Reconstruction of possible realisations of the Late Glacial and Holocene near surface climate in Central Europe

Dissertation

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

vorgelegt von
Daniel Simonis

aus
Sigmaringen

Bonn, Juli 2009

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

1. Gutachter: Prof. Dr. Thomas Litt
2. Gutachter: Prof. Dr. Andreas Hense
Tag der Promotion: 18.12.2009

Erscheinungsjahr: 2010

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Abstract

This thesis presents several aspects of reconstructing physical consistent realisations of past climatological fields. Local climate reconstructions are obtained by a method, which is based on the idea of indicator taxa and uses presence of taxa as proxy variable. In previous studies, the indicator taxa approach has been enhanced to a probabilistic Bayesian reconstruction method, which provides conditional probability density functions as reconstruction result.

Up to now bivariate normal distributions or mixture models have been applied for reconstructing January and July temperatures. A three dimensional copula approach exists for the incorporation of annual precipitation. Now mixture models are embedded into this approach and a new set of three dimensional transfer functions is estimated. The differences to two dimensional mixture models are examined.

The local climate reconstruction results are interpolated in a dynamically consistent way by applying a variational analysis with weak physical constraint. For reconstructing fields of annual precipitation, a different physical constraint is implemented into the analysis.

A new point of view for the interpretation of climate reconstruction results is proposed. It emphasizes, that the analysis result has to be seen as conditional expectation of the desired climatological field. This expectation is only the mean of all possible realisations of the past climate. In this work, possible realisations are presented and it becomes clear, that these can differ considerably from the mean field. The realisations are obtained by resampling from the analysis error covariance matrix of the variational analysis.

Reconstructions of near surface January and July temperature anomalies for two Late Glacial (13000 and 12000 cal. BP) and two Holocene (8000 and 6000 cal. BP) time slices are provided, based on pollen and macrofossil data from 85 different locations in Europe. The variational analysis is for the first time applied for reconstructing a cold climate state. It becomes clear that the sensitivity of the botanical proxies, which are used in this work, is too low for capturing the difference between 13000 and 12000 BP. Both time slices are reconstructed significantly colder than the Holocene. The results for the Holocene time slices agree well with results from other studies. No significant differences to the modern 1961-90 climate can be found.

A successful reconstruction of fields of annual precipitation anomalies is not possible. Apparently the botanical proxies, applied in this work, are not sensitive enough for this purpose. Both, the results for the Late Glacial and

the results for annual precipitation call for the incorporation of other proxy data and a multiproxy approach.

Contents

Chapter 1

Introduction

The period of the early 21st century is a very interesting one in the field of climatology. Most likely due to the increase of greenhouse gases, mainly CO_2, the climate system has already warmed significantly and is expected to reach a temperature level out of the range mankind has ever experienced. This statement would never be possible without paleoclimatological research. Instrumental measurements of climate parameters only reach back to 1850 (Brohan et al., 2006). For receiving information about the climate beyond that date, the field of paleoclimatology was created.
It was stated in the fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC), that in 1990 "...many climatic variations prior to the instrumental record were not that well known or understood. Fifteen years later, understanding is much improved, more quantitative and better integrated with respect to observations and modelling" (Jansen et al., 2007). In these recent years the knowledge has improved a lot, concerning the variability of the past climate. However, there still is a large uncertainty in answers to questions as e.g. what were the absolute differences in the global temperature between maxima of glaciation and deglaciation or what was the regional impact of these transitions of the climate state.
Especially the role of internal climate variability is a key topic of current research. The fact that the global mean temperature increased strongly in the 1990th but nearly remained constant during the first decade of the 21st century, raised the question how large the influence of internal variability in a warming climate is (Easterling and Wehner, 2009; Swanson and Tsonis, 2009). It is also a matter of discussion if climate anomalies or cycles were driven by external forcing or internal variability. Examples for that are the Little Ice Age and Medieval Climate Anomaly (Trouet et al., 2009) or Dansgaard-Oeschger events during glacial periods (Ditlevsen et al., 2007).
To understand the mechanisms of climate variability in the past and the pos-

sible strength of internal cycles is crucial for the prediction of future climate change. More precise reconstructions of the regional impact and the spatial characteristics of past climate changes will help to improve future climate change projections on a regional scale. These aspects motivate this thesis, the reconstruction of physically consistent temperature fields and a detailed analysis of the uncertainties of these reconstructions.
In the following sections, first a short overview of the climate of the Quaternary will be given. Afterwards the scientific background will be discussed, focussing on different methods for climate reconstructions. Finally the development of the method used in this work will be described and its advantages and disadvantages will be discussed.

1.1 A brief overview of the Quaternary climate

The Quaternary is the youngest geological time period in the history of earth. It began about 2.6 million years before present (BP) and spans the epoch of development of the species *Homo sapiens*. It is divided into two parts, the Pleistocene and the Holocene, which endures since 11700 years (Ogg et al., 2008). The climate of the Pleistocene was characterised by periodic glacial cycles. One glacial cycle normally consists of a cold phase and a warm phase, lasting between 10 and 30 ka while a complete cycle has length of about 100 ka. During the cold phases glaciation took place in the northern latitudes. Thus, the cold phases are often referred to as "glacials" and the warm phases as "interglacials". During the last 750 ka, eight glacial cycles have taken place (EPICA Community Members, 2004). Not all glacial cycles show the same behaviour. Between 430 ka and 740 ka the warm phases lasted longer than in the younger cycles, but on the other hand did not reach the same high temperature levels. A curve of isotope ratios, indicating the climate history of the past 740 ka can be found in Fig. (1.1).
The last interglacial, called the "Eemian" interglacial and the most recent glacial are the best examined ones. For the first phase of the Eemian, slightly higher July temperatures and slightly lower January temperatures, compared to the 1961-90 climate, have been reconstructed by Gebhardt et al. (2008) for Central Europe. Climate reconstructions for the Last Glacial Maximum (LGM) around 21 ka BP show a heterogeneous pattern of cooling over the globe. In the tropics anomalies have been between -2 K and -3 K at sea level and below -6 K at high altitudes according to Farrera et al. (1999). For Greenland, anomalies of -23 K were reconstructed by Dahl-Jensen et al.

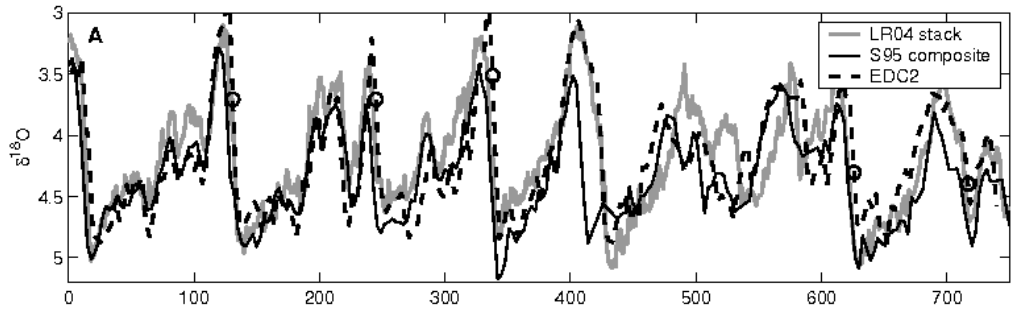

Figure 1.1: Oxygen isotope ratios from the Antarctica for the last 740 ka (Lisiecki and Raymo, 2005). Maxima of the curves indicate warm phase while minima indicate cold phases. The different curves represent different models, what is not of relevance here.

(1998) from data of the GRIP borehole. For the LGM also climate model simulations have been carried out within the Paleoclimate Modeling Intercomparison Project (PMIP2, Braconnot et al. (2007)). The mean results of the climate models show anomalies around -2 K in the tropics, and maxima of cooling over Scandinavia and North America with -20 K and -30 K, respectively. Simulated anomalies for Central Europe are around -10 K and for Greenland around -15 K. These results disagree to the reconstructions mentioned above. Also for Western Europe and the Mediterranean, several climate reconstructions result in lower temperatures than simulated (Ramstein et al., 2007).
For the last glacial cycle ice cores with a high resolution have been analysed (NGRIP-Members, 2005). Such records show that during a glacial phase, the climate does not stay on a stable cold level. Instead abrupt warming (Dansgaard-Oeschger-Events) and cooling (Heinrich-Events) occurs. These variations are also referred to as stadials (cold) and interstadials (warm) and are most likely related to different modes of the Northatlantic Thermohaline Circulation (THC, Stocker et al. (1992); Ganopolski and Rahmstorf (2001)). Also the warming during the deglaciation process from the LGM to the Holocene did not happen steadily. The deglaciation was interrupted during the Younger Dryas (∼12000 BP). Although different theories have been discussed for explaining this event, the most plausible explanation seems to be, that freshwater from the melting Laurentide ice sheet disturbed the THC (Clark et al., 2001). After the Younger Dryas the current interglacial, the Holocene, began. How long it will last, until the next glaciation has to be expected, has been discussed by Loutre and Berger (2000). They pointed

out that, due to orbital parameters, no large variations in solar insolation are expected during the next 20 ka and a critical point for glaciation will not be reached within that period. In dependence of how much the values of CO_2 will increase due to human activities, it is possible that the current interglacial will last for more than 50 ka.

1.2 Methods for climate reconstructions

A variety of methods is used for quantitative climate reconstructions. These methods were discussed extensively by Schölzel (2005), who focused on their statistical framework and their relation to the reconstruction approach presented by him. The most important methods and their properties will be briefly presented in the following.

Regression methods

Regression methods like linear regression, principal component regression or weighted average regression are very common and widely used for reconstructing climatological or environmental quantities. A comprehensive description of these methods is given by Birks (1995). Reconstructions by regression methods are always based on a calibration dataset on which regression parameters, describing a linear relationship between proxy and target variable, are estimated.
Regression methods are applied to many proxy data and for reconstructing many different climatological variables. An example is Seppä et al. (2004) who used weighted average partial least squares to estimate a transfer function between surface pollen samples and the annual mean temperature. In another study Bar-Matthews et al. (2003) applied linear regression to oxygen and carbon isotope ratios for reconstructing annual precipitation. A final example that could be mentioned here is the work of Cook et al. (2001) who reconstructed the North Atlantic Oscillation Index by application of principal component regression to data of tree ring width and oxygen isotope ratios.
It was pointed out by Schölzel (2005) that regression methods represent the roughest approximation of the stochastical relationship between environmental and proxy random variables, because they only regard expectation values instead of probability densities. Why the relation between environmental and proxy variables is stochastic, is discussed in Section (2.2.1).
Additionally, the application of regression methods can lead to biased results (Robertson et al., 1999) or underestimate the variability of reconstructed

variables (von Storch et al., 2004, 2006).

Bayesian methods

Bayesian approaches are relatively new in the field of paleoclimatology. They were introduced at the beginning of this decade by a Finnish research group (Vasko et al., 2000; Toivonen et al., 2001; Korhola et al., 2002), who used a hierarchical Bayesian network for reconstructing temperatures from abundances of chironomides (species of midges).
In a Bayesian approach (see also chapter (2.2.2)), a posterior distribution is calculated by integration of a likelihood function multiplied with a prior density over the parameter space. The integration can be performed by Markov chain Monte Carlo integration (MCMC,Gilks et al. (1996)), which is a method to approximate the posterior distribution by sampling. As Schölzel (2005) underlines, this procedure has the advantage of considering all quantities as random variables. Furthermore the calculation of posterior distributions provides an explicit description of the uncertainty. Finally, prior knowledge can be introduced by prior probability distributions. On the other hand computational problems can come along with MCMC.
These problems are avoided when the likelihood functions are estimated by explicit parameter estimation, like in Robertson et al. (1999) or in this work. Doing this means estimation of the "optimal" model parameters instead of integration over the whole parameter space. Although Schölzel (2005) stated, that this was a rough simplification of the idea of the Bayesian approach, the important aspects of this approach (calculation of posterior distribution, inclusion of prior distribution) are not touched by that.

Modern analogue methods

Another widely applied method for reconstructing climate from pollen samples is the modern analogue technique (MAT), which was introduced by Overpeck et al. (1985). It is based on the assumption that a certain climate state is represented by a certain sample of abundances of pollen from different vegetation elements (taxa) in the sediment. Hence a dataset of modern pollen samples is needed. Then the difference between fossil pollen samples and the modern samples is calculated by dissimilarity coefficients. The modern sample with the lowest difference is regarded as the best analogue and therefore as reconstructed climate. Examples for the application of this method are Guiot and Couteaux (1992), Cheddadi et al. (1997), Isarin et al. (1998) or Davis et al. (2003).
Modern analogues are problematic when they are applied for reconstructing

a climate that has no modern counterpart, as possible for glacial climate states. Anthropogenic influences like land cultivation and deforestation also can be a problem. Finally, the MAT is a deterministic method that disregards the stochastic nature of the climate system. This leads to the problem that only a climate value is reconstructed, without any error measure.

Other approaches

Some other approaches that also should be mentioned here, as they are applied to vegetational proxy data, are biomisation and plant functional types (pft), (Prentice et al., 1992, 1996). Here different taxa are grouped together to pfts and then pfts are grouped to biomes. These methods are generally used for reconstruction of vegetation and vegetation modelling. However, they are also used as a basis for climate reconstruction by combination with the MAT or a regression method. With the development of the Bayesian biome model, Schölzel (2005) showed that it is possible to implement these concepts into a Bayesian approach.
In recent years a new approach for reconstructing climate by inverse vegetation modelling came up (e.g. Wu et al. (2007)). This approach is also based on biomisation, because a vegetation model that uses biomes is applied.

