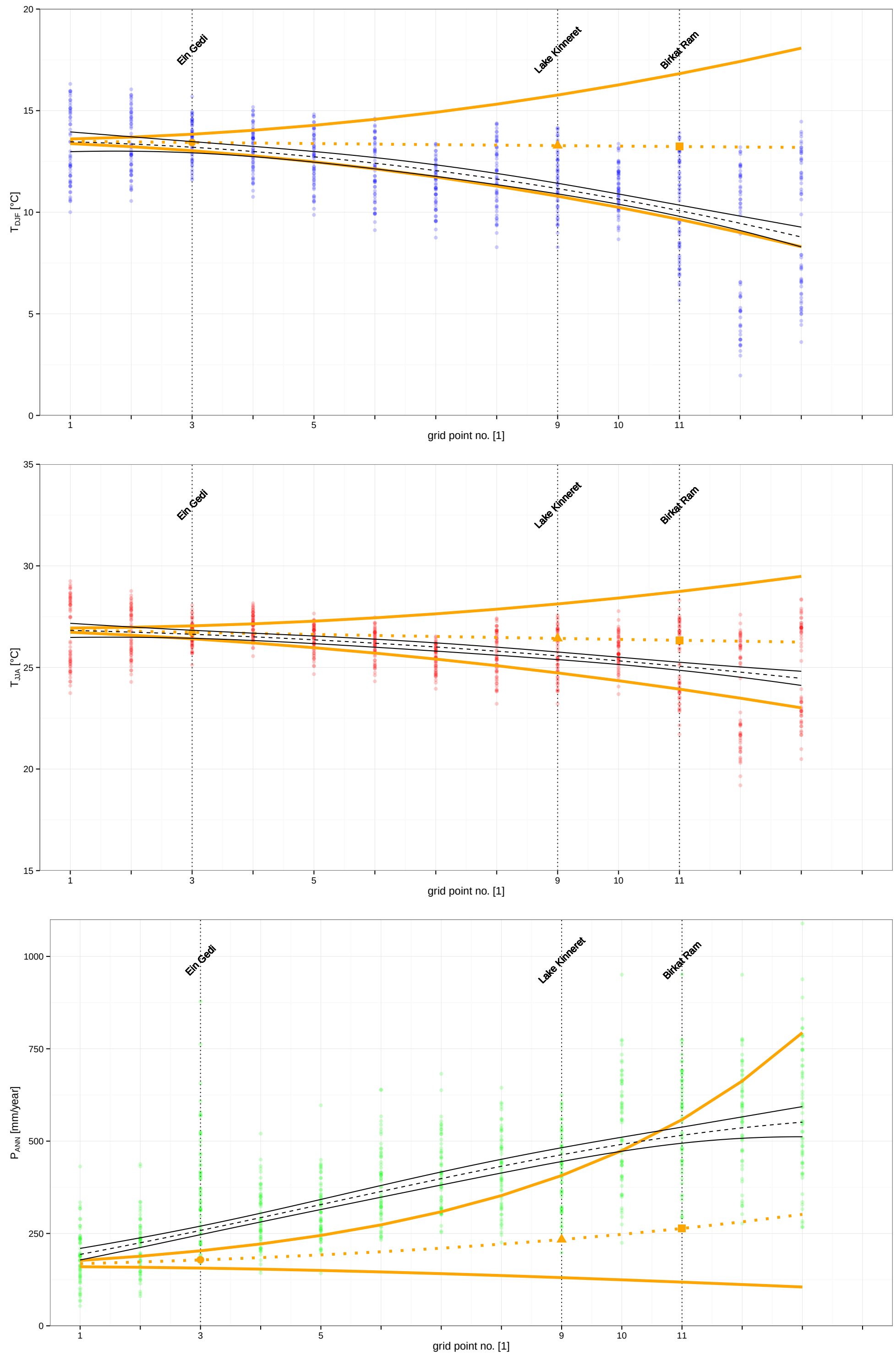


Figure E.6.: The Jordan Valley profile reconstruction result for time slice 1200 y. cal BP and the climate variables $T_{DJF\,h_{mean}}$, $T_{JJA\,h_{mean}}$ and $P_{ANN\,h_{mean}}$. The mapping of the figures is defined analogously to Fig. 6.5 and in detail repeated in the beginning of chapter E.3.

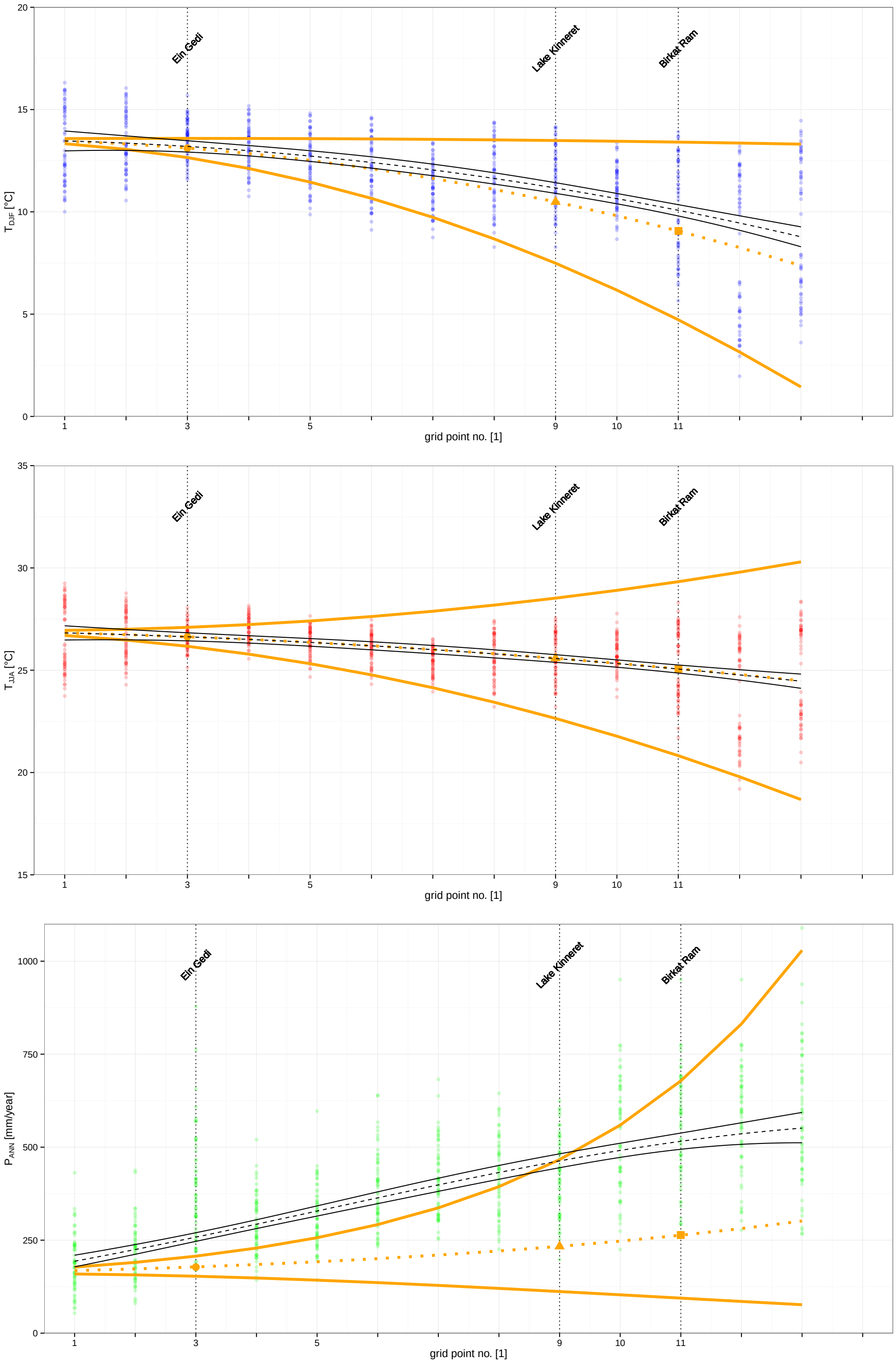


Figure E.7.: The Jordan Valley profile reconstruction result for time slice 1790 y. cal BP and the climate variables $T_{DJF\,h_{mean}}$, $T_{JJA\,h_{mean}}$ and $P_{ANN\,h_{mean}}$. The mapping of the figures is defined analogously to Fig. 6.5 and in detail repeated in the beginning of chapter E.3.

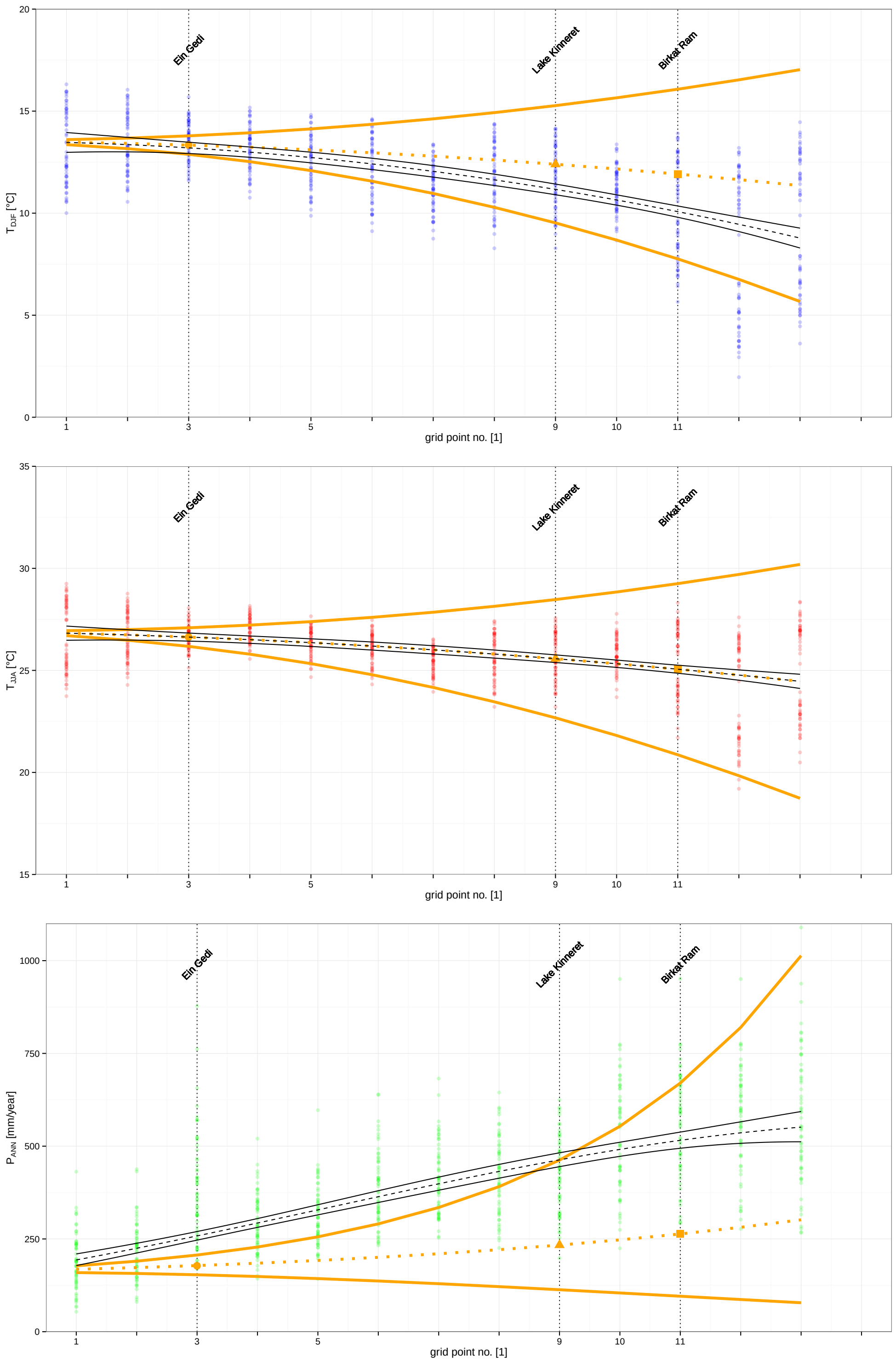


Figure E.8.: The Jordan Valley profile reconstruction result for time slice 1970 y. cal BP and the climate variables $T_{DJF\,h_{mean}}$, $T_{JJA\,h_{mean}}$ and $P_{ANN\,h_{mean}}$. The mapping of the figures is defined analogously to Fig. 6.5 and in detail repeated in the beginning of chapter E.3.