1.3 Fundamentals and motivation for this work

In the previous section a variety of different methods, which are used for climate reconstruction, has been presented. When these methods are applied to paleobotanical data, abundances of pollen are used as proxy variable. In this work the presence of taxa is used as proxy variable, implicating several advantages and one major disadvantage. Advantages of using the presence of taxa are the possibility to better take into account taxa with low pollen production rate and the possibility to include macrofossils. The latter are sparsely represented in the sediment but can normally be assigned to certain species while pollen often only allow for the identification of genus or family. Regarding presence also makes the method robust against anthropogenic influences or non-modern analogue situations. For this robustness, however, one pays by loosing the information that is contained in the abundances, what results in larger error bounds (Litt et al., 2009).
The original idea for this approach, to use certain plants as climate indicators, dates back to Iversen (1944). Based on the idea of climate indicator species, Grichuk (1969) introduced the method of Mutual Climatic Ranges (MCR), which is a graphical approach. In its two-dimensional application

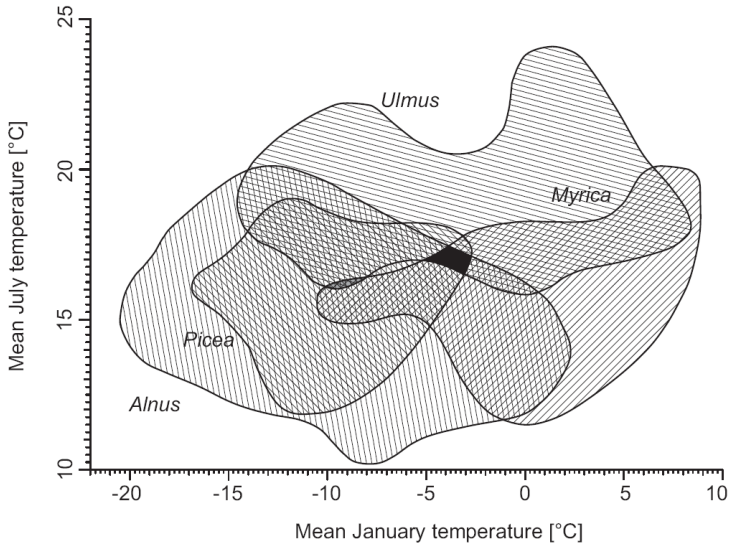

Figure 1.2: Illustration of the MCR method from Schölzel (2005), redrawn after Grichuk (1969). The black area indicates the intersection of all taxa in the climate phase space. A second region intersection that was not indicated in the original figure is marked grey.

all climate values from stations where a certain taxon occurs is represented by a point on a plane. Then a line is drawn around these points, resulting in closed areas. Now all areas of plants occurring at a certain location are placed on the same plane. The intersection of all these areas is regarded as reconstructed climate. This method has not only been applied to vegetational taxa but also to beetle remains (Atkinson et al., 1987), for instance.
An illustration of the MCR method is given in Fig. (1.2). Some disadvantages of the graphical approach are evident. Due to the sharp boundaries, the occurrence of a taxon is equally likely within all of its area but suddenly drops to zero outside of it. Furthermore, clear signs of overfitting are visible, mostly in the area of *Ulmus* with its dented shape. However these problems are overcome by the implementation of these ideas into a statistical concept. This was done by replacing the mutual climatic ranges with probability density functions in the PDF-method (Kühl et al., 2002; Kühl, 2002; Gebhardt, 2003). Within these studies a cooperation between Paleobotanists and Meteorologists at the University of Bonn was initiated for effectively merging two fields of research, which both are essential for reconstructing climate from paleobotanical data.
In the original PDF-method, bivariate normal distributions were used as

transfer functions for January and July temperature. As several taxa need a more flexible statistical model which is able to represent non-gaussian properties like skewness, Gebhardt (2003) applied gaussian mixture models as transfer functions. He also introduced a variational method for reconstructing dynamically consistent temperature fields under a weak physical constraint.
In his work Schölzel (2005) concentrated on the statistical framework of the method and enhanced the PDF-method to the Bayesian indicator taxa model. There he successfully included the gamma distribution for precipitation by using a copula approach for three-dimensional pdfs with mixed marginals (Schölzel and Friederichs, 2008). By applying the method to a sediment core from the Near East, it was for the first time used in an area outside of Europe (Neumann et al., 2007).
The success of Schölzel (2005) in reconstructing annual precipitation in the Near East motivates to combine his three dimensional approach with the more flexible mixture models, what is one goal of this thesis. Another goal is motivated by the successful reconstruction of temperature fields by Gebhardt (2003). The variational analysis should be enhanced to the ability to reconstruct dynamically consistent fields of precipitation. For this purpose also another physical constraint is needed.

This study was embedded in the European Science Foundation project DECVeg (Dynamic European Climate-Vegetation impacts and interactions). At the beginning of this project it was decided, that reconstructions for the four time slices (13000, 12000, 8000 and 6000 cal. BP) should be provided. These times slices are classified in the Earth's history as follows:

- The period around 13000 BP is called Alleroed. It is located at the end of the last glacial period (Late Glacial) and is characterised by a relatively warm climate (interstadial). This is reflected in the vegetation by an increasing amount of trees and shrubs.

- The transition from the last glacial period into the Holocene was interrupted around 12000 BP by a reemergence of glacial conditions (stadial). This is reflected in the vegetation records by decreasing abundances of trees and shrubs. This period is called Younger Dryas.

- The Younger Dryas was followed by a rapid warming, the beginning of the Holocene, that still endures. At 8000 BP the so called Holocene climate optimum started. Therefore this time slice is regarded as interesting.

- Finally the period around 6000 BP marks the beginning of the Neolithic (New Stone Age) in Europe. It also lies in the middle of the Holocene climate optimum.

The time slices 13000 and 12000 BP will be referred to as the "Late Glacial" time slices, the others as the "Holocene" time slices in the following.
While the PDF-method and its successive methods have already been applied to a cold climate phase in a local time series reconstruction (Kühl et al., 2007), this is not the case for the variational analysis. With its application to the Late Glacial time slices, the method will be used for the first time to reconstruct fields for a cold climate state. It was pointed out by Gebhardt (2003) that the variational reconstruction could be problematic in glacial times, because its formulation assumes similarity of reconstructed climate to the modern climate. It is yet unclear, however, where the limit for this similarity lies. Hence the reconstructions for the Late Glacial time slices will also address this question.
The variational reconstruction has already proven that it is sensitive enough to resolve the differences between two phases of the Eemian interglacial. Thus it will be interesting if the method is also sensitive enough to resolve the differences between the two Late Glacial time slices.
A final and very important issue that plays an essential role in this work is the presentation of the analysis error. An idea will be proposed, how to illustrate the spatial uncertainties of a reconstructed two dimensional field.

The outline of this thesis is as follows:

- In chapter 2 the important theoretical aspects are introduced. It will be shown how transfer functions are defined and how three dimensional mixture models are estimated. The problems which can arise with mixture models are discussed and another approach to determine the optimal number of mixture model components is proposed to avoid overfitting more effectively. Finally, the local reconstruction with the indicator taxa model is described.

- Chapter 3 presents the reconstruction of physically consistent fields by the variational analysis. A physical constraint for reconstructing fields of annual precipitation is presented and a resampling approach to visualise the uncertainty of the reconstructed fields is suggested.

- The data basis for this work is discussed in chapter 4.

- Chapter 5 contains the most important results of this thesis. Regarding the transfer functions, differences between a two dimensional and

a three dimensional approach will be analysed. The reconstruction results for the different time slices are presented, following the results of an experiment to reconstruct the 1961-90 climate.

- Chapter 6 addresses certain aspects of the field reconstructions and compares the results to the current state of research. Finally, the attempt to reconstruct annual precipitation is assessed.

- In chapter 7 a final conclusion will be given, together with ideas for future research.

Chapter 2

Transfer functions and local reconstructions

In this chapter the theoretical aspects of climate reconstructions based on probability density functions (pdfs) are described. First, the statistical basics are presented, followed by an introduction of different statistical models that can be applied for fitting statistical transfer functions.
The theory of pdfs as statistical transfer functions presented here is based on several studies done in cooperation between Meteorologists and Paleontologists in recent years, as mentioned in the introduction. Before a detailed description of the method will take place, some important basic principles of the underlying statistics will be introduced. For understanding, the reader should have some basic knowledge about statistics. There are several textbooks which provide these basics, e.g. von Storch and Zwiers (1999).

2.1 Important distributions and statistics

2.1.1 The normal distribution

The normal (or gaussian) distribution is most likely the best known and mostly applied statistical distribution. A lot of quantities in many scientific fields can be described by a normal distribution and a lot of statistical procedures like e.g. the linear regression analysis are based on the normal distribution. The probability density of the multivariate normal distribution is given by

$$f_{N(\vec{\mu},\Sigma)}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^q \det \Sigma}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu}) \right), \qquad (2.1)$$

where $\vec{x}$ is the data vector of the random variable, $\vec{\mu}$ denotes the mean vector and Σ the covariance matrix of the distribution. Some properties of the covariance matrix should be noted. It is symmetric, positive definite and in the case of uncorrelated random variables it becomes a diagonal matrix.
In this work the normal distribution is applied to climatological monthly mean temperature values.

2.1.2 The gamma distribution

Not all climatological variables can be handled by the normal distribution. Annual precipitation has several properties that cannot be captured by a normal distribution. First, it has no negative values, while the density for negative values in the normal distribution will never be zero. Further annual precipitation is non-symmetric as most of its values lie around 500 mm in Europe, but there also exist very high values above 2000 mm.
The gamma distribution is able to represent properties like skewness and is only valid for positive values. Its definition is given by

$$f_{G(\nu,\lambda)}(x) = \frac{1}{\Gamma(\nu)} \lambda^{\nu} x^{\nu-1} \exp(-\lambda x). \tag{2.2}$$

Like the normal distribution, the gamma distribution also is defined by two parameters $\nu > 0$ (shape) and $\lambda > 0$ (rate). The definition uses the gamma function

$$\Gamma(\nu) = \Gamma_{\infty}(\nu) = \int_0^{\infty} t^{\nu-1} \exp(-t) dt. \tag{2.3}$$

The parameters ν and λ are related to mean and variance of the distribution in the following way:

$$\hat{\mu} = \frac{\hat{\nu}}{\hat{\lambda}}, \qquad \hat{\sigma}^2 = \frac{\hat{\nu}}{\hat{\lambda}^2}. \tag{2.4}$$

2.1.3 Mixture models

A more flexible possibility to describe a population is offered by mixture models. These are defined by a weighted sum over a certain number of component densities (Titterington et al., 1985):

$$f_{MIX}(\vec{x}) = \sum_{k=1}^{K} \alpha_k f_k(\vec{x}). \tag{2.5}$$

The density function $f_k(\vec{x})$ is the one of a normal distribution in most cases, but also other distributions like the gamma distribution can be used for a

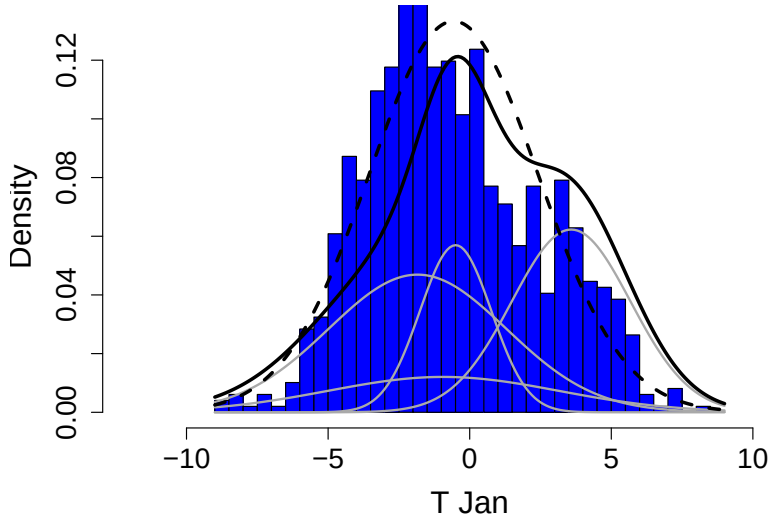

Figure 2.1: Fit to January temperature data displayed as blue histogram: normal distribution (dashed black line), mixture model (black line) and five mixture model components (grey lines).

mixture model. However the parameter estimation is a problem in such cases, especially as it can already be difficult in the case of normal distributions. The problem of parameter estimation is described later on in this chapter.

Figure (2.1) shows the differences between a normal distribution and a Gaussian mixture model, fitted to certain data. It becomes clear that a normal distribution, despite fitting the data quite well, is not able to display skewness and other non Gaussian properties. On contrary the mixture model captures the skewness and provides a closer fit to the data.

2.1.4 Distributions with mixed marginals

In climatological applications one often has to deal with random vectors containing components to be described by different distributions. As mentioned, in this work we have temperature, described by normal distributions and precipitation, described by gamma distributions.
As both, the parameters of the gamma distribution and of the normal distribution can be transformed to standard normal distribution, the component vector can be described by that transformation. A brief description of that procedure was given by Schölzel (2005) and more information can be found in literature by "inverse CDF method", e.g.Gentle (2003).
The transformation from a normal distributed random variable $X_j \sim N(\mu_j, \sigma_j^2)$

to a standard normal distributed random variable $Z_j \sim N(0,1)$ is given by

$$\phi_j(x) = \frac{x - \hat{\mu}_j}{\hat{\sigma}_j}. \tag{2.6}$$

The following rule gives the transformation from a gamma distributed random variable $X_j \sim G(\nu_j, \lambda_j)$ to a standard normal distributed one:

$$\phi_j(x) = F^{-1}_{N(0,1)}(F_{G(\nu_j,\lambda_j)}), \tag{2.7}$$

where $F_{N(0,1)}$ is the cumulative distribution function (cdf) of the standard normal distribution and $G(\nu_j, \lambda_j)$ is the one of the gamma distribution. These cdfs are given by

$$F_{N(0,1)}(x) = \int_{-\infty}^{x} f_{N(0,1)}(x^{'})dx^{'}, \tag{2.8}$$

$$F_{G(\nu_j,\lambda_j)}(x) = \frac{\Gamma_{\lambda_j x}(\nu)}{\Gamma(\nu)}. \tag{2.9}$$

Now we regard a random vector with mixed normal and gamma marginals $\vec{X} \sim D(\vec{\mu}, \Sigma, \nu, \lambda)$. As shown by Schölzel (2005) the pdf of $\vec{X}$ can be written using the transformation $\vec{\phi}(\vec{x})$:

$$f_{\vec{X}}(\vec{x}) = \mid det(J_{\vec{\phi}}(\vec{x})) \mid \cdot f_{N(\vec{\mu},\Sigma)}(\vec{\phi}(\vec{x})), \tag{2.10}$$

where $J_{\vec{\phi}}(\vec{x})$ is the Jacobian of the transformation $\vec{\phi}$ containing its components on the principal diagonal:

$$J_{\vec{\phi}}(\vec{x}) = diag(\frac{\partial \phi_1}{\partial x_1}, ..., \frac{\partial \phi_n}{\partial x_n}). \tag{2.11}$$

Differentiation of equations (2.6) and (2.7) leads to

$$\frac{\partial \phi_j}{\partial x_j} = \frac{1}{\hat{\sigma}_j} \tag{2.12}$$

for the j-th normal component and

$$\frac{\partial \phi_j}{\partial x_j} = f_{G(\hat{\nu}_j,\hat{\lambda}_j)}(x_j) \cdot \left(f_{N(0,1)} \left(F^{-1}_{N(0,1)}(F_{G(\hat{\nu}_j,\hat{\lambda}_j)}(x_j)) \right) \right) \tag{2.13}$$

for the j-th gamma component. More details again can be found in Schölzel (2005).

It is important to note, that the equations presented here are not only valid for single distributions with mixed marginals but also for mixture models. The pdf of the mixture model can just be inserted into Eq. (2.10), what only changes the parameters it depends on:

$$f_{\vec{X}(\vec{x})} = \mid det(J_{\vec{\phi}}(\vec{x})) \mid \cdot f_{MIX(\hat{\mu_1},...,\hat{\mu_N}\,\hat{\Sigma_1},...,\hat{\Sigma_N})}(\vec{\phi}(\vec{x})). \tag{2.14}$$

The transformation vector $\vec{\phi}(\vec{x})$ remains unchanged as the estimation of mixture model parameters is carried out on the transformed variables.

2.2 Estimation of statistical transfer functions

2.2.1 The need for statistics

In this section it will be described how statistical transfer functions, based on botanical data, are estimated. The first question arising is, why the transfer function has to be a statistical one. There are two potential reasons conceivable:

- The vegetational-climatological relation is *stochastic*.
- The vegetational-climatological relation is *deterministic*, but we don't know the exact relation and have only measures for it. These underlie random statistical errors.

Latter would mean that certain environmental conditions would definitely lead to a certain vegetation composition. If we reduce the environmental influences controlling vegetation to climate only, this would mean $(C_1, ..., C_n) \Leftrightarrow (P_1, ..., P_n)$, where C_i are climate variables and P_i the occurring plants. So it would be possible to conclude from a given vector of plants $\vec{T}$ to a certain vector $\vec{P}$ of climate variables. This perception automatically leads to a *modern analogue method*.
However even in the case of a deterministic relationship it would not be possible to avoid the use of statistics. First, on the side of climate variables, there are observational errors which are always connected with any kind of measurement. Second, on the side of vegetation one can almost for sure assume that the historical dataset will never contain all taxa, having occurred in the past. Finally there remains the problem, that climate is not the only factor controlling the occurrence and distribution of plants. Other important factors are, for instance, soil properties, geographical barriers like oceans or mountain ranges and also competition between different plants.
All these factors introduce uncertainty into the relation between climate and

vegetation and thus emphasise the appliance of statistics to describe this relation.

From a more theoretical point of view, several considerations imply that the relation between climate and vegetation is a stochastic one. It came to mind by the work of Lorenz (1963) that the atmosphere (and thus the climate system) has to be regarded as a chaotic/stochastic system. In circulation models, used for weather forecasting, small differences in the initial state lead to completely different solutions after a few weeks. All of these solutions normally don't reflect the real state of the atmosphere at this point. One reason for that is the high dimensionality of the atmosphere. Theoretically every molecule would have to be modelled and of course then its initial condition would have to be known. This is impossible from the point of measuring as well as from the point of modeling.
Another aspect is the nonlinearity of processes in the climate system, causing small differences in the initial state to grow exponentially.
All this is reflected in the field of atmospheric modeling by already a long history of ensemble forecasting (Lewis, 2005). This means that a model is run several times with only small differences in its initial conditions, each representing a different realisation of the state of the atmosphere, reflecting its stochastic nature. Also in the longer time scales of climate model simulations, in recent years the use of ensembles opened up a new era (Collins, 2007).
Still one could argue that the climate system is not really stochastic but only seems so to us due to our lack of understanding and our limited possibilities of simulating processes. However the inherent stochastic nature becomes clear in several processes like the El Niño/Southern Oscillation (ENSO)-circulation (Kleeman, 2008). The power spectrum of long temperature time series reflects red noise patterns. Also climate cycles during the glacials, namely Dansgaard-Oeschger events, are connected with random noise (Ditlevsen et al., 2005).
Consideration of all these factors makes clear that the climate system is indeed stochastic. Similar thoughts apply to the biospheric system. Vegetation, for instance, underlies external radom influences like fires or infestation with vermin. Thus, the relation between climate and vegetation has to be considered as a stochastic one. Hence it is a logical step and even more an absolute must to use statistical transfer functions for describing this relation.

2.2.2 Definition of transfer functions

Transfer functions in the field of climate reconstructions define the relation between climate variables ($\vec{X}$) and proxy variables ($\vec{Y}$). Latter are quantities containing a certain amount of information about the desired climatological variable. There exist a variety of proxy variables like different isotope ratios, width of tree rings or even historical documents. In this work occurrence of vegetation elements (taxa) is used as proxy variable.
As Schölzel (2005) pointed out in his thesis, recent variables and palaeo variables can not be regarded as realisations of the same random variables, as they are recorded under different boundary conditions.
So palaeo climate variables ($\vec{X}_0$) are distinguished from recent climate variables although they normally represent the same quantities like air temperature or precipitation. However, due to different treatment like spatial or temporal averaging, they are considered as realisations of different random variables.
Also palaeo proxy variables ($\vec{Y}_0$) are distinguished from the recent ones. Today occurrence of plants can be observed directly. In palaeo archives like lake sediments only evidence for occurrence of a taxon can be found. Pollen or macro remains of vegetation are present in the archives but there is no guarantee that this is the case for all taxa that occurred at a certain place. Additionally at least the occurrence of pollen of some taxa is not a perfect evidence, as they can be transported over long distances. So here it is quite clear that recent and palaeo proxy variables are realisations of different random variables. The same point of view was also proposed by Toivonen et al. (2001) and Korhola et al. (2002).
A transfer function is defined in the most general way by a joint probability density function,

$$\begin{aligned} f_{\vec{X},\vec{Y}}(\vec{x},\vec{y}) &= f_{\vec{X}|\vec{Y}}(\vec{x} \mid \vec{y}) \cdot f_{\vec{Y}}(\vec{y}) \\ &= f_{\vec{Y}|\vec{X}}(\vec{y} \mid \vec{x}) \cdot f_{\vec{X}}(\vec{x}), \end{aligned} \tag{2.15}$$

as a product of conditional and marginal probability density functions. The transfer function is defined analogously for the palaeo variables:

$$\begin{aligned} f_{\vec{X}_0,\vec{Y}_0}(\vec{x}_0,\vec{y}_0) &= f_{\vec{X}_0|\vec{Y}_0}(\vec{x}_0 \mid \vec{y}_0) \cdot f_{\vec{Y}_0}(\vec{y}_0) \\ &= f_{\vec{Y}_0|\vec{X}_0}(\vec{y}_0 \mid \vec{x}_0) \cdot f_{\vec{X}_0}(\vec{x}_0). \end{aligned} \tag{2.16}$$

A rule for conditional probabilities is given by the Bayes theorem, based on the work of Bayes (1763). One can find a good overview over the Bayesian theory in Berger (1985). The Bayes theorem for continuous random variables

is given by

$$f_{\vec{X}|\vec{Y}}(\vec{x} \mid \vec{y}) = \frac{f_{\vec{Y}|\vec{X}}(\vec{y} \mid \vec{x}) \cdot \pi_{\vec{X}}(\vec{x})}{m_{\vec{Y}}(\vec{y})} \tag{2.17}$$

and can be formulated analogously for the palaeo variables.
As pointed out by Schölzel (2005), the full conditional probability density function depends on recent and palaeo variables and is examined by introducing a parameter set $\vec{\Theta}$, containing all possible transfer function parameters, and the according parameter space ν_Θ. Further, the recent random vectors $\vec{X}$ and $\vec{Y}$ are represented by the sample matrices $\mathbf{X}$ and $\mathbf{Y}$. The full conditional probability density is obtained by integration over the parameter space:

$$f_{\vec{X}_0|\vec{Y}_0,\vec{X},\vec{Y}}(\vec{x}_0 \mid \vec{y}_0, \mathbf{X}, \mathbf{Y}) = \int_{\nu_\Theta} f_{\vec{X}_0,\vec{\Theta}|\vec{Y}_0,\vec{X},\vec{Y}}(\vec{x}_0, \vec{\theta} \mid \vec{y}_0, \mathbf{X}, \mathbf{Y}) d\vec{\theta}. \tag{2.18}$$

The according formulation of the Bayes theorem is given by

$$f_{\vec{X}_0,\vec{\Theta}|\vec{Y}_0,\vec{X},\vec{Y}}(\vec{x}_0, \vec{\theta} \mid \vec{y}_0, \mathbf{X}, \mathbf{Y}) = \\ f_{\vec{Y}_0|\vec{X}_0,\vec{X},\vec{Y},\vec{\Theta}}(\vec{y}_0 \mid \vec{x}_0, \mathbf{X}, \mathbf{Y}, \vec{\theta}) \cdot \frac{\pi_{\vec{X}_0,\vec{\Theta}|\vec{X},\vec{Y}}(\vec{x}_0, \vec{\theta} \mid \mathbf{X}, \mathbf{Y})}{m_{\vec{Y}_0|\vec{X},\vec{Y}}(\vec{y}_0 \mid \mathbf{X}, \mathbf{Y})}. \tag{2.19}$$

Inserting Eq. (2.19) into Eq. (2.18) leads to

$$\begin{aligned} & f_{\vec{X}_0|\vec{Y}_0,\vec{X},\vec{Y}}(\vec{x}_0 \mid \vec{y}_0, \mathbf{X}, \mathbf{Y}) = \\ & = \int_{\nu_\Theta} f_{\vec{Y}_0|\vec{X}_0,\vec{X},\vec{Y},\vec{\Theta}}(\vec{y}_0 \mid \vec{x}_0, \mathbf{X}, \mathbf{Y}, \vec{\theta}) \cdot \frac{\pi_{\vec{X}_0,\vec{\Theta}|\vec{X},\vec{Y}}(\vec{x}_0, \vec{\theta} \mid \mathbf{X}, \mathbf{Y})}{m_{\vec{Y}_0|\vec{X},\vec{Y}}(\vec{y}_0 \mid \mathbf{X}, \mathbf{Y})} d\vec{\theta} \\ & = \int_{\nu_\Theta} f_{\vec{Y}_0|\vec{X}_0,\vec{X},\vec{Y},\vec{\Theta}}(\vec{y}_0 \mid \vec{x}_0, \mathbf{X}, \mathbf{Y}, \vec{\theta}) \cdot \frac{\pi_{\vec{X}_0}(\vec{x}_0)}{m_{\vec{Y}_0}(\vec{y}_0)} \pi_{\vec{\Theta}|\vec{X},\vec{Y}}(\vec{\theta} \mid \mathbf{X}, \mathbf{Y}) d\vec{\theta}. \end{aligned} \tag{2.20}$$

In the case of explicit parameter estimation, the "optimal" parameters $\vec{\theta}_{opt}$ are estimated. The according distribution in dependence of the recent random variables $\vec{X}$ and $\vec{Y}$ is described by Dirac's delta function:

$$\pi_{\vec{\Theta}|\vec{X},\vec{Y}}(\vec{\theta} \mid \mathbf{X}, \mathbf{Y}) = \delta(\vec{\theta} - \vec{\theta}_{opt}). \tag{2.21}$$

Dirac's delta function satisfies the criteria for probability density functions, as it is non-negative and its integral over the parameter space ν_Θ is unity. Insertion into Eq. (2.20) gives

$$\begin{aligned} f_{\vec{X}_0|\vec{Y}_0,\vec{X},\vec{Y}}(\vec{x}_0 \mid \vec{y}_0, \mathbf{X}, \mathbf{Y}) &\propto \int_{\nu_\Theta} f_{\vec{Y}_0|\vec{X}_0,\vec{X},\vec{Y},\vec{\Theta}}(\vec{y}_0 \mid \vec{x}_0, \mathbf{X}, \mathbf{Y}, \vec{\theta}) \pi_{\vec{X}_0}(\vec{x}_0) \delta(\vec{\theta} - \vec{\theta}_{opt}) d\vec{\theta} \\ &\propto f_{\vec{X}_0|\vec{Y}_0}(\vec{x}_0 \mid \vec{y}_0) \pi_{\vec{X}_0}(\vec{x}_0). \end{aligned} \tag{2.22}$$

The first part, $f_{\vec{X}_0|\vec{Y}_0}(\vec{x}_0 \mid \vec{y}_0)$, is called response or likelihood function and the second part, $\pi_{\vec{X}_0}(\vec{x}_0)$, is the climate prior function.
How the likelihood functions are estimated, is addressed in the following sections.