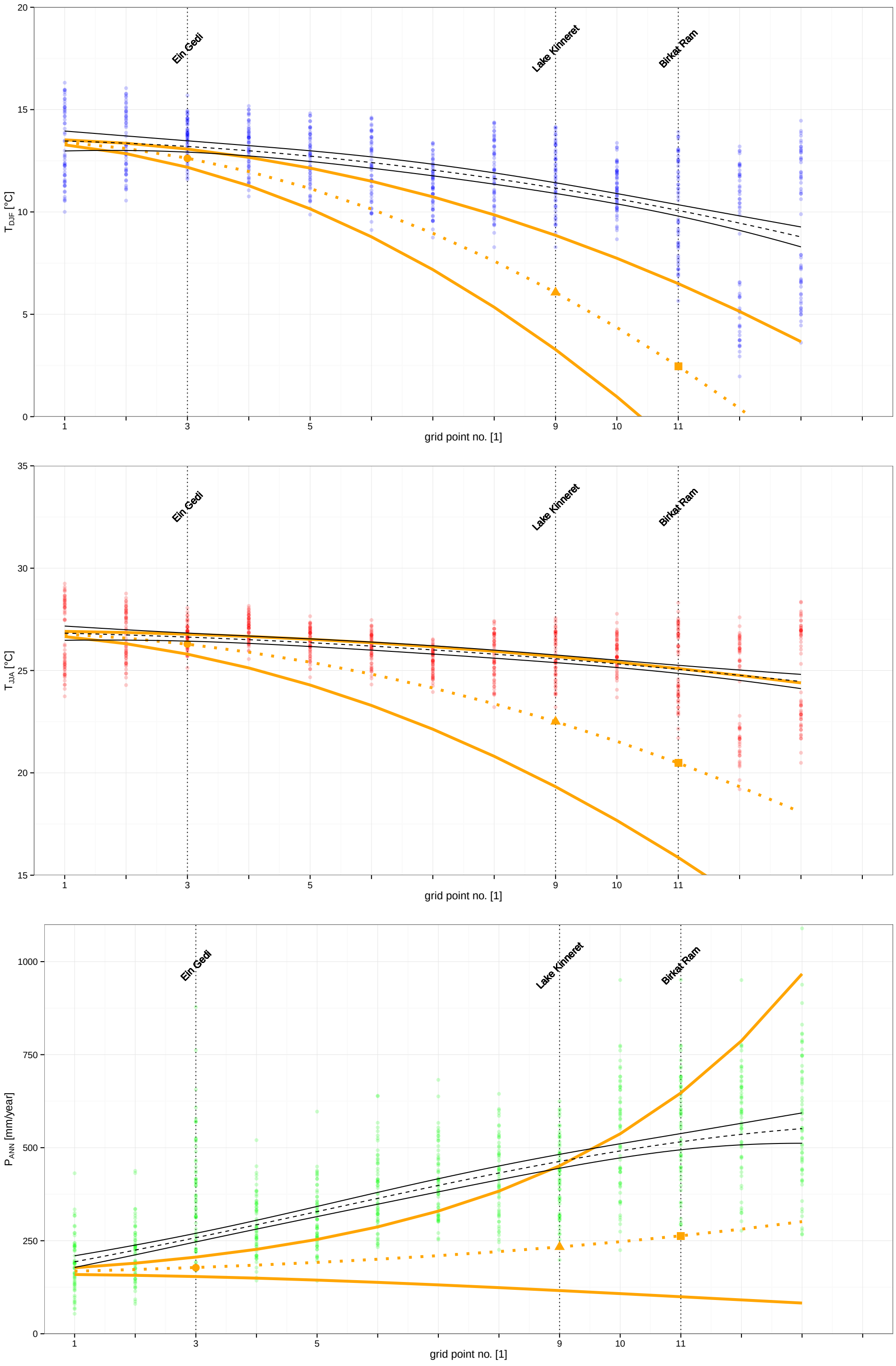


Figure E.9.: The Jordan Valley profile reconstruction result for time slice 6740 y. cal BP and the climate variables $T_{DJF\,h_{mean}}$, $T_{JJA\,h_{mean}}$ and $P_{ANN\,h_{mean}}$. The mapping of the figures is defined analogously to Fig. 6.5 and in detail repeated in the beginning of chapter E.3.

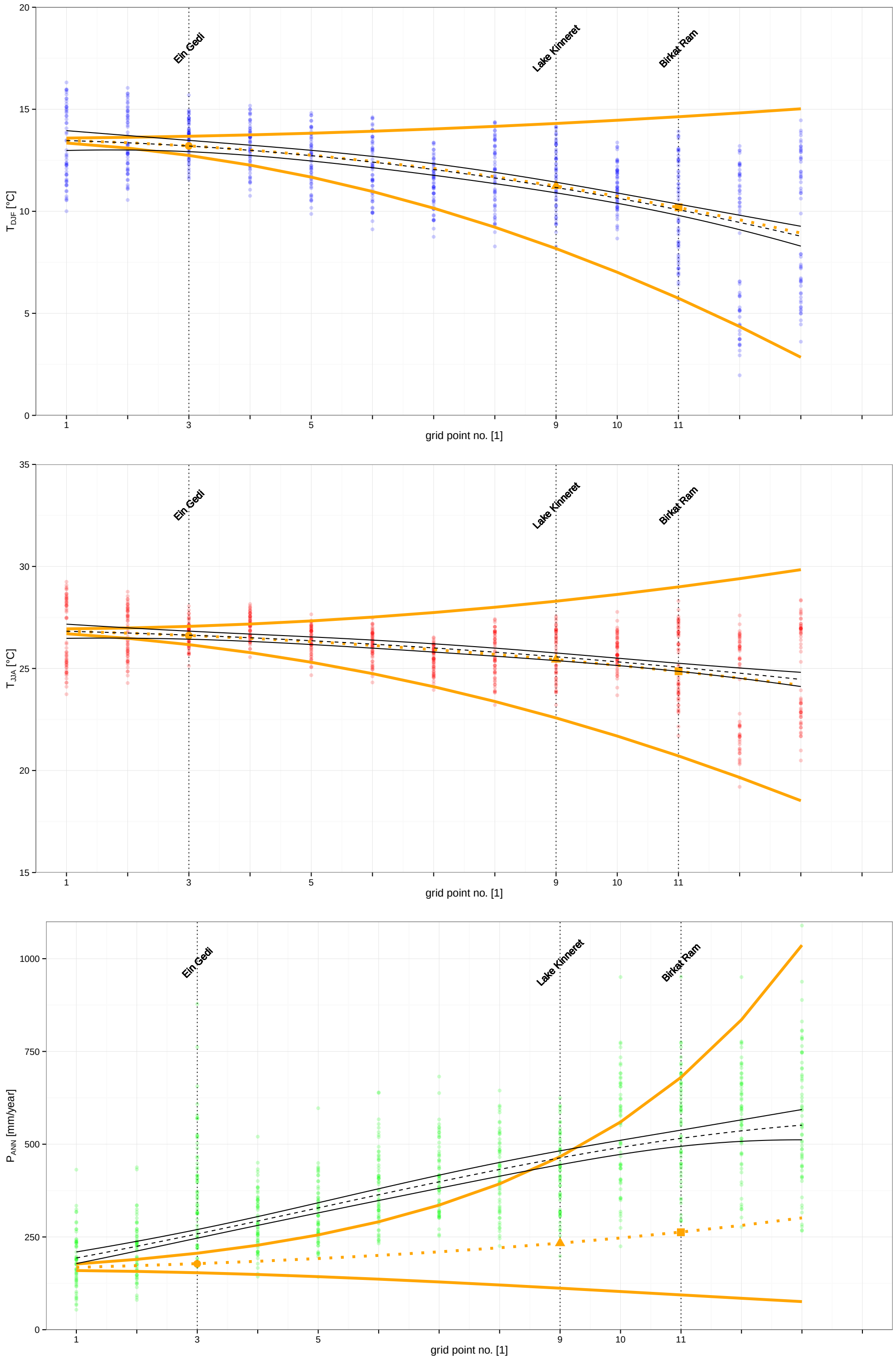


Figure E.10.: The Jordan Valley profile reconstruction result for time slice 7190 y. cal BP and the climate variables $T_{DJF\,h_{mean}}$, $T_{JJA\,h_{mean}}$ and $P_{ANN\,h_{mean}}$. The mapping of the figures is defined analogously to Fig. 6.5 and in detail repeated in the beginning of chapter E.3.