2.2.3 Mixture models as transfer functions

Several statistical models are thinkable as statistical transfer functions. In his thesis Gebhardt (2003) examined the practicability of normal distributions, kernel densities and mixture models. The work of Kühl et al. (2002) showed that transfer functions are estimated sufficiently by normal distributions for many taxa. For some important taxa, though, normal distributions lack of flexibility as not all information can be captured by mean and covariance matrix. Kernel densities and mixture models both are able to do so.
Gebhardt found a lot of kernel densities to be skewed and therefore having a different maximum than the respective normal distribution. Given the skewness to be reasonable ecological information, kernel densities provide better approximations of the transfer functions than normal distributions. On contrary he judged the non-parametric nature of kernel densities being a disadvantage because it makes them inconvenient for mathematical operations and difficult to store.
Differences between kernel densities and mixture models were found to be small. As mixture models yet have the advantage of a parametric nature, Gebhardt decided for them as best choice for the transfer functions.

Mixture models are normally used in cluster analysis, e.g. Fraley and Raftery (2002) or Vrac et al. (2005), to identify different clusters in a complex dataset. Here we are not interested in the clusters itself but use the ability of mixture models to provide good images of the dataset as a whole.
The mixture model approximates the probability density function for the modern climate vector $\vec{x}$ given the occurrence of taxon y_k

$$f_{\vec{X}|Y_k}(\vec{x} \mid y_k) = \sum_{k=1}^{M_{opt}} \alpha_k \cdot f_{N(\vec{\mu_k},\Sigma_k)}(\vec{x}). \tag{2.23}$$

Here M_{opt} is the optimal number of mixture model components. Its determination is a difficult task, that will be described later in this chapter. The EM-algorithm (Dempster et al., 1977) was already used by Gebhardt (2003) to estimate the mixture model parameters. He implemented the algorithm after Smyth et al. (1999) who used mixture models to identify Northern Hemispheric circulation regimes.

For the description of the EM-algorithm procedure we regard a dataset $\mathbf{X} = \{\vec{x_1}, ..., \vec{x_N}\}$ of N independent observations of random variable $\vec{X}$. The EM-algorithm maximises iteratively the log-likelihood function of parameter vector $\hat{\Phi} = \{\hat{\mu}_1, \hat{\Sigma}_1, \hat{\alpha}_1, ..., \hat{\mu}_{M_{opt}}, \hat{\Sigma}_{M_{opt}}, \hat{\alpha}_{M_{opt}}\}$, given by

$$L^{M_{opt}}(\hat{\Phi} \mid \mathbf{X}) = \sum_{n=1}^{N} \log \left(\sum_{k=1}^{M_{opt}} \hat{\alpha}_k \cdot f_{N(\hat{\mu}_k, \hat{\Sigma}_k)}(\vec{x}_n) \right). \tag{2.24}$$

Now, the likelihood that data point $\vec{x}_n$ belongs to the class $w_k (k = 1, ..., M_{opt})$ at iteration r is calculated from the parameters, estimated at iteration $r-1$:

$$\hat{P}^r(w_k \mid \vec{x}_n) = \frac{\hat{\alpha}_k^{(r-1)} f_{N(\hat{\mu}_k^{(r-1)}, \hat{\Sigma}_k^{(r-1)})}(\vec{x}_n)}{\sum_{i=1}^{M_{opt}} \hat{\alpha}_i^{(r-1)} f_{N(\hat{\mu}_i^{(r-1)}, \hat{\Sigma}_i^{(r-1)})}(\vec{x}_n)}. \tag{2.25}$$

This is called the expectation (E)-step of the EM-algorithm. In the maximisation (M)-step the parameters of iteration r are estimated as

$$\hat{\alpha}_k^r = \frac{1}{N} \sum_{n=1}^{N} \hat{P}^r(w_k \mid \vec{x}_n), \tag{2.26}$$

$$\hat{\vec{\mu}}_k^r = \frac{\sum_{n=1}^{N} \hat{P}^r(w_k \mid \vec{x}_n) \vec{x}_n}{\sum_{n=1}^{N} \hat{P}^r(w_k \mid \vec{x}_n)}, \tag{2.27}$$

$$\hat{\Sigma}_k^r = \frac{\sum_{n=1}^{N} \hat{P}^r(w_k \mid \vec{x}_n)(\vec{x}_n - \hat{\vec{\mu}}_k^r)(\vec{x}_n - \hat{\vec{\mu}}_k^r)^T}{\sum_{n=1}^{N} \hat{P}^r(w_k \mid \vec{x}_n)}. \tag{2.28}$$

As far as M_{opt} is known, the procedure of estimating the mixture model parameters is not very complicated. Although the EM-algorithm only guarantees convergence to some maximum that has not necessarily to be a global one, a reasonable initialisation avoids this problem.

2.2.4 The optimal number of components

Unfortunately the estimation of the optimal number of components is not that easy as it might seem. In principle a higher number of mixture model components offer a higher degree of flexibility to the mixture model and thus should be an advantage. The difference between a mixture model with a low number of components ($M_{opt} = 2$) and one with a relatively high number ($M_{opt} = 5$) can be seen in Fig. (2.2). The model with less components shows a smoother fit with the maximum density lying more in the centre of the data

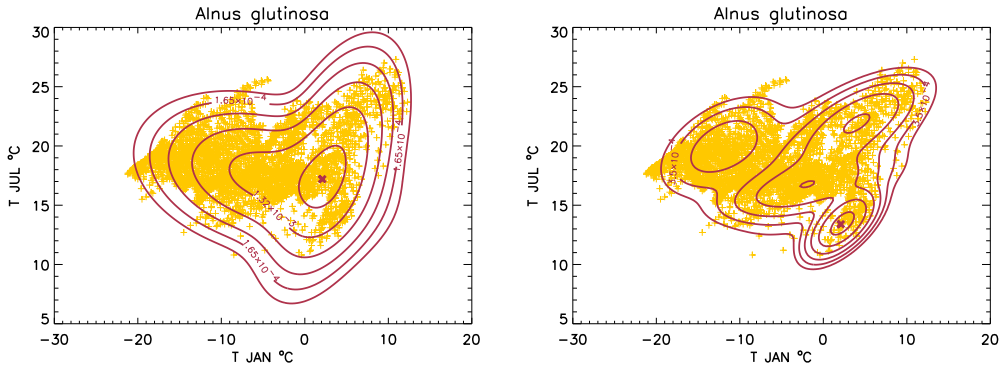

Figure 2.2: Mixture model fit to occurrence data of *Alnus glutinosa* (orange crosses), two model components (left) and five model components (right). Maxima are indicated by a red X, respectively.

points. With five components, signs of overfitting occur. So the maximum is shifted to an isolated cluster at the lower right corner. From an ecological point of view it is not very convincing that January temperatures of 2°C and July temperatures of 13°C should be more likely than slightly higher July and slightly lower January temperatures. Especially as the taxon *Alnus glutinosa* obviously exists very frequently at lower winter temperatures. From the view of maximum log-likelihood, the fit with more components would be the better, indicating that the log-likelihood might not be the best criterion for estimating the optimal number of components.
Generally, when estimating M_{opt}, one has to make a compromise between a close fit to the data on the one hand and avoiding overfitting and isolated clusters on the other hand. Gebhardt (2003) used a crossvalidation approach after Smyth (2000) to determine the optimal number of components. In this approach the dataset to be fitted is divided into two parts, one for training and one for verification. Then the mixture model parameters are estimated based on the training dataset while the corresponding likelihood for different numbers of components is calculated on the verification dataset. This procedure is repeated several times by crossvalidation and finally the sum over the likelihood of all crossvalidated datasets is used as criterion for the optimal number of components. Further the maximum was constrained to $M_{opt} = 4$. It turned out that nearly 84% of all regarded taxa received four components as optimal number and overfitting remained a problem.
Therefore Gebhardt adjusted the EM-algorithm by giving a lower limit for the smallest eigenvalue of the component covariance matrix. A small eigenvalue means that the component has become very narrow, at least in one

dimension, and this is something that has to be avoided. With the EM-algorithm modified in this way still nearly 70% of the taxa got $M_{opt} = 4$ as optimal component number.
While including precipitation in the fitting procedure of the transfer functions it became clear that calculation time increases dramatically. This is related to the longer calculation time of the three dimensional EM-algorithm in every crossvalidation sample.

A different criterion for model selection

In the field of information technology and machine intelligence, there exist a variety of publications about clustering and selecting the optimal number of clusters. Many authors like Fraley and Raftery (2002, 2007) or Hu and Xu (2004) suggest using the Bayesian Information Criterion (BIC) for determining M_{opt}. Moreover Pernkopf and Bouchaffra (2005) propose to use the Minimum Description Length (MDL), which is formally equivalent to the BIC, given by

$$BIC = 2l_M(x, \hat{\theta}) - m_M \log(n). \tag{2.29}$$

Here $l_M(x, \hat{\theta})$ is the maximised log-likelihood for the model M, m_M is the number of independent parameters to be estimated and n the number of data points. The advantage of using the BIC is quite clear: It does not focus alone on maximising the log-likelihood. This is illustrated in Fig. (2.3). The log-likelihood reaches its absolute maximum at five components while the BIC has two similar maxima at two and five components. Taking into account the number of degrees of freedom and the size of the dataset itself obviously leads to higher values of the BIC on low component numbers, compared to the log-likelihood. For reducing the danger of overfitting it seems to be a pragmatical procedure to choose the maximum with the lower number of components. In the illustrated case one would regard $M_{opt} = 2$ as optimal.
To sum this up briefly, the BIC first has the advantage of higher values at low component numbers. So it enhances the probability of choosing a lower number of components. Second, compared to evaluating the loglikelihood by crossvalidation, the procedure of using the BIC is computational much faster.

Final efforts to avoid overfitting

Using the BIC as model selection criterion avoids overfitting from the outset in many cases, when a lower number of components is chosen. However, there remain still a considerable number of cases, where a relative high M_{opt} is chosen. In these cases overfitting still can occur. Overfitting always is a

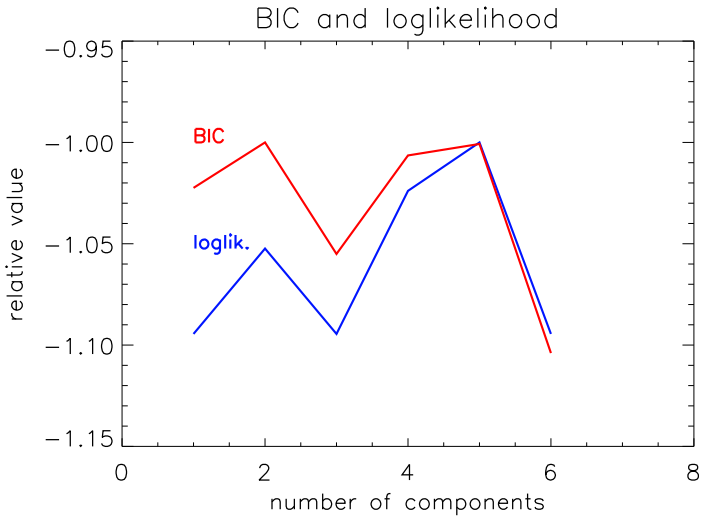

Figure 2.3: Values of BIC and log-likelihood against the number of mixture model components for *Alnus glutinosa*.

problem of a cluster becoming too small during the iteration process of the EM-algorithm. The worst case in this context would be a singularity, meaning that one cluster contains only one single datapoint. When this happens the EM-algorithm crashes, because no estimation of a covariance matrix is possible. However it is also a problem, when the algorithm reaches convergence with a small cluster that gets a high weight. From the ecological point of view this leads to an unrealistic transfer function. From the mathematical point of view it could mean that the EM-algorithm only has found a local maximum. The ability of the algorithm to find the global maximum can be related to its initialisation. So, for example, an initialisation by random starting values might be problematic. For avoiding this problem, Lee et al. (2006) propose using k-means clustering to find the initial values for the EM-algorithm. A somewhat more detailed description of the k-means clustering algorithm is to be found in Likas et al. (2003).

In Fig. (2.4) an example of the initial k-means clustering is displayed. The clustering generates clusters of similar size, resulting in nearly equidistant cluster centres. For the EM-algorithm the cluster centres are used as initial means and the initial covariance matrices are estimated on basis of the initial clusters. Thus, it is unlikely for the EM-algorithm to run into a singularity or not to find the global maximum of the log-likelihood function.