E.3.2. Time Series

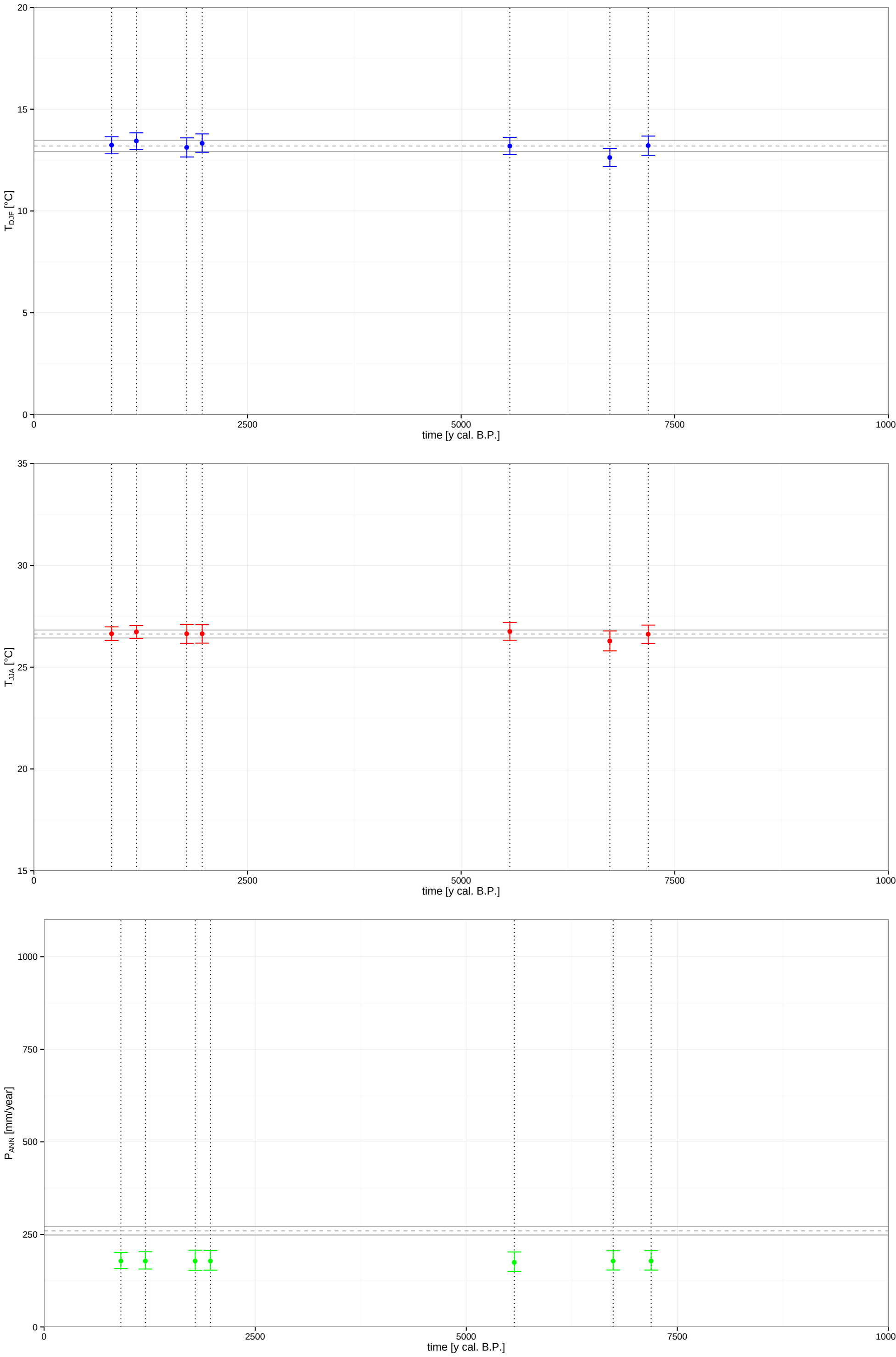


Figure E.11.: The palaeoclimate CFR results ${}^{EG}T_{DJF,R,0,h_{mean}}$, ${}^{EG}T_{JJA,R,0,h_{mean}}$ and ${}^{EG}P_{ANN\,R,0,h_{mean}}$ for Ein Gedi summarized as a time series. The mapping of the figures is defined analogously to Fig. 6.7 and in detail repeated in the beginning of chapter E.3.

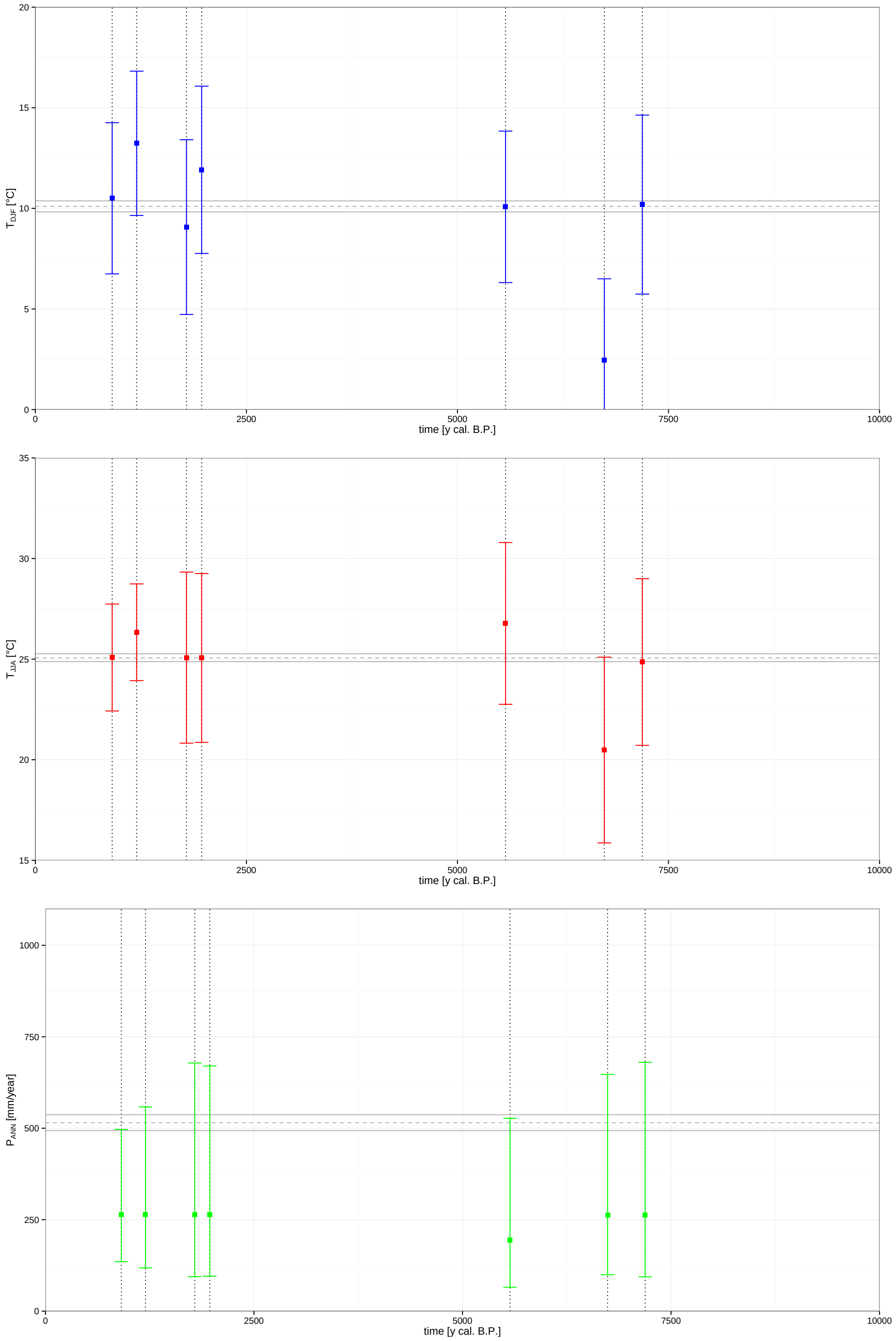


Figure E.12.: The palaeoclimate CFR results ${}^{BR}T_{DJF,R,0,h_{mean}}$, ${}^{BR}T_{JJA,R,0,h_{mean}}$ and ${}^{BR}P_{ANN\,R,0,h_{mean}}$ for Birkat Ram summarized as a time series. The mapping of the figures is defined analogously to Fig. 6.7 and in detail repeated in the beginning of chapter E.3.

Table E.2.: The Jordan Valley CFR result for all climate variables ($c_{i,R,0,h_{mean}}$) and for all considered time slices and fossil sites s. The first row in this table contains ${}^{s}c_{i,h_{mean}}$ the values for the modern time slice GLM and the other rows the minimization result $c_{i,R,0,h_{mean}}$ of the cost function $\mathcal{J}$. The errors of the modern time slice represent the 99.5 % confidence interval (s. chap. 6.1 for details). The other errors base on equation 6.25 and represent the $68.3\% = \pm\sigma_{u_i} = \pm\sigma_{...}$ confidence interval (s. chap. 6.5.1 for details).

No.	time slice [y. cal BP]	Ein Gedi			Lake Kinneret			Birkat Ram		
		$T_{DJF} \pm \sigma_{T_{DJF}}$ [°C]	$T_{JJA} \pm \sigma_{T_{JJA}}$ [°C]	$P_{ANN} \pm \sigma_{P_{ANN}}$ $\left[\frac{\text{mm}}{\text{year}}\right]$	$T_{DJF} \pm \sigma_{T_{DJF}}$ [°C]	$T_{JJA} \pm \sigma_{T_{JJA}}$ [°C]	$P_{ANN} \pm \sigma_{P_{ANN}}$ $\left[\frac{\text{mm}}{\text{year}}\right]$	$T_{DJF} \pm \sigma_{T_{DJF}}$ [°C]	$T_{JJA} \pm \sigma_{T_{JJA}}$ [°C]	$P_{ANN} \pm \sigma_{P_{ANN}}$ $\left[\frac{\text{mm}}{\text{year}}\right]$
0	0	13.2 ± 0.3	26.6 ± 0.2	$259.5\pm^{11.8}_{11.3}$	11.1 ± 0.3	25.6 ± 0.2	$464.3\pm^{19.0}_{18.3}$	10.1 ± 0.3	25.1 ± 0.2	$514.9\pm^{22.0}_{21.2}$
1	910	13.2 ± 0.4	26.6 ± 0.3	$178.5\pm^{22.9}_{20.5}$	11.4 ± 2.6	25.6 ± 1.9	$233.7\pm^{141.9}_{91.1}$	10.5 ± 3.8	25.1 ± 2.7	$263.7\pm^{232.9}_{128.6}$
2	1200	13.4 ± 0.4	26.7 ± 0.3	$178.3\pm^{24.8}_{22.0}$	13.3 ± 2.5	26.4 ± 1.7	$233.5\pm^{173.3}_{103.0}$	13.2 ± 3.6	26.3 ± 2.4	$263.5\pm^{294.5}_{145.2}$
3	1790	13.1 ± 0.5	26.6 ± 0.5	$178.2\pm^{28.9}_{25.1}$	10.5 ± 3.0	25.6 ± 2.9	$233.4\pm^{233.2}_{121.5}$	9.1 ± 4.3	25.1 ± 4.3	$263.3\pm^{414.9}_{169.1}$
4	1970	13.3 ± 0.5	26.6 ± 0.5	$178.2\pm^{28.4}_{24.8}$	12.4 ± 2.9	25.6 ± 2.9	$233.4\pm^{229.0}_{120.3}$	11.9 ± 4.2	25.1 ± 4.2	$263.4\pm^{406.7}_{167.8}$
5	5570	13.2 ± 0.4	26.8 ± 0.4	$174.1\pm^{28.2}_{24.5}$	11.2 ± 2.6	26.7 ± 2.8	$189.9\pm^{201.5}_{101.8}$	10.1 ± 3.8	26.8 ± 4.0	$193.8\pm^{333.2}_{128.6}$
6	6740	12.6 ± 0.4	26.3 ± 0.5	$178.1\pm^{27.8}_{24.3}$	6.1 ± 2.8	22.5 ± 3.2	$233.3\pm^{217.9}_{117.2}$	2.5 ± 4.0	20.5 ± 4.6	$263.2\pm^{383.8}_{163.7}$
7	7190	13.2 ± 0.5	26.6 ± 0.4	$178.1\pm^{28.1}_{24.5}$	11.2 ± 3.1	25.4 ± 2.9	$233.3\pm^{233.0}_{121.4}$	10.2 ± 4.4	24.9 ± 4.1	$263.2\pm^{417.0}_{169.4}$

Registers

List of Figures and Tables

List of Figures

List of Tables

Acronyms

AMSL	above mean sea level
AR	autoregressive
AR5	Fifth Assessment Report
B2	Collaborative Research Centre 806- subproject B2
B3	Collaborative Research Centre 806- subproject B3
BADC	British Atmospheric Data Centre
BBM	Bayesian Biome Model
BHM	Bayesian Hierarchical Model
BITBM	Bayesian Indicator Taxa and Biome Model
BITM	Bayesian Indicator Taxa Model
BMSL	below mean sea level
CCA	canonical correlation analysis
CCSM1v4	Community Climate System Model, version 1.4
CCSM3	Community Climate System Model, version 3
CDF	cumulative distribution function
CFR	climate field reconstruction
Co1215	Lake Megali Prespa composite sediment core or shortened Lake Prespa composite sediment core
CRC	Collaborative Research Centre
CRU	Climate Research Unit
CRU TS	CRU Time Series
CWD	climatic water deficit
CWD_{ANN}	annual climatic water deficit
DFG	German Research Foundation (German: Deutsche Forschungsgemeinschaft)
DSDDP	Dead Sea Deep Drilling Project
ECA&D	European Climate Assessment & Dataset project
ECDF	empirical cumulative distribution function

ECMWF	European Centre for Medium-Range Weather Forecasts
EM	expectation maximization
E-OBS	The ENSEMBLES Observations gridded dataset (E-OBS) provided by the ECA&D (*Data Policy for ECA&D and E-OBS* 2016).
EOF	empirical orthogonal function
ERA	ECMWF Reanalysis
ERA-Interim	An ERA dataset from January 1989 onward.
GCM	general circulation model
GDD5	growing degree days above 5 °C
GLM	generalized linear model
ICDP	International Continental Scientific Drilling Program
iid	independent and identically distributed
IPCC	Intergovernmental Panel on Climate Change
L-BFGS-B	limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm for box-constrained optimization
LOVECLIM1.2	dynamical climate system model with the possibility to include proxy information via data assimilation
LPAZ	local pollen assemblage zones (LPAZ), Lake Kinneret and Birkat-Ram, s. Schiebel (2013)
MAT	modern analogue technique
MCMC	Markov Chain Monte Carlo
MCR	mutual climatic range
MLE	maximum likelihood estimation
MTCO	mean temperature of the coldest month
NCAR	National Center for Atmospheric Research
NGDC	National Geophysical Data Center
P	precipitation amount
PalMod	Paleo Modelling
P_{ANN}	annual precipitation amount
PAZ	pollen assemblage zones (PAZ), Lake Prespa, s. Panagiotopoulos *et al.* (2013)

PCS	probability mutual climatic sphere
pdf	probability density function
PET	potential evapotranspiration
PET_{ANN}	annual potential evapotranspiration
PFT	plant functional types
PMIP3	Paleoclimate Modelling Intercomparison Project Phase III
PRM	pollen-ratio model
QDA	quadratic discriminant analysis
QR	quantile regression
s.no.	setup number for the verification of the biome likelihood
$SPEI$	Standardized Precipitation Evapotranspiration Index
T_{DJF}^{850hpa}	winter temperature at the 850 hPa-level, the middle of the troposphere
T_{DJF}	winter temperature
T_{Jan}	January temperature
T_{JJA}^{850hpa}	summer temperature at the 850 hPa-level, the middle of the troposphere
T_{JJA}	summer temperature
T_{Jul}	July temperature
T_{ANN}^{mean}	annual mean temperature
T_{JA}^{mean}	mean temperature of July and August
WG3	Paleo Modelling project WG3
y. BP	years before present
y. cal BP	calibrated years before present
YD	Younger Dryas

List of Symbols

Scalar Quantities

$a\left(\cdot\right)$	natural parameter in the definition of exponential family of distributions
$a_{\iota\, l_0}$	affinity score entry of $\boldsymbol{\mathcal{A}}^{N_\iota \times N_{l_0}}$
α_j	parameter of the generalized linear model (GLM) and element of the parameter vector $\vec{\alpha}$
$B_{\dots}$	shortcut for biomes
$b\left(\cdot\right)$	natural parameter in the definition of exponential family of distributions
Bern	Bernoulli distribution
$\beta_{\dots}$	coefficients of the GLM estimated for the modern Jordan Valley climate state
Bin	Binomial distribution
$\mathcal{B}_{l_0}$	vector component of the random variable for Biomes $\vec{\mathcal{B}}$
BS_{B_l}	Brier score for each biome
$BS_{ref}^{RF_{B_l}}$	reference Brier score calculated with the relative frequency of each biome
BS_{ref}^{ZV}	reference Brier score calculated with the reference probability ZV
$BSS_{B_l}^{RF}$	Brier skill score calculated with the reference score $BS_{ref}^{RF_{B_l}}$
$BSS_{B_l}^{ZV}$	Brier skill score calculated with the reference score BS_{ref}^{ZV}
$c\left(\cdot\right)$	specific parameter in the definition of exponential family of distributions
$d_{\cdot}\left(\cdot\right)$	discriminant function
$d\left(\cdot\right)$	specific parameter in the definition of exponential family of distributions
$\mathcal{D}$	mahalanobis distance
$\delta\left(\cdot\right)$	Dirac's delta function
$\Delta_{\mathcal{D}}$	threshold mahalanobis distance for filtering

δ_{ij}	Kronecker delta, a function which returns only one if $i \equiv j$ and zero otherwise
${}^{s}\Delta T$	value of the temperature height correction at the location of fossil site $\vec{r}_s$
$E_i(l)$	Brier score function that returns one if biome B_l is present at grid point i and zero otherwise
$\mathrm{E}(\cdot)$	expectation value
${}^{n}\varepsilon_{\dots}$	error of the estimated GLM at a grid box in the Jordan Valley
$f(\cdot)$	general symbol for a probability density function
$\mathcal{G}$	Gamma distribution
g	link function
G	a general additional condition
γ_d	temperature gradient of the standard atmosphere
γ_M	weighting factor in the cost function
h_{mean}	mean grid box height above sea level
h_{site}	height of the fossil site above sea level
i	general counting index
ι	pollen sample number $\iota = 1, \dots N_\iota$
$Ind(\cdot)$	index shift function for taxa
$\mathcal{J}$	cost function
j	general counting index
$\mathcal{J}_{Mod}$	model cost function
$\mathcal{J}_R$	value of the cost function at the minimum
$\mathcal{J}_{veg}$	vegetational cost function
$\mathcal{K}$	sample space for the recent climate variable or more general environmental variable
$\widetilde{k}$	new counting index for taxa
k	general counting index for taxa, $k = 1, \dots N_k$

$\mathcal{K}_0$	sample space for the past climate variable or more general environmental variable
$\mathscr{L}$	Likelihood
l	counting index for biomes
ℓ	log-likelihood
λ	counting index for longitude
$\lambda^{prior}_{\dots}$	the rate, a parameter of the Gamma distribution which defines the climate prior distribution
$m_{\dots}(\cdot)$	marginal distributions
$m_{k_0 l_0}$	biome assignment entry of $\boldsymbol{\mathcal{M}}^{N_{k_0}\times N_{l_0}}$
μ	the mean, a parameter of the Normal distribution
$\mu^{prior}_{\dots}$	the mean, a parameter of the Normal distribution which defines the climate prior distribution
n	grid box number in the Jordan Valley
$\mathcal{N}$	Normal distribution
${}^{n}\dots = \dots(\vec{r}_n)$	index notation which reflects the evaluation of a quantity at the Jordan Valley location $\vec{r}_n$
N_l	number of different biomes, assigned to l
$N_{B_l\,obs}$	number of grid points with presence information of biome type B_l
N_{GB}	number or grid boxes in the Jordan Valley
N_ι	number of samples
N_i	a general number for summation or multiplication assigned to the general counting index i
N_j	a general number assigned to the general counting index j
$N_{\widetilde{k}}$	new number of taxa, assigned to $\widetilde{k}$
N_k	number of taxa assigned to k
$N_{\lambda\varphi}$	number of grid points

$\mathcal{N}_{MV}$	symbol for the multivariate normal distribution with the parameters $\vec{\mu}$ and $\mathbf{\Sigma}$
N_{obs}	number of grid points with presence information of any biome type B_l, so that $N_{obs} = \sum_{l=1}^{N_l} N_{B_l\,obs}$
$\nu_{\ldots}^{prior}$	the shape, a parameter of the Gamma distribution which defines the climate prior distribution
$\widetilde{\omega}_{\iota\, k_0}$	relative abundance of taxon k_0 in pollen sample number ι
$\mathbb{P}\left(\cdot\right)$	general symbol for probability
p	probability of success, Binomial and Bernoulli distribution
$\mathcal{P}$	sample space for the recent proxy variable
$\mathcal{P}_0$	sample space for the past proxy variable
${}^{n}P_{ANN}$	annual precipitation amount at a grid box in the Jordan Valley
φ	counting index for latitude
ϕ_s	latitude of the northern resp. southern grid box boundary
$\pi_{\ldots}\left(\cdot\right)$	prior distributions
$q\left(\cdot\right)$	function which is a general condition to extract the past and the modern climate
$\mathbb{R}$	set of real numbers
RF	shortcut for relative frequency of each biome in the definition of the Brier skill score
s	counting index fossil sites
$\mathcal{S}$	general symbol for the sample space
S	number of fossil sites s
${}^{s}\ldots = \ldots\left(\vec{r}_s\right)$	index notation which reflects the evaluation of a quantity at the location of the fossil site $\vec{r}_s$
${}^{s}\cdots_{h_{\ldots}}^{\ldots}$	index notation which reflects the evaluation of a quantity at the location of fossil site $\vec{r}_s$ at height $h_{\ldots}$
σ	the standard deviation, a parameter of the Normal distribution

$\sigma_{\dots}^{prior}$	the standard deviation, a parameter of the Normal distribution which defines the climate prior distribution
${}^{n}T_{DJF}$	winter temperature at a grid box in the Jordan Valley
ϑ_{k_0}	threshold for each taxon
$\mathcal{V}_{\Theta}$	sample space for the parameters
ϑ_{k_0}	threshold value
$t_{\iota\, k_0}$	past proxy realisation of taxon k_0 in pollen sample number ι
${}^{n}T_{JJA}$	summer temperature at a grid box in the Jordan Valley
$to\ell\,(\cdot)$	tolerance function
$u_{\dots}$	component of transferred climate state vector $\vec{u}$
$W_{B_{l_0}}$	biome probability
X	explanatory random variable
x_{ij}	explanatory variable and element of the design Matrix $\boldsymbol{X}$
y	realisation of responses
Y	random variable for the response variable
ZV	shortcut for the reference probability in the definition of the Brier skill score
z	general symbol for a realisation
Z	general symbol for a random variable