Now, all arrangements for avoiding overfitting and getting the best transfer

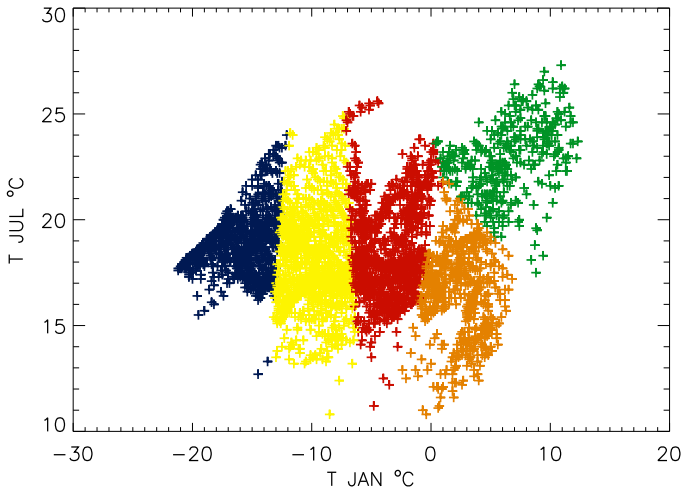

Figure 2.4: Initial k-means clustering for *Alnus glutinosa* with five clusters.

function possible, should be shortly summed up. First the maximum number of model components should be constrained to four, as a higher number of components nearly always produces overfitting. Further the BIC should be taken as model selection criterion because it has the tendency to select fewer clusters. For avoiding singularities and not reaching local maxima during the iteration process, the EM-algorithm is initialised by k-means clustering.

This course of action produces satisfying results in most but not in all cases. Some taxa remain difficult to fit and overfitting still occurs. The only solution for these cases still is giving a lower bound for the smallest eigenvalue of the component covariance matrices, as mentioned at the beginning of this section. The general idea was limiting the eigenvalue to not falling below the smallest eigenvalue of the initial clusters: $\lambda_{min} \geq \lambda_{min}^{init}$. It can be seen in Fig. (2.5) that the taxon *Empetrum nigrum* is a good example for this limit not automatically leading to a better solution. The mixture model fit without any limitations reveals a multimodal distribution with the maximum shifted to warm winter temperatures. With the constraint $\lambda_{min} \geq \lambda_{min}^{init}$ the result becomes even worse. The maxima of the respective clusters are more pronounced and there is no improvement concerning the position of the global maximum. A better result can be achieved in this case by enhancing the limit to $\lambda_{min} \geq 2 \cdot \lambda_{min}^{init}$. In all cases $M_{opt} = 4$ was chosen as optimal, while in the last case two model components get a very low weight.

Unfortunately, there seems to be no procedure for estimating M_{opt} or con-

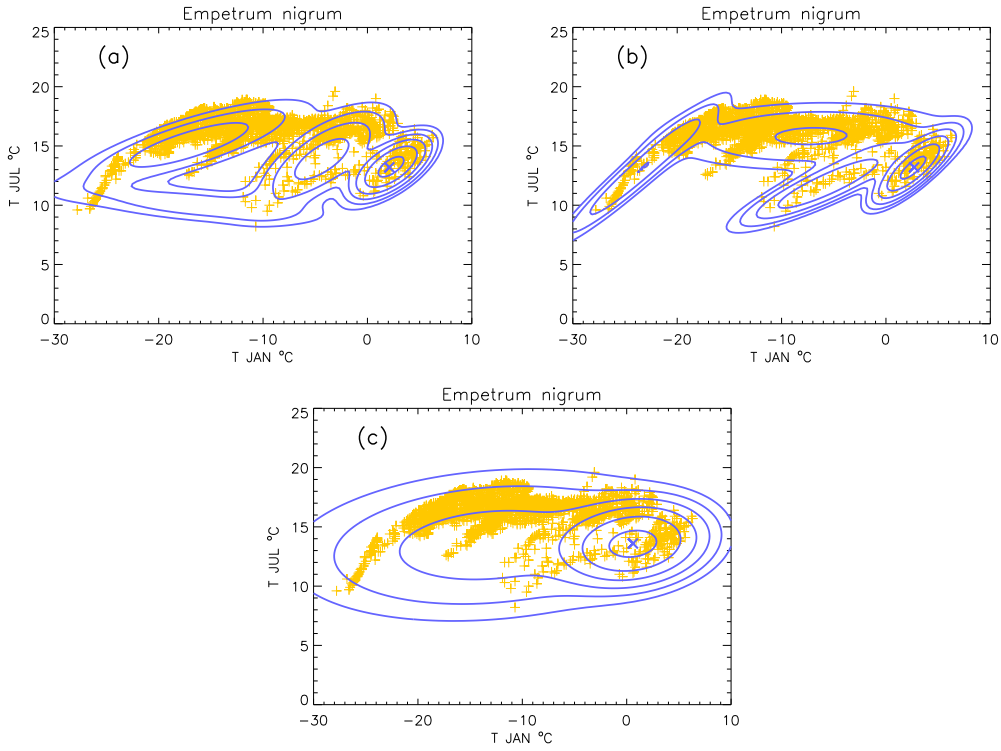

Figure 2.5: Mixture Model fit to occurrence data of *Empetrum nigrum*, without limiting the eigenvalues (a), with $\lambda_{min} \geq \lambda_{min}^{init}$ (b) and with $\lambda_{min} \geq 2 \cdot \lambda_{min}^{init}$ (c).

straining the EM-algorithm in a way that overfitting never occurs. Also it is not feasible to automatically detect overfitting. Thus, at the moment there is no alternative to visually checking the results for overfitting. The limit for λ_{min} has to be adjusted for these cases until a satisfying result for the respective transfer function is achieved.

2.3 Local climate reconstructions

A local climate reconstruction normally is a point reconstruction from data, gained at a certain place. This could be a pollen sample from a sediment core, drilled in a lake, or also data from several cores of the same site. The representativeness of local reconstructions depends on several factors. So it is important what kind of data is used. While macrofossils normally originate from the direct neighbourhood of a location, some pollen are able to

be transported over long distances of about 100 km or even more. Thus, if only macrofossils are used, the reconstruction represents the climate exactly at the given location. If pollen are used one has to regard the reconstruction as representative for a box of 100 or 200 km edge length.
A lot of fossil data comes from lake sediments. Here the size of the lake, normally connected with the size of its catchment area plays a role. Both, pollen and macrofossils can be transported by rivers to the respective lake. So also the size of a lake has to be taken into account when thinking about the representativeness.
What remains somewhat problematic are locations in the mountains. Here different climatic regions vary rapidly with altitude and thus on small horizontal scales. Pollen data have to be used with care here, as they can distort the result when coming from another altitude as the one of the location.
When it is clear which taxa can be used, the local reconstruction is done with the Bayesian indicator taxa model (Schölzel, 2005). This model is based on the assumption, that the likelihood functions depending on the modern variables are equal to their palaeo counterparts. Therefore it is not distinguished between palaeo and recent variables in the notation from now on. We regard a discrete random vector $\vec{Y} = (Y_1, ..., Y_n)$ with values $\vec{y} \in \{0, 1\}$ representing all n taxa. A value $y_k = 1$ represents the presence of a taxon while $y_k = 0$ indicates its absence. Thus, at this point it is possible to regard the problem in terms of presence or absence of taxa. Inserting this into the Bayes theorem (Eq. (2.17)) gives

$$f_{\vec{X}|\vec{Y}}(\vec{x} \mid y_1, ..., y_n) = \frac{f_{\vec{Y}|\vec{X}}(y_1, ..., y_n \mid \vec{x}) \cdot \pi_{\vec{X}}(\vec{x})}{m_{\vec{Y_0}}(y_1, ..., y_n)}. \tag{2.30}$$

When statistical independence can be assumed, the likelihood function can be rewritten by the multiplication rule:

$$f_{\vec{Y}|\vec{X}}(y_1, ..., y_n \mid \vec{x}) = \prod_{k=1}^{n} f_{Y_k|\vec{X}}(y_k \mid \vec{x}). \tag{2.31}$$

Some discussion about the statistical independence of different taxa can be found in the data chapter.
Now a final application of the Bayes Theorem leads to a model that allows for the calculation of the posterior distribution by estimating the recent conditional pdf of climate given the existence of each taxon. Based on the previous equations the final equation for the posterior distribution is given by

$$f_{\vec{X}|\vec{Y}}(\vec{x} \mid y_1, ..., y_n) = \frac{\pi_{\vec{X}}(\vec{x})}{m_{\vec{Y}}(y_1, ..., y_n)} \cdot \prod_{k=1}^{n} \frac{f_{\vec{X}|\vec{Y}}(\vec{x} \mid y_k) \cdot \pi_{Y_k}(y_k)}{m_{\vec{X}}(\vec{x})}. \tag{2.32}$$

As pointed out by Schölzel (2005) from the mathematical point of view it would be possible to work with absence of taxa as well as with presence. However there are some good reasons for not doing so. The absence of taxa can have several different reasons than climate. For instance, soil conditions can be inappropriate or plant competition can be too strong. Concerning pollen, the pollen production rate of a plant can be too low for being represented in the sediment. When macrofossils are used it can be a problem, that they are normally not transported over longer distances. So the taxon could have been present at a certain site but nevertheless is not present everywhere in the sediment.
Based on these considerations the posterior pdf is rewritten as a pdf only conditional on the present taxa. These are formally the taxa $Y_{i(\tilde{k})}$ with

$$i(\tilde{k}) \in \{1, ..., n_k\} : y_{i(\tilde{k})} = 1. \tag{2.33}$$

In the final step of the Indicator taxa model Schölzel (2005) introduced a general condition C for further random variables like e.g. unknown soil properties. Then he stated that the climate state vector given the condition C should be conditionally independent to the absence of taxa ($Y_k = 0$). Then he omitted all terms not dependent on the climate state vector and it remains

$$f_{\vec{X}|Y_{i(1)},...,Y_{i(n_{\tilde{k}})}}(\vec{x} \mid 1, ..., 1, C) \propto \pi_{\vec{X}}(\vec{x}) \cdot \prod_{k=1}^{n} \frac{f_{\vec{X}|\vec{Y}_{i(\tilde{k})}}(\vec{x} \mid 1, C)}{m_{\vec{X}}(\vec{x})}. \tag{2.34}$$

Now there are three important terms left that have to be quantified for the reconstruction. These are the climate prior $\pi_{\vec{X}}(\vec{x})$, the marginal climate distribution $m_{\vec{X}}(\vec{x})$ and the likelihood functions $f_{\vec{X}|\vec{Y}_{i(\tilde{k})}}(\vec{x} \mid 1, C)$ which are the transfer functions, described in the previous sections. The others are specified briefly below.

The marginal climate distribution

The marginal climate distribution describes the observed climate state vector of the modern climate, which is used to estimate the transfer functions. The presence of this distribution in the Indicator taxa model is necessary as it is not an uniform distribution, meaning that the different climate values in the area of observation are not equally likely. Thus the estimated likelihood functions are also dependent on the frequency of occurrence of different climate states.
This problem can be addressed in different ways. One possibility would be to

estimate the marginal distribution directly from the climate data. Another way, used by Schölzel (2005), is to calculate the marginal distribution as mixture of all likelihood functions. Finally it is also possible to give a certain weight to the climate values, depending on the frequency of their occurrence. How this weighting is done exactly will be described in the data chapter.

The climate prior

The climate prior can have a substantial influence on the reconstruction result. The classical Bayesian philosophy is, that one has always some kind of prior knowledge which is modified by new or updated data. This can also be seen as learning from data.
If there is prior knowledge, informative (or subjective) priors are used. These could be any kind of probability density function. However it makes sense, when dealing e.g. with temperatures where we have to do with normally distributed likelihoods, that also the priors are normally distributed. Depending on the level of prior knowledge one can choose a broad pdf, reflecting weak prior knowledge. On the other hand, a narrow pdf would allow for less influence of the likelihood and thus reflect a strong prior knowledge.
When there is no prior knowledge, the so called uninformative prior is chosen. This means that the prior distribution is set to an uniform distribution. Different approaches have been used in previous studies. Gebhardt (2003) leaved it at the uninformative prior, while Schölzel (2005) chose an informative prior based on typical Holocene climate variations. As his reconstructions where confined to the Holocene this was a logical approach. However when reconstructing the climate of the Late Glacial, the uninformative prior might be the better choice.

Chapter 3

Reconstruction of fields

Point reconstructions are able to provide information about the climate of the past only for a limited area. For some studies this might be satisfying but in most cases one is interested in reconstructions for a larger area, at least for certain countries, mostly for (parts of) continents and sometimes even for the whole northern hemisphere.
In these cases some studies leave it at showing a compilation of several point reconstructions together with a more or less subjective interpretation, like Tarasov et al. (1999) or Wu et al. (2007). In other studies like Zagwijn (1996) or Kuhlemann et al. (2008) isolines between sites are drawn, which are often not based on a quantitative analysis and thus are not reproducible. Further, such an approach can lead to very strange patterns, e.g. in the case of outliers.
At least in some studies, quantitative interpolation techniques were used. So Cheddadi et al. (1997) applied a non-linear interpolation method based on neural networks. Another approach of Davis et al. (2003) used a 4D-interpolation technique based on smoothing splines.
A method for field reconstruction should have several characteristics as mostly already mentioned by Gebhardt (2003):

- It should be objective, quantitative and thus reproducible.
- It should compensate outliers or heterogeneous results from point reconstructions caused by local peculiarities of the sites.
- The information about uncertainties of the point reconstructions provided by the statistical reconstruction approach should be taken into account.
- Additional information about spatial properties of the reconstructed quantity should be able to be taken into consideration.

- The reconstruction should provide results on scales, climate model simulations are carried out. This would allow for comparison to climate model output.