Vectorial Quantities

$\vec{\alpha}$	parameter vector of GLM
$\vec{\beta}$	coefficient vector which defines the Jordan Valley GLM for the modern climate state and summarizes all $\beta_{\dots}$
$\vec{\beta}_0$	coefficient vector which defines the Jordan Valley GLM for the past climate state and summarizes all $\beta_{0\dots}$
$\vec{\beta}\,'$	coefficient vector which defines the Jordan Valley GLM deviation and summarizes all $\beta'_{\dots}$

$\vec{\beta}_{R,0}$	coefficient vector for the Jordan Valley GLM at the minimum of the cost function
$\vec{c}$	recent climate or more general environmental realisation
$\vec{C}$	recent climate ore more general environmental random variable
$\vec{c}_0$	past climate or more general environmental realisation
$\vec{C}_0$	past climate ore more general environmental random variable
$\vec{c}\,'$	climate anomaly between past and modern climate state
$\vec{e}_{\beta_j}$	a column vector with the same dimension as $\vec{\beta}_0$ in which only the j-th component is one and the rest is zero
$\vec{e}_h$	vector for the definition of the height correction
$\vec{\varepsilon}$	vector which summarizes the errors ${}^n\varepsilon_{\dots}$
$\vec{\varepsilon}_0$	vector which summarizes the errors ${}^n\varepsilon_{0\dots}$
$\vec{\varepsilon}\,'$	vector which summarizes the errors ${}^n\varepsilon'_{\dots}$
$\vec{\mu}$	mean value vector, parameter of the multivariate Normal distribution
$\vec{n}$	location vector of the Jordan Valley described by the grid box numbers
$\vec{n^2}$	component wise squared of the location vector of the Jordan Valley described by the grid box numbers
$\vec{\nabla}_{\vec{\beta}_0}\mathcal{J}$	gradient of the cost function with respect to $\vec{\beta}_0$
$\vec{\tilde{\omega}}_{k_0}$	column vector of $\mathbf{\Omega}^{N_\iota \times N_{k_0}}$
$\vec{p}$	recent proxy realisation
$\vec{P}$	recent proxy random variable
$\vec{p}_0$	past proxy realisation
$\vec{P}_0$	past proxy random variable
$\vec{\varphi}_{prior}/\vec{\varphi}^{-1}_{prior}(\cdot)$	function for the transformation with the inverse CDF
$\vec{R}$	random variable for the recent time frame
$\vec{R}_0$	random variable for the past time frame

$\vec{r}_n$	Jordan Valley location vector
$\vec{r}_s$	location of fossil site s
$\vec{t}$	taxon proxy realisation
$\vec{T}$	taxon proxy random variable
$\vec{\theta}$	realisation of a set of parameters which describe the statistical model
$\vec{\Theta}$	random variable for a set of parameters which describe the statistical model
$\vec{\theta}_{opt}$	maximum likelihood estimated set of parameters which describe the statistical model
$\vec{u}$	with the inverse CDF method transferred climate state vector
$\vec{y}$	realisation of responses
$\vec{Y}$	multivariate random variable for the response variable

Matrices

$\boldsymbol{\mathcal{A}}^{N_\iota \times N_{l_0}}$	biome affinity score matrix
$\boldsymbol{\mathcal{H}}_{\vec{\beta}_0}(\mathcal{J})$	Hessian matrix of the cost function with respect to $\vec{\beta}_0$
$\boldsymbol{\mathcal{I}}_{\vartheta_0}^{N_\iota \times N_{k_0}}$	Indicator Taxa Matrix defined with threshold exceeding
$\boldsymbol{\mathcal{I}}_{diff}^{N_\iota \times N_{k_0}}$	difference between the threshold defined Indicator Taxa Matrix and the mahalanobis distance filtered Indicator Taxa Matrix
$\boldsymbol{\mathcal{I}}_{filtered}^{N_\iota \times N_{k_0}}$	Indicator Taxa Matrix filtered with the mahalanobis distance algorithm
$\boldsymbol{\mathcal{M}}^{N_{k_0} \times N_{l_0}}$	biome assignment matrix
$\boldsymbol{\Omega}^{N_\iota \times N_{k_0}}$	matrix which summarizes in one matrix all information of the sediment core, in detail the relative abundances $\widetilde{\omega}_{\iota k_0}$
$\boldsymbol{\Omega}_{max(k_0)}^{N_\iota \times N_{k_0}}$	maximum value matrix for each column vector $\vec{\widetilde{\omega}}_{k_0}$
${}^{s}\boldsymbol{P}$	interpolation matrix with the positions of the fossil sites
$\boldsymbol{R}$	analysis area matrix which summarizes the definition of the analysis area
$\boldsymbol{S}$	a shortcut defined by the matrix product ${}^{s}\boldsymbol{P}\boldsymbol{R}$
$\boldsymbol{\Sigma}$	covariance matrix, parameter of the multivariate Normal distribution

$\boldsymbol{T}$	Matrix for the definition of the tolerance function $\mathcal{tol}\left(\cdot\right)$ with the same properties as the design matrix $\boldsymbol{X}$
$\boldsymbol{\vartheta}^{N_t \times N_{k_0}}$	taxa threshold value matrix with entries ϑ_{k_0}
$\boldsymbol{X}$	design Matrix which defines the GLM

List of Taxa

Abies	Abies_Prm - *Lake Prespa*
Acacia	no taxon shortcut- *Ein Gedi* ⇒ Saharo-Arabian biome
Adonis type	Adoni_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Alisma cf. A. plantago-aquatica	Alpl_Prm - *Lake Prespa*
Alnus	Algl_m - *Lake Prespa*
Anemone type	Anemo_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Apiaceae	Apiac_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Arbutus	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome
Artemisia	Arte_BR - *Birkat Ram*, Artem_K - *Lake Kinneret* & *Ein Gedi* ⇒ Irano-Turanian biome
Betula	Betu_m - *Lake Prespa*
Brassicaceae	Brass_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Buxus	Buse_m - *Lake Prespa*
Campanulaceae	Campa_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Cannabis type	Canna_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Caryophyllaceae	Caryo_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Cedrus	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome
Centaurea	Centa_K - *Lake Kinneret* & *Ein Gedi* ⇒ Irano-Turanian biome

Ceratonia	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome
Cerealea type	Cerea_K - *Lake Kinneret* & *Ein Gedi* ⇒ Irano-Turanian biome
Chenopodiaceae	Cheno_K - *Lake Kinneret* & *Ein Gedi* ⇒ Irano-Turanian biome- *Lake Kinneret* ⇒ Saharo-Arabian biome- *Ein Gedi*
Chenopodium	Cheno_Pr - *Lake Prespa*
Cistus type	Cistu_K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome
Cornus mas	Coms_m - *Lake Prespa*
Corylus	Coav_m - *Lake Prespa*
Cupressaceae	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome
Elatine type	Elati_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Ephedra	Ephed_Pr - *Lake Prespa*, Ep_dt_K - *Ein Gedi* ⇒ Irano-Turanian biome
Ephedra dyst	Ep_dt_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Ephedra frag	Ep_fg_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Ericaceae	Erica_K - *Lake Kinneret* ⇒ Mediterranean biome
Fabaceae	Fabac_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Fagus	Fasy_f - *Lake Prespa*
Fraxinus	Frax_Prm - *Lake Prespa*
Fumana type	Fuman_K - *Lake Kinneret* ⇒ Mediterranean biome
Hedera	Hehe_m - *Lake Prespa*
Helianthemum	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome

Hippophae	Hirh_Prm - *Lake Prespa*
Hypericum type	Hyper_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Juglans	Jugla_K - *Lake Kinneret* ⇒ Mediterranean biome
Laurus	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome
Liguliflorae	Ligul_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Linum type	Linum_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Malvaceae	Malva_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Myrica	no taxon shortcut
Nymphaea	Nyal_m - *Lake Prespa*
Olea	Oleu_m - *Lake Prespa*, Oleu_BR - *Birkat Ram*, Olea__K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome
Papaver type	Papav_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Phillyrea	Phla_m - *Lake Prespa*, Phill_K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome
Phoenix	no taxon shortcut- *Ein Gedi* ⇒ Saharo-Arabian biome
Picea	Picea_Prmf - *Lake Prespa*
Pinus	Pinus_K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome

Pistacia	Pista_Pr - *Lake Prespa*, Pieu_BR - *Birkat Ram*, Pista_K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome
Plantago	PlSM_BR - *Birkat Ram*, Plant_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Poaceae	Poace_K - *Lake Kinneret* & *Ein Gedi* ⇒ Irano-Turanian biome
Polemoniaceae	Poca__Prm - *Lake Prespa*
Polygonaceae undiff.	Poav_Prh - *Lake Prespa*
Quercus calliprinos type	Que_BR - *Birkat Ram*, Qu_ca_K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome
Quercus cerris	Quce_Prm - *Lake Prespa*
Quercus ilex	Quil_m - *Lake Prespa*
Quercus indet	Queru_K - *Lake Kinneret* ⇒ Mediterranean biome
Quercus ithaburensis type	Qud_BR - *Birkat Ram*, Qu_it_K - *Lake Kinneret* & *Ein Gedi* ⇒ Mediterranean biome
Quercus robur	Quru_Prm - *Lake Prespa*
Rhamnus type	Rhamn_K - *Lake Kinneret* ⇒ Mediterranean biome
Rubiaceae	Rubia_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Rumex	Rumex_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Sambucus	Sani_m - *Lake Prespa*
Sanguisorba type	Sangu_K - *Lake Kinneret* ⇒ Mediterranean biome
Sarcopoterium	no taxon shortcut- *Ein Gedi* ⇒ Mediterranean biome
Scabiosa type	Scabi_K - *Lake Kinneret* ⇒ Irano-Turanian biome

Styrax type	Styra_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Tamarix	no taxon shortcut- *Ein Gedi* ⇒ Saharo-Arabian biome
Thalictrum type	Thali_K - *Lake Kinneret* ⇒ Irano-Turanian biome
Tilia	Tilia_Prm - *Lake Prespa*
Tubuliflorae	Tubul_K - *Lake Kinneret* & *Ein Gedi* ⇒ Irano-Turanian biome
Typha	Tyla_Prh - *Lake Prespa*
Ulmus	Ulmu_Prmf - *Lake Prespa*
Vitis	Vitis_K - *Lake Kinneret* ⇒ Mediterranean biome
Zizyphus	no taxon shortcut- *Ein Gedi* ⇒ Saharo-Arabian biome
Zygophylum	no taxon shortcut- *Ein Gedi* ⇒ Saharo-Arabian biome

Glossary

Climate Proxy

Indicators which are influenced essentially by climatic conditions. The nature of these could be for example biological, chemical or also historical reports.

Climatic Water Deficit

Describes the effective usable water for the plants (s. e. g. Crimmins *et al.* 2011)

Holocene

Time period in Earth history defined as the time period from approximate 10 000 years before present (y. BP) - present (s. *NOAA Palaeoclimatology* 2016).

Mid Holocene

Time period in Earth history: roughly from 7 000 y. BP - 5 000 y. BP (s. *NOAA Palaeoclimatology* 2016).

Middle and Late Holocene

Time period in Earth history defined in Panagiotopoulos *et al.* (2013) as 7 900 y. cal BP - present

Younger Dryas

Time period in Earth history defined in Panagiotopoulos *et al.* (2013) as 7 900 y. cal BP 13 200 y. cal BP - 11 500 y. cal BP

Bibliography

Bibliography

Scientific Books

Dobson, A.J., Barnett, A.G. (2008). *An Introduction to Generalized Linear Models*. Texts in statistical science. CRC Press.

Feinbrun-Dothan, Nqaomi, Danin, Avinoam (1998). *Analytical Flora of Eretz-Israel*. Ed. by Plitmann, Uzi. 2nd ed. Carta, The Israel Map and Publishing Company Limited, Israel.

Gebhardt, Christoph (2003). *Variational Reconstruction of Quaternary Temperature Fields using Mixture Models as Botanical-Climatological Transfer Functions*. Ed. by Hense, Andreas. 60. Asgard.

Gentle, J.E. (2003). *Random Number Generation and Monte Carlo Methods*. Statistics and Computing. Springer.

Meusel, H., Jäger, E. J., Weinert, E. (1965). *Vergleichende Chorologie der zentraleuropäischen Flora*. VEB Fischer, Jena.

Meusel, H., Jäger, E., Weinert, E (1974). *Vergleichende Chorologie der zentraleuropäischen Flora*. Vol. Bd. II, Karten.

Schölzel, Christian (2006). *Palaeoenvironmental transfer functions in a Bayesian framework with application to holocene climate variability in the Near East*. Ed. by Hense, Andreas. 62. Asgard.

Venables, W.N., Ripley, Brain D. (2002). *Modern applied statistics with S*. Springer verlag. DOI: `10.1007/978-0-387-21706-2`.

Wilks, Daniel.S. (2011). *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Academic Press, 676.

Parts of Scientific Collections

Hense, Andreas (2005). “Processing of observational data and its implication for climate analysis.” *Landolt-Börnstein - Group V Geophysics*. Vol. 6. SpringerMaterials - The Landolt-Börnstein Database. DOI: `10.1007/10730550_3`.

Weinstein, M. (1979). “Recent Pollen spectra.” *The Quaternary of Israel*. Ed. by Horowitz, A. Academic Press, New York, 180–210.

Scientific Articles

Andreev, A.A., Klimanov, V.A. (2000). Quantitative Holocene climatic reconstruction from Arctic Russia. English. *Journal of Paleolimnology* **24**:1, 81–91. DOI: `10.1023/A:1008121917521`.

Aufgebauer, Anne, Panagiotopoulos, Konstantinos, Wagner, Bernd, Schäbitz, Frank, Viehberg, Finn A., Vogel, Hendrik, Zanchetta, Giovanni, Sulpizio, Roberto, Leng, Melanie J., Damaschke, Magret (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**: 122–135. DOI: `10.1016/j.quaint.2012.02.015`.

Bartlein, Patrick J., Harrison, S.P., Brewer, S., Connor, S., Davis, B.A.S., Gajewski, K., Guiot, J., Harrison-Prentice, T.I., Henderson, A., Peyron, O., Prentice, I. Colin, Scholze, M., Seppä, H., Shuman, B., Sugita, S., Thompson, R.S., Viau, A.E., Williams, J., Wu, H. (2011). Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. English. *Climate Dynamics* **37**:3-4, 775–802. DOI: `10.1007/s00382-010-0904-1`.

Bartlein, Patrick J., Webb III, Thompson, Fleri, E. (1984). Holocene climatic change in the northern Midwest: Pollen-derived estimates. *Quaternary Research* **22**:3, 361–374. DOI: `10.1016/0033-5894(84)90029-2`.

Bayes, Thomas (1763). An Essay towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions* **53**: 370–418. DOI: `10.1098/rstl.1763.0053`.

Bernabo, J.Christopher (1981). Quantitative estimates of temperature changes over the last 2700 Dates in Michigan based on pollen data. *Quaternary Research* **15**:2, 143–159. DOI: `10.1016/0033-5894(81)90101-0`.

Birks, H. John B. (1998). D. G. Frey and E. S. Deevey Review 1: Numerical tools in palaeolimnology - Progress, potentialities, and problems. English. *Journal of Paleolimnology* **20**:4, 307–332. DOI: `10.1023/A:1008038808690`.

Birks, H. John B., Heiri, Oliver, Seppa, Heikki, Bjune, Anne E. (2010). Strengths and weaknesses of quantitative climate reconstructions based on late-Quaternary biological proxies. *Open Ecology Journal* **3**: 68–110.

Brier, Glenn W. (1950). Verification of Forecast Expressed in Terms of Probability. *Monthly Weather Review* **78**:1, 1–3. DOI: `10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2`.

Byrd, Richard H., Lu, Peihuang, Nocedal, Jorge, Zhu, Ciyou (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **16**:5, 1190–1208.

Cheddadi, R., Lamb, H. F., Guiot, J., Kaars, S. van der (1998). Holocene climatic change in Morocco: a quantitative reconstruction from pollen data. English. *Climate Dynamics* **14**:12, 883–890. DOI: `10.1007/s003820050262`.

Cook, Edward R., Briffa, Keith R., Jones, Philip D. (1994). Spatial regression methods in dendroclimatology: A review and comparison of two techniques. *International Journal of Climatology* **14**:4, 379–402. DOI: `10.1002/joc.3370140404`.

Crimmins, Shawn M., Dobrowski, Solomon Z., Greenberg, Jonathan A., Abatzoglou, John T., Mynsberge, Alison R. (2011). Changes in Climatic Water Balance Drive Downhill Shifts in Plant Species Optimum Elevations. *Science* **331**:6015, 324–327. DOI: `10.1126/science.1199040`.

Crucifix, Michel (2012). Traditional and novel approaches to palaeoclimate modelling. *Quaternary Science Reviews* **57**: online: 2012-10-18, 1–16. DOI: `10.1016/j.quascirev.2012.09.010`.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., Berg, L. van de, Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P. de, Tavolato, C., Thépaut, J.-N., Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**:656, 553–597. DOI: `10.1002/qj.828`.

Garreta, Vincent, Guiot, J., Mortier, F., Chadœuf, J., Hély, C. (2012). Pollen-based climate reconstruction: Calibration of the vegetation–pollen processes. *Ecological Modelling* **235–236**: online: 2012-05-04, 81–94. DOI: `10.1016/j.ecolmodel.2012.03.031`.

Garreta, Vincent, Miller, PaulA., Guiot, Joël, Hély, Christelle, Brewer, Simon, Sykes, MartinT., Litt, Thomas (2010). A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model. English. *Climate Dynamics* **35**:2–3, 371–389. DOI: `10.1007/s00382-009-0629-1`.

Gauthier, Pierre (1992). Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus A* **44**:1, 2–17. DOI: `10.1034/j.1600-0870.1992.00002.x`.

Gebhardt, Christoph, Kühl, Norbert, Hense, Andreas, Litt, Thomas (2008). Reconstruction of Quaternary temperature fields by dynamically consistent smoothing. *Climate Dynamics* **30**:4 (4), 421–437. DOI: `10.1007/s00382-007-0299-9`.

Guillot, Dominique, Rajaratnam, Bala, Emile-Geay, Julien (2013). Statistical paleoclimate reconstructions via Markov random fields. *ArXiv e-prints*, 23.

Harris, Ian, Jones, Philip D., Osborn, T.J., Lister, D.H. (2014). Updated high-resolution grids of monthly climatic observations. *International Journal of Climatology* **34**:3, 623–642. DOI: `10.1002/joc.3711`.

Haslett, J., Whiley, M., Bhattacharya, S., Salter-Townshend, M., Wilson, Simon P., Allen, J. R. M., Huntley, B., Mitchell, F. J. G. (2006). Bayesian palaeoclimate reconstruction. *Journal of the*

Royal Statistical Society: Series A (Statistics in Society) **169**:3, 395–438. DOI: `10.1111/j.1467-985X.2006.00429.x`.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research: Atmospheres* **113**:D20, n/a–n/a. DOI: `10.1029/2008JD010201`.

Hense, Andreas (2002). Klimavariabilität durch interne Wechselwirkungen. *PROMET* **28**:3, 108–116.

Hintikka, V. (1963). Über das Großklima einiger Pflanzenareale in zwei Klimakoordinatensystemen dargestellt. *Annales Botanici Societatis Zoologicae Botanicae Fennicae*, 1–64.

Howe, Sally, Webb III, Thompson (1983). Calibrating pollen data in climatic terms: Improving the methods. *Quaternary Science Reviews* **2**:1, 17–51. DOI: `10.1016/0277-3791(83)90003-3`.

Hughes, M. K., Ammann, Caspar M. (2009). The future of the past - an earth system framework for high resolution paleoclimatology: editorial essay. *Climatic Change* **94**:3-4, 247–259. DOI: `10.1007/s10584-009-9588-0`.

Iversen, Johs (1944). Viscum, Hedera and Ilex as Climate Indicators: A Contribution to the Study of the Post-Glacial Temperature Climate. *Geologiska Foereningens Stockholm Foerhandlingar* **66**:3, 463–483. DOI: `10.1080/11035894409445689`.

Janson, Lucas, Rajaratnam, Bala (2014). A Methodology for Robust Multiproxy Paleoclimate Reconstructions and Modeling of Temperature Conditional Quantiles. *Journal of the American Statistical Association* **109**:505, 63–77. DOI: `10.1080/01621459.2013.848807`.

Klotz, Stefan, Müller, Ulrich, Mosbrugger, Volker, Beaulieu, Jacques-Louis de, Reille, Maurice (2004). Eemian to early Würmian climate dynamics: history and pattern of changes in Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology* **211**:1-2, 107–126. DOI: `10.1016/j.palaeo.2004.04.009`.

Korhola, Atte, Vasko, Kari, Toivonen, Hannu T.T., Olander, Heikki (2002). Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modelling. *Quaternary Science Reviews* **21**:16–17, 1841–1860. DOI: `10.1016/S0277-3791(02)00003-3`.

Kühl, Norbert, Gebhardt, Christoph, Litt, Thomas, Hense, Andreas (2002). Probability Density Functions as Botanical-Climatological Transfer Functions for Climate Reconstruction. *Quaternary Research* **58**:3, 381–392. DOI: `10.1006/qres.2002.2380`.

Kühl, Norbert, Litt, Thomas, Schölzel, Christian, Hense, Andreas (2007). Eemian and Early Weichselian temperature and precipitation variability in northern Germany. *Quaternary Science Reviews* **26**:25–28, 3311 –3317. DOI: `10.1016/j.quascirev.2007.10.004`.

Li, Bo, Nychka, Douglas W., Ammann, Caspar M. (2010). The Value of Multiproxy Reconstruction of Past Climate. *Journal of the American Statistical Association* **105**:491, 883–895. DOI: `10.1198/jasa.2010.ap09379`.

Litt, Thomas, Ohlwein, Christian, Neumann, Frank, Hense, Andreas, Stein, Mordechai (2012). Holocene climate variability in the Levant from the Dead Sea pollen record. *Quaternary Science Reviews* **49**: 95–105. DOI: `10.1016/j.quascirev.2012.06.012`.

Litt, Thomas, Schölzel, Christian, Kühl, Norbert, Brauer, Achim (2009). Vegetation and climate history in the Westeifel Volcanic Field (Germany) during the past 11 000 Dates based on annually laminated lacustrine maar sediments. *Boreas* **38**:4, 679–690. DOI: `10.1111/j.1502-3885.2009.00096.x`.