Especially the point about including additional information is very interesting. It can address the fact that our knowledge about paleoclimatologic fields is not restricted to the available proxy data. Beyond that there is also knowledge about physical laws on which these fields depend. Therefore Gebhardt (2003) imposed a simple stationary advection-diffusion model to his analysis. As analysis tool he applied a variational analysis technique and he successfully reconstructed temperature fields for the Eemian interglacial (Gebhardt et al., 2008). The inclusion of additional information is a substantial difference to the interpolation techniques mentioned above, as those are based on the proxy data alone.

3.1 Variational Analysis

It is now more than 50 years ago that the variational analysis was first mentioned in the field of meteorological applications by Sasaki (1958). The same author introduced the variational analysis with strong and weak constraint about a decade later (Sasaki, 1970). However, it took another fifteen years till the method became more important in the field of meteorology. This finally happened with the rise of numerical modelling. The first investigations highlighted the theoretical background and several practical issues, like Le Dimet and Talagrand (1986), Courtier and Talagrand (1987), Thacker and Long (1988), Stauffer and Seaman (1994) or Zhu and Kamachi (2000). Later, when more computational power was available and numerical weather prediction became more and more important, powerful data assimilation procedures were necessary. On the basis of the variational analysis, three dimensional (3D-VAR) methods, e.g. Andersson et al. (1998) or Gauthier et al. (1999), and four dimensional (4D-VAR) methods, e.g. Thépaut and Courtier (1991) or Li et al. (2000), were developed. In the field of ozone analysis, Eskes et al. (1999) and Elbern and Schmidt (2001) applied the variational analysis as well as Gebhardt and Hense (2001) for the analysis of climatological data.

The variational analysis is based on setting up a cost function quantifying the inconsistency between data which have to be assimilated into a given field and the field itself. When the field is calculated with respect to a constraint like the advection-diffusion model, the cost function quantifies the inconsistency between data and the constraint. During the analysis the cost function

is minimised and thus the best compromise between data and constraint is achieved. To be precise this would be the case when a weak constraint is used and the constraint has only to be approximately fulfilled. When the constraint has to be fulfilled exactly, it is called a strong constraint.
The variation analysis for reconstructing temperature fields is presented by Gebhardt (2003) in a detailed description. Therefore its presentation here will be restricted to the aspects that have relevance for this work. Further, a possible constraint for the reconstruction of precipitation fields will be introduced.

3.1.1 Specification of the cost function

In terms of reconstructing fields of climatological quantities the variational analysis is used to assimilate the local reconstructions into a field based on the advection-diffusion model. The cost function quantifies the inconsistency of either the local reconstruction and the constraint to the reconstructed field. Thus the cost function consists of a part representing the local reconstructions based on vegetation data, hence called "vegetational part". The other part represents the advection-diffusion model and is called "model part".
Let Ω be the area of analysis, then vegetational and model part are defined as:

$$J_{veg} = \int_{\Omega} H_{veg} d\Omega, \tag{3.1}$$

$$J_{mod} = \int_{\Omega} H_{mod} d\Omega. \tag{3.2}$$

Here H_{veg} and H_{mod} just represent the costs at the different points of the analysis area and remain unspecified for the moment. For practical reasons the analysis is performed in terms of anomalies $\vec{x}'(\vec{r})$, given by

$$\vec{x}' := \vec{x}_{past}(\vec{r}) - \vec{x}_0(\vec{r}). \tag{3.3}$$

Note that the modern values $\vec{x}_0(\vec{r})$ are represented by the mean of the period 1961-90.
The decision to use anomalies despite absolute values is based on the fact that some inhomogeneously distributed sites can only resolve large-scale patterns. Fields of absolute values, though, contain small scale features, especially in mountainous regions. One can expect anomalies to be much smoother and thus better representable by the available data.
The usage of anomalies involves the implicit assumption that the modern values can serve as first guess for the reconstruction. When a deviation of the analysed quantity from a first guess is analysed, it is called incremental

approach (Courtier et al., 1994). This approach is common also in numerical weather prediction and was used by Courtier et al. (1998) or Gauthier et al. (1999) within their 3D-VAR systems.
When large-scale increments are analysed, the mathematical dimension of the problem often can be reduced. In these cases the minimisation of the cost function is computationally less expensive.

The cost function, consisting of vegetational- and model-part, finally is given by

$$J[\vec{x}'(\vec{r})] = \int_\Omega \left\{ H_{veg}[\vec{x}'(\vec{r}), \vec{r}] + \gamma_M \cdot H_{mod}[\vec{x}'(\vec{r}), \vec{r}] \right\} d\Omega. \tag{3.4}$$

The parameter γ_M is called regularisation parameter (Gebhardt et al., 2008). It serves for controlling the influence of the constraint, meaning that the model is implemented as a weak constraint. The value of γ_M has to be chosen carefully, as it has substantial influence on the analysis result. In this work, always a value is chosen which leads to equal contribution to the total costs from both parts.
At the end of the minimisation process, one receives the anomaly field $\vec{x}'_R(\vec{r})$ which is the one that produces the lowest costs. Now the sum of $\vec{x}'_R(\vec{r})$ and $\vec{x}_0(\vec{r})$ is regarded as the reconstructed field of absolute values.

3.1.2 Vegetational costs

The vegetational costs quantify the deviation of the actual anomaly field to the paleobotanical data. With the Bayesian Indicator taxa model (Eq. (2.34)), there is already a tool for translation of paleobotanical data into climatological values. Although Gebhardt (2003) defined the vegetational part of the cost function independently from the Indicator taxa model, a formulation with the Indicator taxa model is equivalent. The vegetational part was defined by Gebhardt as

$$\begin{aligned} J_{veg}[\vec{x}'(\vec{r})] &= \int_\Omega H_{veg}[\vec{x}'(\vec{r}), \vec{r}] d\Omega \\ &= -\int_\Omega ln\left[f_{\vec{X}|Y_1,\ldots,Y_n}(\vec{x}'(\vec{r}) + \vec{x}_0(\vec{r}) \mid y_1(\vec{r}), \ldots, y_n(\vec{r})) \right] d\Omega + \beta. \end{aligned} \tag{3.5}$$

This is the formulation of the vegetational costs in continuous space, which assumes that paleobotanical data are available at every point of the analysis area Ω. Of course this is not a realistic assumption, but as the analysis will be discretised later, it is enough to regard the case of sparse vegetational observations later. The parameter β contains all constant values, which can

be skipped, as they do not influence the minimisation process.
The smallest value of the vegetational costs is reached in case of $f_{\vec{X}|Y_1,...,Y_n}(\vec{x}'(\vec{r})+\vec{x}_0(\vec{r}) \mid y_1(\vec{r}),...,y_n(\vec{r})) = 1$. This would mean that the analysed field is absolutely consistent with the paleobotanical data.
Now, insertion of Eq. (2.34) leads to

$$J_{veg}[\vec{x}'(\vec{r})] = -\int_{\Omega} ln\left[\prod_{i=1}^{n} \frac{f_{\vec{X}|Y_i}(\vec{x}'(\vec{r}) + \vec{x}_0(\vec{r}) \mid y_i(\vec{r}))}{m_{\vec{X}}(\vec{x})}\right] d\Omega, \qquad (3.6)$$

in the case of uninformative priors (priors set to uniform distribution). Both, likelihood-function $f_{\vec{X}|Y_i}$ and marginal distribution $m_{\vec{X}}$ are based on the modern (1961-90) climate. Thus, it has to be assumed that the relation between climate and vegetation is the same today and in the period whose climate has to be reconstructed. Further it has to be assumed that the marginal distribution of both is the same. At least the second assumption is wrong. The reconstruction of climate accompanies the possibility of a different global climate than today, meaning at least the mean of $m_{\vec{X}}$ to be different than today.
However, Gebhardt (2003) argues that the characteristics of the marginal distribution already influence the estimation of the likelihood-functions and thus their ratio is independent of $m_{\vec{X}}$.
The first assumption of a constant relation between climate and vegetation over time can be regarded as sufficiently true, when they are in equilibrium state.

3.1.3 Advection-diffusion model

After Gebhardt et al. (2008) have succeeded to reconstruct consistent temperature fields of the Eemian under the constraint of the linear advection-diffusion model, the same constraint for reconstructing temperature fields is also used in this work. It should be noted clearly that the model only serves as weak constraint. It only introduces additional information about atmospheric thermodynamics into the analysis. So the result is dynamically consistent and stabilised. The crucial information for the reconstruction result always comes from the paleobotanical data.
The advection-diffusion model is a simplified linear model, introduced to the field of meteorology about 30 years ago by Egger (1977) and Webster (1981). It turned out to be a useful tool for analysis of climatological and GCM data during the work of Hense et al. (1990) and Klaßen et al. (1994). It might be, and also turned out to be, a strange idea to some people to apply such a model to climate reconstructions. However, when a model can

be successfully applied to analysing modern climatological data, there is no fundamental reason for not doing so with historical climatological data.
The derivation of the advection-diffusion model shall be presented here briefly. It starts with the assumption that the temporal rate of change of the temperature $\frac{dT}{dt}$ is controlled by a spatial forcing field $Q(\vec{r})$. This leads to the basic equation

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \nabla(\vec{v}T) = Q. \tag{3.7}$$

Here the advection term $\vec{v}\nabla T$ was replaced by the temperature flux divergence $\nabla(\vec{v}T)$ under the assumption of a non-divergent, two dimensional flow. Now we regard two realisations of this equation, for T_1 and T_2, respectively:

$$\frac{dT_1}{dt} = \frac{\partial T_1}{\partial t} + \nabla(\vec{v_1}T_1) = Q_1, \tag{3.8}$$

$$\frac{dT_2}{dt} = \frac{\partial T_2}{\partial t} + \nabla(\vec{v_2}T_2) = Q_2. \tag{3.9}$$

Calculating the difference between Eq. (3.8) and Eq. (3.9) gives

$$\frac{dT_1}{dt} - \frac{dT_2}{dt} = \frac{\partial T_1}{\partial t} - \frac{\partial T_2}{\partial t} + \nabla(\vec{v}_1T_1 - \vec{v}_2T_2) = Q_1 - Q_2. \tag{3.10}$$

This equation can be split up to

$$\begin{aligned} \frac{dT_1}{dt} - \frac{dT_2}{dt} &= \frac{\partial T_1}{\partial t} - \frac{\partial T_2}{\partial t} + \nabla\left[\frac{1}{2}(\vec{v}_1 + \vec{v}_2)\cdot(T_1 - T_2)\right] \\ &+ \nabla\left[(\vec{v}_1 - \vec{v}_2)\cdot\frac{1}{2}(T_1 + T_2)\right]. \end{aligned} \tag{3.11}$$

Now we define an anomalous temperature, horizontal flow and forcing by $T' = T_1 - T_2$, $\vec{v}' = \vec{v}_1 - \vec{v}_2$ and $Q' = Q_1 - Q_2$. Further the averages of horizontal flow and temperature are given by $\bar{\vec{v}} = \frac{1}{2}(\vec{v}_1 + \vec{v}_2)$ and $\bar{T} = \frac{1}{2}(T_1 + T_2)$. Inserting these definitions into Eq. (3.11) leads to

$$\frac{dT'}{dt} = \frac{\partial T'}{\partial t} + \nabla(\bar{\vec{v}}T') + \nabla(\vec{v}'\bar{T}) = Q'. \tag{3.12}$$

Now the assumption is made that the gradient of the anomalous temperature T' is balanced by the anomalous flow $\vec{v}'$ with transport coefficient k:

$$\vec{v}' = -k\nabla T'. \tag{3.13}$$

Insertion leads to

$$\frac{\partial T'}{\partial t} + \nabla(\bar{\vec{v}}T') - \nabla(\bar{T}k\nabla T') = Q'. \tag{3.14}$$

Parametrising the third term as diffusion term leads to the advection-diffusion equation:

$$\frac{\partial T^{'}}{\partial t} + \nabla(\bar{\vec{v}}T^{'}) - \alpha\nabla^2 T^{'} = Q^{'}. \tag{3.15}$$

It represents a linear model, forced by an anomalous heating $Q^{'}$ that induces a local change of temperature with time, an advection of $T^{'}$ with the mean flow and a sub-scale diffusive transport given by $\alpha\nabla^2 T^{'}$.
Within the reconstruction of temperatures, the assumption can be made that one deals with long-time temperature averages. The modern temperatures are mean values over 30 years and the paleotemperatures represent time slices of several hundred years. Thus, $T^{'}$ is assumed to be stationary, meaning $\frac{\partial T^{'}}{\partial t} = 0$. Finally, the anomalous heating can be split up into a known part (e.g. solar insolation) and an unknown part ϵ:

$$\nabla(\bar{\vec{v}}T^{'}) - \alpha\nabla^2 T^{'} = Q^{'}_{sol} + \epsilon. \tag{3.16}$$

To split up the anomalous heating like this is justified, as solar insolation is agreed to be the most important driver of climate changes in the past (Muscheler et al., 2004).
Now, how is Eq. (3.16) implemented into the variational analysis? First, it is assumed that ϵ remains small, compared to the other terms. The spatial structure of $T^{'}$ is basically forced by $Q^{'}_{sol}$, leading to high costs for anomaly fields that do not balance the solar forcing. Following this point of view, the quantity $\epsilon(\vec{r})$ is interpreted as model imbalance.
As the analysis is performed in terms of monthly mean values and will provide reconstructions for January- and July-temperatures, from now on it will be distinguished between values for January and July. So there are temperature anomalies $T^{'}_{Jan} = T_{Jan,past} - T_{Jan,0}$ and $T^{'}_{Jul}$ accordingly, which give the temperature anomaly vector $\vec{T}^{'}(\vec{r}) = (T^{'}_{Jan}; T^{'}_{Jul})$.
When the mean wind vectors $\bar{\vec{v}}_{Jan}$ and $\bar{\vec{v}}_{Jul}$ are defined, the model imbalances $\vec{\epsilon}(\vec{r}) = (\epsilon_{Jan}; \epsilon_{Jul})$ can be calculated. These can be regarded as unbiased, Gaussian distributed model errors with mean zero and error variances $\sigma^2_{M,Jan}(\vec{r}), \sigma^2_{M,Jul}(\vec{r})$. Finally the model error covariance matrix is given by

$$\Sigma_M(\vec{r}) = \begin{pmatrix} \sigma^2_{M,Jan} & c_M \\ c_M & \sigma^2_{M,Jul} \end{pmatrix}, \tag{3.17}$$

where c_M is the model error covariance. With these definitions the contribution of the temperature anomaly $\vec{T}^{'}(\vec{r})$ at $\vec{r}$ to the model costs is

$$H_{mod}(\vec{T}, \vec{r}) = \vec{\epsilon}^{T}(\vec{r})\Sigma^{-1}_M(\vec{r})\vec{\epsilon}(\vec{r}). \tag{3.18}$$

The complete model part of the cost function is defined by integration over the analysis area Ω:

$$J_{mod} = \int_{\Omega} H_{mod} d\Omega = \int_{\Omega} \vec{\epsilon}^{T}(\vec{r}) \Sigma_M^{-1}(\vec{r}) \vec{\epsilon}(\vec{r}) d\Omega. \tag{3.19}$$

Some more details are presented in the section about discretisation of the variational analysis.

3.1.4 A constraint for precipitation

Additional physical information should also be used for the reconstruction of precipitation. However the advection-diffusion model cannot be applied to the reconstruction of precipitation fields. Unlike temperature, precipitation is not a certain property of the atmosphere, but an indication for other atmospherical properties.
Mainly two factors are important for the development of precipitation. First, there has to be enough humidity in the atmosphere. Second, vertical motions play a critical role. Due to the decreasing pressure with altitude, the air expands when it rises and thus cools. On the contrary, when the air is sinking, it is compressed and thus warms. Now, warm air can borrow much more water than cold air. So when an air parcel is cooling, at a certain point it reaches saturation and its water starts to condensate. Drops of water are built and finally it starts to rain. Therefore precipitation is linked with upward vertical motions in the very most cases.

In the field of atmospheric dynamics, vertical motions in mid-latitudes are described by the omega equation. Its derivation will be presented here after Holton (1992). It starts from the geostrophic energy equation with omitted diabatic heating term and the geostrophic vorticity equation. These equations are expressed in terms of the two dependent variables Φ (geopotential) and ω (vertical velocity). So the thermodynamic equation is given by

$$\frac{\partial \chi}{\partial p} = -\mathbf{V}_g \cdot \nabla \left(\frac{\partial \Phi}{\partial p} \right) - \sigma \omega, \tag{3.20}$$

where $\mathbf{V}_g$ is the geostrophical wind speed component, σ the static stability and $\chi = \partial\Phi/\partial p$. The geopotential vorticity equation is expressed as

$$\nabla^2 \chi = -f_0 \mathbf{V}_g \cdot \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) + f_0^2 \frac{\partial \omega}{\partial p}, \tag{3.21}$$

where f is the Coriolis parameter.
Now the horizontal Laplacian ∇^2 is applied to Eq. (3.20) and $\partial/\partial p$ is applied

to Eq. (3.21). This leads to

$$\nabla^2 \frac{\partial \chi}{\partial p} = -\nabla^2 \left[\mathbf{V}_g \cdot \nabla \left(\frac{\partial \Phi}{\partial p} \right) \right] - \sigma \nabla^2 \omega \tag{3.22}$$

and

$$\frac{\partial}{\partial p} \nabla^2 \chi = -f_0 \frac{\partial}{\partial p} \left[\mathbf{V}_g \cdot \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) \right] + f_0^2 \frac{\partial^2 \omega}{\partial p^2}. \tag{3.23}$$

The order of operators in these both equations may be reversed. So χ is eliminated by subtracting Eq. (3.22) from Eq. (3.23), giving

$$-f_0 \frac{\partial}{\partial p} \left[\mathbf{V}_g \cdot \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) \right] + f_0^2 \frac{\partial^2 \omega}{\partial p^2} + \nabla^2 \left[\mathbf{V}_g \cdot \nabla \left(\frac{\partial \Phi}{\partial p} \right) \right] - \sigma \nabla^2 \omega = 0.$$

Rearrangement of this equation leads to the omega equation:

$$\left(\nabla^2 + \frac{f_0^2}{\sigma} \frac{\partial^2}{\partial p^2} \right) \omega = \frac{f_0}{\sigma} \frac{\partial}{\partial p} \left[\mathbf{V}_g \cdot \nabla \left(\frac{1}{f_0} \nabla^2 \Phi + f \right) \right] + \frac{1}{\sigma} \left[\mathbf{V}_g \cdot \nabla \left(\frac{\partial \Phi}{\partial p} \right) \right]. \tag{3.24}$$

The omega equation does not contain any derivatives in time, making it a diagnostic equation. The first term on the right hand side of the equation describes the influence of differential vorticity advection to ω, as the geostrophic vorticity can be expressed as $\zeta_g = \frac{1}{f_0} \nabla^2 \Phi$.
The second term on the right hand side represents the temperature advection, as the change of geopotential with pressure is directly related to the temperature.

For the purpose of reconstructing climate it is not advisable to make an assumption of the geopotential field in the past, as this would be more or less pure guessing. So the right side remains unspecified and is represented by Q, leading to

$$\left(\nabla^2 + \frac{f_0^2}{\sigma} \frac{\partial^2}{\partial p^2} \right) \omega = Q. \tag{3.25}$$

Now for the treatment of the second derivative with pressure, a simple two layer model (Fig. (3.1)) is applied. This model is the most simple discretisation possible and allows for direct calculation of the second derivative. For doing that, a sinus-shaped profile of vertical velocity is assumed as

$$\omega = \omega_h(\vec{r}) \cdot sin \left(2\pi m \frac{p}{p_0} \right). \tag{3.26}$$

Here $\omega_h(\vec{r})$ is the horizontal field of vertical velocity at the middle of the atmosphere, indicated by the dashed line in Fig. (3.1). As boundary condition it is assumed, that ω vanishes at the top at the atmosphere and is controlled

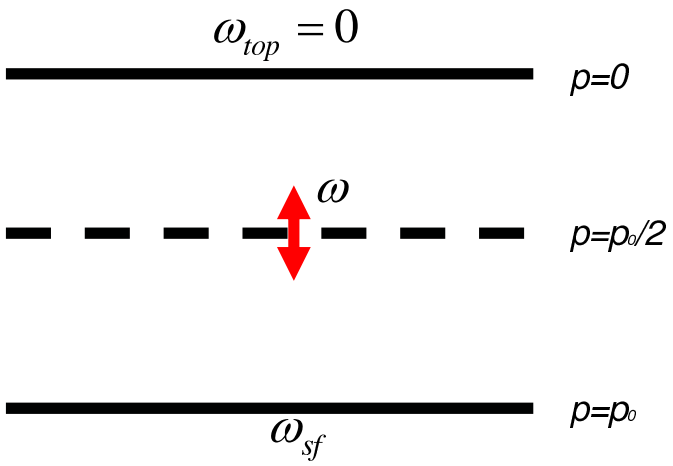

Figure 3.1: Illustration of the simple two-layer atmospheric model, applied to the omega equation.

by the orography at the bottom. The lower boundary condition can be expressed as

$$\omega_{sf} = g\rho_0 \vec{v_h} \nabla h, \tag{3.27}$$

where ∇h is the orography gradient and $\vec{v_h}$ is the horizontal wind velocity. Now, having specified the boundary conditions, the second derivative can be calculated as

$$\frac{f_0^2}{\sigma}\left(\frac{\partial^2}{\partial p^2}\right)\omega = \frac{f_0^2}{\sigma \Delta p^2}(\omega_{top} + \omega_{sf} - 2\omega). \tag{3.28}$$

Hence the final equation for the precipitation constraint (under the assumption of proportionality between precipitation and ω) reads

$$\nabla^2 \omega - \frac{2f_0^2}{\Delta p^2 \sigma}\omega = -\frac{f_0^2}{\Delta p^2 \sigma} g\rho_0 \vec{v_h} \nabla h + \epsilon', \tag{3.29}$$

leaving ϵ' for unspecified sources and defining the orography as main source of ω. Leaving other potential sources of omega unspecified, means not constraining the vertical motions too much. The assumption that vertical motions are basically driven by the orography should be applicable for mean values over several decades or hundreds of years.

3.2 Discretisation of the analysis

In the last section the variational analysis in continuous space was described. However for numerical solution of the problem, the analysis has to be discre-

tised.
First, one has to decide on which grid the discretisation will be carried out. Gebhardt (2003) chose a regular latitude-longitude grid with 1° grid spacing. He decided, a 1° grid spacing would reasonably resolve the locations with available paleodata. This is also true for all four time slices in this work. Enhancing the grid spacing to 0.5° would increase the number of grid boxes by factor four, while the grid boxes with data would only increase by an absolute number around five. The reconstruction area is different for the different time slices and will be described in the data chapter.
For discretisation on the regular grid the location vectors $\vec{r}_n, (n = 1, ..., N)$ are defined at the centres of the N grid boxes, respectively. For the analysis, January and July temperatures are summarised in an anomaly vector of length $2N$:

$$\vec{X}' = (T'_{Jan,1}, ..., T'_{Jan,N}, T'_{Jul,1}, ..., T'_{Jul,N}). \tag{3.30}$$

This is done accordingly in the three-dimensional case when precipitation is included. Then the anomaly vector is enhanced to a length of $3N$:

$$\vec{X}' = (T'_{Jan,1}, ..., T'_{Jan,N}, T'_{Jul,1}, ..., T'_{Jul,N}, P_{Ann,1}, ..., P_{Ann,N}). \tag{3.31}$$

Now the cost function (Eq. 3.4) is discretised by replacing the integral with a sum over all grid boxes:

$$J_D(\vec{X}', \vec{r}) = \sum_{n=1}^{N} \Big(H_{veg}(\vec{x}'_n(\vec{r}_n)) + H_{mod}(\vec{x}'_n(\vec{r}_n)) \Big) \Delta\Omega_n, \tag{3.32}$$

where $\vec{x}'_n$ is the vector of all anomalies at grid box n - $(T'_{Jan,n}, T'_{Jul,n}, P'_{Ann,n})$ in the three dimensional case. Multiplied at each grid box is the factor $\Delta\Omega_n = r_E^2 \Delta\lambda \Delta sin\phi_n$ containing the distance of longitudes $\Delta\lambda$ and the difference of the sines of the latitudinal bounds. This weights the influence of each grid box according to its area on the surface of the earth's sphere. As the earth's radius r_E^2 and $\Delta\lambda$ are constant factors, they can be omitted in the cost function, leading to $\Delta\Omega_n = \Delta sin\phi_n$.

Discretisation of the vegetational part

As the vegetational part of the cost function was defined for continuous space, it was implicitly assumed, that paleovegetation data were available everywhere in the analysis area. This is of course a hypothetical idealised situation that in reality never will occur. It would require paleodatabases at least every 1° in every direction over land. As very special conditions are necessary for useful preservation of vegetation material in the sediments, this

is not even theoretically possible. So in any given case within this work, only certain K grid boxes of the analysis area are filled with vegetation information. The vegetational costs are calculated as sum over these locations:

$$J_{veg}[\vec{x}^{'}(\vec{r})] = -\sum_{k=1}^{K} ln\Big[\prod_{i=1}^{n} \frac{f_{\vec{X}|Y_i}(\vec{x}_k^{'}(\vec{r}_k) + \vec{x}_0(\vec{r}_k) \mid y_i(\vec{r}_k))}{m_{\vec{X}}(\vec{x})}\Big]\Delta sin\phi_k. \quad (3.33)$$

Discretisation leads to a problem, because temperature is dependent on height. The mean altitude of an 1° grid box is normally not representative for a certain site within this box. The reconstructed value for a given location, however, is valid only for the altitude of the location. Thus, a temperature correction to the mean altitude of the grid box has to be applied. In the standard atmosphere, valid for long term means of the midlatitude atmosphere, the temperature decreases by 0.0065 K/m with increasing height. Gebhardt (2003) applied a correction due to this temperature gradient of the standard atmosphere.
Calculating temperature gradients from reanalysis data reveals that corrections with the standard atmosphere gradient might be problematic, at least for the January temperature. These gradients often are smaller than 0.0065 K/m, especially in Scandinavia, probably due to winter inversions. Therefore it is advisable to work with a lower temperature correction for avoiding positive biases. However, also a correction calculated from reanalysis data might be insufficient, as the vertical resolution is relatively low.
As precipitation is only observed at the surface, no altitude correction has to be applied there.