Mahalanobis, Prasanta Chandra (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India* **12**: 49–55.

Mairesse, A., Goosse, H., Mathiot, P., Wanner, H., Dubinkina, S. (2013). Investigating the consistency between proxy-based reconstructions and climate models using data assimilation: a mid-Holocene case study. *Climate of the Past* **9**:6, 2741–2757. DOI: `10.5194/cp-9-2741-2013`.

Masson-Delmotte, Valérie, Schulz, Michael, Abe-Ouchi, Ayako, Beer, Juerg, Ganopolski, Andrey, González Rouco, Jesus Fidel, Jansen, Eystein, Lambeck, Kurt, Luterbacher, Jürg, Naish, Tim, Osborn, Timothy, Otto-Bliesner, Bette, Quinn, Terrence, Ramesh, Rengaswamy, Rojas, Maisa, Shao, XueMei, Timmermann, Axel (2013). Information from Paleoclimate Archives. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., 383–464.

McShane, Blakeley B., Wyner, Abraham J. (2011). A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable? *The Annals of Applied Statistics* **5**:1, 5–44. DOI: `10.1214/10-AOAS398`.

Migowski, Claudia, Agnon, Amotz, Bookman, Revital, Negendank, Jörg F. W., Stein, Mordechai (2004). Recurrence pattern of Holocene earthquakes along the Dead Sea transform revealed by varve-counting and radiocarbon dating of lacustrine sediments. *Earth and Planetary Science Letters* **222**:1, 301–314. DOI: `10.1016/j.epsl.2004.02.015`.

Migowski, Claudia, Stein, Mordechai, Prasad, Sushma, Negendank, Jörg F. W., Agnon, Amotz (2006). Holocene climate variability and cultural evolution in the Near East from the Dead Sea sedimentary record. *Quaternary Research* **66**:3, 421–431. DOI: `10.1016/j.yqres.2006.06.010`.

Mukhopadhyay, Sabyasachi, Bhattacharya, Sourabh (2013). An Improved Bayesian Semiparametric Model for Palaeoclimate Reconstruction: Cross-validation Based Model Assessment. *ArXiv e-prints*. file: 2013-12-12, 48.

Müller, Ulrich C., Pross, Jörg, Tzedakis, Polychronis C., Gamble, Clive, Kotthoff, Ulrich, Schmiedl, Gerhard, Wulf, Sabine, Christanis, Kimon (2011). The role of climate in the spread of modern humans into Europe. *Quaternary Science Reviews* **30**:3-4, 273–279. DOI: `10.1016/j.quascirev.2010.11.016`.

Neumann, Frank, Schölzel, Christian, Litt, Thomas, Hense, Andreas, Stein, Mordechai (2007). Holocene vegetation and climate history of the northern Golan heights (Near East). *Vegetation History and Archaeobotany* **16**:4, 329–346. DOI: `10.1007/s00334-006-0046-x`.

New, Mark, Hulme, Mike, Jones, Philip D. (2000). Representing Twentieth-Century Space–Time Climate Variability. Part II: Development of 1901–96 Monthly Grids of Terrestrial Surface Climate. *Journal of Climate* **13**:13, 2217–2238. DOI: `10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2`.

Ohlwein, Christian, Wahl, Eugene R. (2012). Review of probabilistic pollen-climate transfer methods. *Quaternary Science Reviews* **31**: online: 2011-12-16, 17–29. DOI: `10.1016/j.quascirev.2011.11.002`.

Panagiotopoulos, Konstantinos, Aufgebauer, Anne, Schäbitz, Frank, Wagner, Bernd (2013). Vegetation and climate history of the Lake Prespa region since the Lateglacial. *Quaternary International* **293**: 157–169. DOI: `10.1016/j.quaint.2012.05.048`.

Parnell, Andrew C., Sweeney, James, Doan, Thinh K., Salter-Townshend, Michael, Allen, Judy R. M., Huntley, Brian, Haslett, John (2012). On Bayesian Modelling of the Uncertainties in Palaeoclimate Reconstruction. *ArXiv e-prints*, 25.

Pross, Jörg, Klotz, Stefan, Mosbrugger, Volker (2000). Reconstructing palaeotemperatures for the Early and Middle Pleistocene using the mutual climatic range method based on plant fossils. *Quaternary Science Reviews* **19**:17-18, 1785 –1799. DOI: `10.1016/S0277-3791(00)00089-5`.

Rabier, Florence, Courtier, Philippe (1992). Four-Dimensional Assimilation In the Presence of Baroclinic Instability. *Quarterly Journal of the Royal Meteorological Society* **118**:506, 649–672. DOI: `10.1002/qj.49711850604`.

Raichlen, David A., Wood, Brian M., Gordon, Adam D., Mabulla, Audax Z. P., Marlowe, Frank W., Pontzer, Herman (2014). Evidence of Lévy walk foraging patterns in human hunter–gatherers. *Proceedings of the National Academy of Sciences* **111**:2, 728–733. DOI: `10.1073/pnas.1318616111`.

Robertson, I., Lucy, D., Baxter, L., Pollard, A. M., Aykroyd, R. G., Barker, A. C., Carter, A. H.C., Switsur, V. R., Waterhouse, J. S. (1999). A kernel-based Bayesian approach to climatic reconstruction. *The Holocene* **9**:4, 495–500. DOI: `10.1191/095968399676373488`.

Schölzel, Christian, Hense, Andreas, Hübl, Peter, Kühl, Norbert, Litt, Thomas (2002). Digitization and geo-referencing of botanical distribution maps. *Journal of Biogeography* **29**:7, 851–856. DOI: `10.1046/j.1365-2699.2002.00696.x`.

Schwab, Markus J., Neumann, Frank, Litt, Thomas, Negendank, Jorg F.W., Stein, Mordechai (2004). Holocene palaeoecology of the Golan Heights (Near East): investigation of lacustrine sediments from Birkat Ram crater lake. *Quaternary Science Reviews* **23**:16–17, 1723 –1731. DOI: `10.1016/j.quascirev.2004.05.001`.

Seppä, Heikki, Birks, H. J.B. (2001). July mean temperature and annual precipitation trends during the Holocene in the Fennoscandian tree-line area: pollen-based climate reconstructions. *The Holocene* **11**:5, 527–539. DOI: `10.1191/095968301680223486`.

Simonis, Daniel, Hense, Andreas, Litt, Thomas (2012). Reconstruction of late Glacial and Early Holocene near surface temperature anomalies in Europe and their statistical interpretation. *Quaternary International* **274**: 233–250. DOI: `10.1016/j.quaint.2012.02.050`.

Stefanova, Lydia, Krishnamurti, Tiruvalum N. (2002). Interpretation of Seasonal Climate Forecast Using Brier Skill Score, The Florida State University Superensemble, and the AMIP-I Dataset. *Journal of Climate* **15**:5, 537–544. DOI: `10.1175/1520-0442(2002)015<0537:IOSCFU>2.0.CO;2`.

Stein, Mordechai, Ben-Avraham, Zvi, Goldstein, Steven L. (2011a). Dead Sea deep cores: A window into past climate and seismicity. *Eos, Transactions American Geophysical Union* **92**:49, 453–454. DOI: `10.1029/2011EO490001`.

Stein, Mordechai, Ben-Avraham, Zvi, Goldstein, Steven L., Agnon, Amotz, Ariztegui, Daniel, Brauer, Achim, Haug, Gerald, Ito, Emi, Yasuda, Yoshinori (2011b). Deep Drilling at the Dead Sea. *Scientific Drilling* **11**: 46–47. DOI: `10.2204/iodp.sd.11.04.2011`.

Telford, R.J., Birks, H.J.B. (2011). A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages. *Quaternary Science Reviews* **30**:9–10, 1272–1278. DOI: `10.1016/j.quascirev.2011.03.002`.

Tingley, Martin P., Craigmile, Peter F., Haran, Murali, Li, Bo, Mannshardt, Elizabeth, Rajaratnam, Bala (2012). Piecing together the past: statistical insights into paleoclimatic reconstructions. *Quaternary Science Reviews* **35**: online: 2012-02-10, 1 –22. DOI: `10.1016/j.quascirev.2012.01.012`.

Tingley, Martin P., Huybers, Peter (2010). A bayesian algorithm for reconstructing climate anomalies in space and time. part I: development and applications to paleoclimate reconstruction problems. *Journal of Climate* **23**: 2759–2781. DOI: `10.1175/2009JCLI3015.1`.

– (2013). Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature* **496**:7444. online: 2013-04-10, 201–205. DOI: `10.1038/nature11969`.

Toivonen, Hannu T. T., Mannila, Heikki, Korhola, Atte, Olander, Heikki (2001). Applying Bayesian Statistics To Organism-based Environmental Reconstruction. *Ecological Applications* **11**:2, 618–630. DOI: `10.1890/1051-0761(2001)011[0618:ABSTOB]2.0.CO;2`.

Tolwinski-Ward, S. E., Tingley, Martin P., Evans, M. N., Hughes, Malcolm K., Nychka, D. W. (2014). Probabilistic reconstructions of local temperature and soil moisture from tree-ring data with potentially time-varying climatic response. English. *Climate Dynamics*, 1–16. DOI: `10.1007/s00382-014-2139-z`.

Torfstein, Adi, Goldstein, Steven L., Kushnir, Yochanan, Enzel, Yehouda, Haug, Gerald, Stein, Mordechai (2015). Dead Sea drawdown and monsoonal impacts in the Levant during the last interglacial. *Earth and Planetary Science Letters* **412**: 235 –244. DOI: `10.1016/j.epsl.2014.12.013`.

Torfstein, Adi, Goldstein, Steven L., Stein, Mordechai, Enzel, Yehouda (2013). Impacts of abrupt climate changes in the Levant from Last Glacial Dead Sea levels. *Quaternary Science Reviews* **69**: 1–7. DOI: `10.1016/j.quascirev.2013.02.015`.

van den Besselaar, E. J. M., Haylock, M. R., Schrier, G. van der, Klein Tank, A. M. G. (2011). A European daily high-resolution observational gridded data set of sea level pressure. *Journal of Geophysical Research: Atmospheres* **116**:D11, n/a–n/a. DOI: `10.1029/2010JD015468`.

Vasko, Kari, Toivonen, HannuT.T., Korhola, Atte (2000). A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. English. *Journal of Paleolimnology* **24**:3, 243–250. DOI: `10.1023/A:1008180500301`.

Vicente-Serrano, Sergio M., Beguerisa, Santiago, Lopez-Moreno, Juan I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate* **23**:7, 1696–1718. DOI: `10.1175/2009JCLI2909.1`.

Werner, Johannes P., Luterbacher, Juerg, Smerdon, Jason E. (2013). A Pseudoproxy Evaluation of Bayesian Hierarchical Modeling and Canonical Correlation Analysis for Climate Field Reconstructions over Europe. *J. Climate* **26**:3. online: 2012-07-26, 851–867. DOI: `10.1175/JCLI-D-12-00016.1`.

Scientific Theses

Garreta, Vincent (2010). “Bayesian approach of pollen-based palaeoclimate reconstructions: toward the modelling of ecological processes.” PhD thesis. University of Aix-Marseille III, France.

Kühl, Norbert (2002). “Die Bestimmung botanisch-klimatologischer Transferfunktionen und die Rekonstruktion des bodennahen Klimazustands in Europa während der Eem-Warmzeit.” PhD thesis. University of Bonn.

Migowski, Claudia (2001). “Untersuchungen laminierter holozäner Sedimente aus dem Toten Meer: Rekonstruktion von Paläoklima und -seismizität.” PhD thesis. Universität Potsdam.

Schiebel, Vera (2013). “Vegetation and climate history of the southern Levant during the last 30,000 years based on palynological investigation.” PhD thesis. University of Bonn.

Schult, Wiebke (2013). “Intercontinental Comparison of Plant-Climate-Relationships of the Northern Hemisphere.” master thesis. Rheinische Friedrich-Wilhelms-Universität Bonn.

Simonis, Daniel (2009). “Reconstruction of possible realisations of the Late Glacial and Holocene near surface climate in Central Europe.” PhD thesis. University of Bonn.

Stolzenberger, Sophie (2011). “Untersuchungen zu botanischen Paläoklimatransferfunktionen.” Diplomarbeit. Rheinische Friedrich-Wilhelms-Universität Bonn.

Electronic Data Sources

Atmospheric Research (UCAR), University Corporation for (2004). *News Release - NCAR Releases New Version of Premier Global Climate Model*. URL: `http://www.ucar.edu/news/releases/2004/ccsm.shtml`.

CRC806-Database Satellite WMS. Collaborative Research Centre 806 Our Way To Europe. URL: `http://sfb806srv.uni-koeln.de/mapproxy/demo/?tms_layer=MapQuestAerial&format=png&srs=EPSG%3A900913`.

Data Policy for ECA&D and E-OBS. URL: `http://eca.knmi.nl/documents/ECAD_datapolicy.pdf` (visited on 04/21/2016).

Digital Terrain Data (2013). National Oceanic and Atmospheric Administration. URL: `http://www.ngdc.noaa.gov/mgg/topo/topo.html`.

European Pollen Database (EPD). URL: `http://www.europeanpollendatabase.net/` (visited on 05/09/2012).

Google Maps (2012). Google Incorporated. URL: `https://maps.google.de/`.

Kapala, Alice (2003). *Großräumige Klimaänderungen und ihre Bedeutung für die Umwelt*. German. Meteorologisches Institut der Universität Bonn. URL: `http://www2.meteo.uni-bonn.de/deutsch/forschung/gruppen/klimaaenderung/files/inhalt/de/archiv/Historie/Historie-idx.html`.

NOAA Palaeoclimatology. National Oceanic and Atmospheric Administration. URL: `https://www.ncdc.noaa.gov/paleo/globalwarming/interglacial.html` (visited on 10/24/2016).

NOAA Palaeoclimatology. National Oceanic and Atmospheric Administration. URL: `https://www.ncdc.noaa.gov/paleo/globalwarming/holocene.html` (visited on 04/22/2016).

PANGAEA. URL: `http://www.pangaea.de/` (visited on 05/09/2012).

Other Sources

Breheny, Patrick, Burchett, Woodrow (2016). *visreg: Visualization of Regression Models*. R package version 2.2-2.

Grichuk, V. P. (1969). “An attempt of reconstruction of certain climatic indexes of the northern hemisphere during the atlantic stage of the holocene.” *Proceedings of the 3rd International Palynological Congress*. Ed. by Neishtadt, M., 41–57.

Harrell Jr., Frank E., with contributions from Charles Dupont and many others. (2013). *Hmisc: Harrell Miscellaneous*. R package version 3.13-0.

Ohlwein, Christian (2012). *Different pictures provided by Christian Ohlwein (Schölzel)*.

Strauch, Nina. *Das Meteorologische Observatorium Hohenpeißenberg*. English. TECHREPORT. Germany, 63067 Offenbach, Frankfurter Straße 135: Deutscher Wetterdienst.

Team, R Core (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Danksagung

Viele Menschen haben dazu beigetragen, dass diese Arbeit zu einem erfolgreichen Abschluss gekommen ist. Danken möchte ich vor allem meinem Doktorvater Prof. Dr. Andreas Hense, der es mir ermöglicht hat, diese Arbeit zu verfassen. Deine Unterstützung war so vielfältig, dass ich sie hier gar nicht einzeln aufzählen kann. Ebenfalls ein großer Dank gilt Prof. Dr. Thomas Litt, der mir oft weitergeholfen hat wenn es bei den Rekonstruktionen nicht weiterging oder es an meinem Verständnis haperte. Beiden auch vielen Dank, dass ihr diese Arbeit als Gutachter bewertet. An dieser Stelle möchte ich auch den beiden anderen Mitgliedern meiner Promotionskommission danken: Prof. Dr. Clemens Simmer und Prof. Dr. Andreas Zimmermann.

Damit ich meine Klimarekonstruktionen überhaupt durchführen konnte, haben zwei Personen viele Stunden am Mikroskop verbracht: Vera Schiebel und Kostas Panagiotopoulos. Vielen Dank, dass ich eure Daten auswerten durfte! Ein weiteres Dankeschön gilt Norbert Kühl, der mich bei der Auswahl der Taxa für die Lake Prespa Rekonstruktion tatkräftig unterstützt hat. Christian Ohlwein möchte ich dafür danken, dass er viele Fragen zu R beantwortet hat, mir bei den Programmen für die Rekonstruktionen wertvolle Tipps gegeben hat und ein paar tolle Grafiken überlassen hat. Auch die vielen Diskussionen mit ihm, Sophie Stolzenberger und Nils Weitzel, die im Bereich der Klimarekonstruktionen arbeiten, möchte ich nicht missen. Sie waren sehr hilfreich, danke! Ein weiterer Dank gilt meinen aktuellen, aber auch einigen ehemaligen Bürokollegen: Mohammad Marami-Milani, Chris Weijenborg, Lilo Bach, Marco Milan und Thorsten Simon. Danke für die vielen Diskussionen, egal ob sie wissenschaftlicher Natur, technischer Natur oder anderer Art waren! Danke auch den vielen anderen aus dem MIUB die Kicker mit mir gespielt haben.

Diese Doktorarbeit wäre ohne Korrekturleser sicherlich mit viel unverständlichem Kauderwelsch abgegeben worden. Vielen Dank an meine liebe Ehefrau Christina Thoma, meine Schwiegereltern Helga und Fritz Mohr, meinen Schwager Thomas Mohr, Sophie Stolzenberger und Chris Weijenborg.

Ohne echte Freunde, mit denen man alle Sorgen bequatschen, aber auch feiern kann, ist das Leben ein großes Stück leerer. Danke Sabrina und Michael Wahl! (An deinen neuen Namen muss ich mich auch erst gewöhnen! Aber ich freue mich von ganzem Herzen für euch!). Ein weiteres Danke gilt unseren ehemaligen Nachbarn Cindy, Benjamin und Julian Frings, Ihr habt auch dazu beigetragen, dass Bonn inwzischen mein zu Hause ist!

Danke auch an meine Eltern, dass Ihr mich so weit habt ziehen lassen und immer für mich da seid, wenn es mal eng wird! Danke Bruderherz, dass du am Schluss extra noch mal über die Arbeit geschaut hast und auch für den technischen Support (diverser Art) während der letzten Jahre. Ein winziges Dankeschön auch an meine heimliche und kleine (aber nicht einzige!!!) Motivation, mich überhaupt erst an eine Doktorarbeit zu wagen: Meine Patentante Dr. med. vet. Rita Renner. Ob du noch weißt, wie du mich motiviert hast? Ich hoffe schon!

Falls ich jemanden vergessen habe: Danke!

Danke an meine beiden großen Schätze: **Christina**! Ich weiß nicht, wo ich heute wäre ohne dich! Du hast mich immer wieder aufgebaut und gestärkt, wenn ich wieder einmal daran gezweifelt habe, ob ich das Ziel (jemals) erreichen werde. **Aaron**, du hast mich oftmals zum Lachen gebracht hat, wenn alles andere nicht mehr geholfen hat.

BONNER METEOROLOGISCHE ABHANDLUNGEN

Herausgegeben vom Meteorologischen Institut der Universität Bonn 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 1-63: siehe `http://www.meteo.uni-bonn.de/bibliothek/bma`

64-77: open access, verfügbar unter `https://uni-bn.de/kpSDaQffe1`

Heft 64: ***Michael Weniger***: Stochastic parameterization: a rigorous approach to stochastic three-dimensional primitive equations, 2014, 148 S. + XV.

Heft 65: ***Andreas Röpnack***: Bayesian model verification: predictability of convective conditions based on EPS forecasts and observations, 2014, 152 S. + VI.

Heft 66: ***Thorsten Simon***: Statistical and Dynamical Downscaling of Numerical Climate Simulations: Enhancement and Evaluation for East Asia, 2014, 48 S. + VII. + Anhänge

Heft 67: ***Elham Rahmani***: The Effect of Climate Change on Wheat in Iran, 2014, [erschienen] 2015, 96 S. + XIII.

Heft 68: ***Pablo A. Saavedra Garfias***: Retrieval of Cloud and Rainwater from Ground-Based Passive Microwave Observations with the Multi-frequency Dual-polarized Radiometer ADMIRARI, 2014, [erschienen] 2015, 168 S. + XIII.

Heft 69: ***Christoph Bollmeyer***: A high-resolution regional reanalysis for Europe and Germany - Creation and Verification with a special focus on the moisture budget, 2015, 103 S. + IX.

Heft 70: ***A S M Mostaquimur Rahman***: Influence of subsurface hydrodynamics on the lower atmosphere at the catchment scale, 2015, 98 S. + XVI.

Heft 71: ***Sabrina Wahl***: Uncertainty in mesoscale numerical weather prediction: probabilistic forecasting of precipitation, 2015, 108 S.

Heft 72: ***Markus Übel***: Simulation of mesoscale patterns and diurnal variations of atmospheric CO_2 mixing ratios with the model system TerrSysMP-CO_2, 2015, [erschienen] 2016, 158 S. + II

Heft 73: ***Christian Bernardus Maria Weijenborg***: Characteristics of Potential Vorticity anomalies associated with mesoscale extremes in the extratropical troposphere, 2015, [erschienen] 2016, 151 S. + XI

Heft 74: ***Muhammad Kaleem***: A sensitivity study of decadal climate prediction to aerosol variability using ECHAM6-HAM (GCM), 2016, 98 S. + XII

Heft 75: ***Theresa Bick***: 3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale, 2016, [erschienen] 2017, 96 S. + IX

Heft 76: ***Zied Ben Bouallegue***: Verification and post-processing of ensemble weather forecasts for renewable energy applications, 2017, 119 S.

Heft 77: ***Julia Lutz***: Improvements and application of the STatistical Analogue Resampling Scheme STARS, 2016, [erschienen] 2017, 103 S.

Heft 78: ***Benno Michael Thoma***: Palaeoclimate Reconstruction in the Levant and on the Balkans, 2016, [erschienen] 2017, XVI, 266 S.

METEOROLOGISCHES INSTITUT
MATHEMATISCH NATURWISSENSCHAFTLICHE FAKULTÄT
UNIVERSITÄT BONN