Discretisation of the model part

In this paragraph the discretisation of the model part should be briefly described for the three dimensional analysis. Here also the method of finite differences is used, like it was done by Gebhardt (2003) did. This method assumes that spatial partial derivation on a regular grid with spacing Δx can be expressed as

$$\frac{\partial f}{\partial x}_{|x=x_i} \approx \frac{f(x_{i+1}) - f(x_{i-1})}{x_{i+1} - x_{i-1}} = \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x}. \quad (3.34)$$

The spatial arrangement of the Arakawa-C grid (Arakawa, 1972) ensures that operators and quantities like temperature anomalies are defined on the same grid. The two equations that have to be discretised should be brought back to mind at this point. These are the advection-diffusion equation

$$\nabla(\bar{\vec{v}}T^{'}) - \alpha\nabla^2 T^{'} = Q_{sol}^{'} + \epsilon,$$

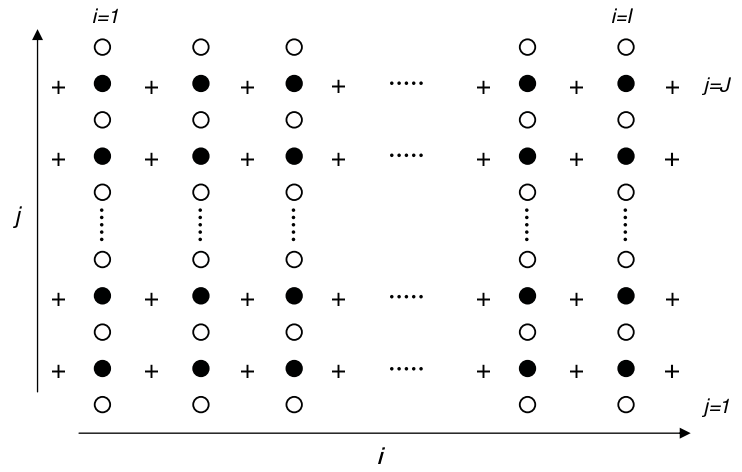

Figure 3.2: Illustration of the discretisation on an Arakawa-C grid.

and the omega equation, already discretised in the vertical:

$$\nabla^2\omega - \frac{2f_0^2}{\Delta p^2\sigma}\omega = -\frac{f_0^2}{\Delta p^2\sigma}g\rho_0\vec{v_h}\nabla h + Q^{'}.$$

Both equations share the Laplace operator and the advection-diffusion equation also contains the flux divergence. Thus, the discretisation of two operators is necessary.

For the theory of discretisation on the Arakawa-C grid, a scalar quantity q is regarded. This quantity is defined at discrete points at the centre of each grid box which is marked by '$\bullet$' in Fig. (3.2). The west-east and south-nord position are given by the indices $i = 1, ..., I$ and $j = 1, ..., J$, so that the position vector $\vec{r}_k$ is replaced by $\vec{r}_{i,j}$. Thus, the scalar quantity, discretised on the centre of each grid box, is marked $q_{i,j}$, respectively.
Two other grids, one shifted half a grid size to east-west direction and marked by '$\circ$', the other shifted to south-north direction and marked by '+', are present. On the '+'-grid, the zonal wind components $u_{i,j}$ are defined, as well as the meridional wind components $v_{i,j}$ on the '$\circ$'-grid. A final convention completes the discretisation rules: $u_{i,j}$ refers to the grid point west of $\vec{r}_{i,j}$ and $v_{i,j}$ to the grid point south of it.
It is assumed within the method of finite differences that the change of q between two grid points is linear in space. With a constant zonal ($\Delta\lambda$) and meridional ($\Delta\varphi$) grid spacing, the discretised derivatives are written as

$$\frac{\partial q}{\partial\lambda_{i,j,+}} = \frac{q_{i,j} - q_{i-1,j}}{\Delta\lambda},$$

$$\frac{\partial q}{\partial \varphi_{i,j,\circ}} = \frac{q_{i,j} - q_{i,j-1}}{\Delta\varphi}.$$

Flux divergence

The flux divergence of q in spherical coordinates is given by

$$\nabla(\vec{v}q) = \frac{1}{rcos\varphi}\frac{\partial}{\partial\lambda}(uq) + \frac{1}{rcos\varphi}\frac{\partial}{\partial\varphi}(cos\varphi vq). \tag{3.35}$$

Its discretisation is executed in two steps. In the first step the products uq and $cos\varphi vq$ are calculated on their respective grids:

$$(u\bar{q}^{\lambda})_{i,j} := u_{i,j} \cdot \frac{q_{i,j} + q_{i-1,j}}{2} \tag{3.36}$$

$$(v\bar{q}^{\varphi})_{i,j} := cos\varphi_{j_\circ} v_{i,j} \cdot \frac{q_{i,j} + q_{i,j-1}}{2}. \tag{3.37}$$

Here the notation $\bar{q}$ means an average over two grid boxes in the respective direction, while $cos\varphi_{j_\circ}$ defines the cosine of the respective latitude on the '$\circ$'-grid.
In the second step the flux divergence is defined on the '$\bullet$'-grid as

$$\nabla(\vec{v}q)_{i,j} = \frac{1}{rcos\varphi_j}\left(\frac{1}{\Delta\lambda}[(u\bar{q}^{\lambda})_{i+1,j} - (u\bar{q}^{\lambda})_{i,j}] + \frac{1}{\Delta\varphi}[(v\bar{q}^{\varphi})_{i,j+1} - (v\bar{q}^{\varphi})_{i,j}]\right). \tag{3.38}$$

As the discretised flux divergence is a linear combination of $q_{i,j}$ and its neighbours in both directions, it can be expressed as a linear transformation

$$\vec{F}_q = \mathbf{F}\vec{q}. \tag{3.39}$$

Here the vector $\vec{F}_q$ contains the values of the flux divergence for each grid point and $\mathbf{F}$ is the flux divergence operator for the complete field.

Laplace Operator

The continuous Laplacian in spherical coordinates is given by

$$\nabla^2 q = \frac{1}{r^2 cos\varphi}\left[\frac{1}{cos\varphi}\frac{\partial^2 q}{\partial\lambda^2} + \frac{\partial}{\partial\varphi}\left(cos\varphi\frac{\partial q}{\partial\varphi}\right)\right]. \tag{3.40}$$

As described above, the zonal and meridional derivatives of q are discretised on the '+'- and '$\circ$'-grid, respectively. The remaining expressions $\frac{\partial q}{\partial\lambda}$ and $\frac{\partial q}{\partial\varphi}$

are then differentiated on the '•'-grid, analogously to the discretisation of the flux divergence. This finally leads to

$$(\nabla^2 q)_{i,j} = \frac{1}{r^2 cos\varphi}\left(\frac{1}{cos\varphi_j \Delta\lambda}\left[\left(\frac{\partial q}{\partial\lambda}\right)_{i+1,j} - \left(\frac{\partial q}{\partial\lambda}\right)_{i,j}\right] + \frac{1}{\Delta\varphi}\left[cos\varphi_{j+1_\circ}\left(\frac{\partial q}{\partial\varphi}\right)_{i,j+1} - cos\varphi_{j_\circ}\left(\frac{\partial q}{\partial\varphi}\right)_{i,j}\right]\right). \tag{3.41}$$

Like the flux divergence, the Laplacian is defined on the same grid as q is. So the vector of Laplacians $\vec{L}_q$ is also a linear transformation of $\vec{q}$:

$$\vec{L}_q = \mathbf{L}\vec{q}. \tag{3.42}$$

Application to the given problem

In this paragraph the application of the discretised analysis to the given three-dimensional application is described. The scalar quantity q is replaced by the temperature anomalies T'_{Jan} and T'_{Jul} and the annual precipitation anomaly P'_{ann}. These quantities are written as vectors $\vec{T}'_{Jan}$, $\vec{T}'_{Jul}$, $\vec{P}'_{Ann}$ containing the values at each grid point and are summarised in the anomaly vector $\vec{X}' = (\vec{T}'_{Jan}; \vec{T}'_{Jul}; \vec{P}'_{Ann})$.
With respect to the given problem the linear transformations described above are specified in more detail. For the flux divergence this gives

$$\vec{F}_{\vec{X}'} = \begin{pmatrix} \mathbf{F}_{Jan} & 0 & 0 \\ 0 & \mathbf{F}_{Jul} & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{X}'. \tag{3.43}$$

The flux divergence operator distinguishes between January and July because of the different wind field. All values of the third line remain zero, as the flux divergence is not applied to the fields of precipitation anomalies. In the omega equation a constant factor is added to the Laplacian. Hence we define $\mathbf{L}^\omega = L + \frac{2f_0^2}{\Delta p^2 \sigma}\mathbf{I}$, where $\mathbf{I}$ is the unity matrix. This leads to

$$\vec{L}_{\vec{X}'} = \begin{pmatrix} \mathbf{L} & 0 & 0 \\ 0 & \mathbf{L} & 0 \\ 0 & 0 & \mathbf{L}^\omega \end{pmatrix} \vec{X}'. \tag{3.44}$$

Concerning the model costs, the discretisation again replaces the integral in the cost function by a sum over all grid points:

$$J_{mod} = \sum_{i=1}^{N} \vec{\epsilon}_i^T \Sigma_{M,i}^{-1} \vec{\epsilon}_n \Delta sin\varphi_n. \tag{3.45}$$

The model error covariance matrix $\Sigma_{M,i}$ is only valid for grid point i. It can be replaced by a complete model error covariance matrix, containing all modell errors:

$$\Sigma_M = \begin{pmatrix} \Sigma_{M,\vec{X}_{1,1}} & \Sigma_{M,\vec{X}_{1,2}} & \Sigma_{M,\vec{X}_{1,3}} \\ \Sigma_{M,\vec{X}_{2,1}} & \Sigma_{M,\vec{X}_{2,2}} & \Sigma_{M,\vec{X}_{2,3}} \\ \Sigma_{M,\vec{X}_{3,1}} & \Sigma_{M,\vec{X}_{3,2}} & \Sigma_{M,\vec{X}_{3,3}} \end{pmatrix}. \tag{3.46}$$

This matrix has the dimension $3N \times 3N$ and consists of three $N \times N$ submatrices. So for instance the submatrix $\Sigma_{M,\vec{X}_{1,2}}$ contains the covariances between all grid points of the T_{Jan} and T_{Jul} fields.
With the definition of the complete model operator

$$\mathbf{M} = \begin{pmatrix} \mathbf{F}_{Jan} & 0 & 0 \\ 0 & \mathbf{F}_{Jul} & 0 \\ 0 & 0 & 0 \end{pmatrix} - \alpha \begin{pmatrix} \mathbf{L} & 0 & 0 \\ 0 & \mathbf{L} & 0 \\ 0 & 0 & \mathbf{L}^{\omega} \end{pmatrix} \tag{3.47}$$

and the difference vector between operator and forcing (see below)

$$\vec{\epsilon} = \mathbf{M}\vec{\mathbf{X}}^{'} - \vec{\mathbf{Q}}^{'}, \tag{3.48}$$

the model costs can be expressed as

$$J_{mod}(\vec{\mathbf{X}}^{'}) = \vec{\epsilon}^{T}\mathbf{W}^{T}\Sigma_M^{-1}\mathbf{W}\vec{\epsilon}. \tag{3.49}$$

The $3N \times 3N$ matrix $\mathbf{W}$ contains the latitude dependent weights of the respective grid points.
It should be noted that Gebhardt (2003) did not estimate Σ_M, but set it to the unity matrix. He decided so, because there was no proper method available for estimating the covariance structure of a field that should be reconstructed. Making assumptions about the covariance structure is linked with the danger of influencing the reconstruction result too much. Within this work, this unsolved problem was not addressed further.

Forcing

The forcing vector $\vec{\mathbf{Q}}^{'}$ from Eq. (3.48) still has to be specified. It consists of radiative forcing anomalies for January and July and of an orography forcing term for precipitation:

$$\vec{\mathbf{Q}}^{'} = (Q'_{Jan,1} \cdots Q'_{Jan,N} Q'_{Jul,1} \cdots Q'_{Jul,N} Q_{Oro,1} \cdots Q_{Oro,N}) \tag{3.50}$$

The forcing anomaly for temperature is controlled by the divergence of the anomalous energy flux $\vec{S}'$. In agreement with Gebhardt (2003) it is assumed

that only the vertical component $S^{'}$ contributes to the divergence leading to a temporal change of the temperature anomaly profile $T^{'}(z)$ of

$$\left(\frac{dT^{'}}{dt}\right) = -\frac{g}{c_p}\frac{\partial S^{'}}{\partial p}. \tag{3.51}$$

The necessary assumption of hydrostatic equilibrium is valid on the temporal and spatial scales that are considered here (Liou, 2002). Further, it is assumed that the effect of the anomalous forcing is homogeneously distributed in the vertical. This would surely be a risky assumption for a complex model, but is applicable in the case of the advection-diffusion model, where only near surface values are regarded.
Finally one can approximate the anomalous surface energy flux to be zero leading to an approximated forcing term of

$$Q^{'} = -\frac{g}{c_p p_0}(1-a)S. \tag{3.52}$$

Here the albedo a, which is the fraction of solar radiation that is reflected by the earth, has to be specified. It is set to $a = 0.35$, regarded as a realistic value for Central Europe by Peixoto and Oort (1992).
The forcing for precipitation (right side of Eq. (3.29)) is discretised as above, leading to

$$Q_{Oro} = -\frac{f_0^2}{\Delta p^2 \sigma} g\rho_0 \left(u\frac{h_{i,j} - h_{i-1,j}}{\Delta\lambda} + v\frac{h_{i,j} - h_{i,j-1}}{\Delta\varphi}\right) \tag{3.53}$$

More about the data used for calculating the forcing terms can be found in the data chapter.

Boundary conditions

At the end of the discretisation section it should be mentioned, which boundary conditions are applied to the analysis. First the boundary is defined between the two outermost row of grid points. The last row inside the boundary contains the outermost analysed field values. According to the boundary conditions, the values are extrapolated to the respective grid point, lying out of the boundary.
Generally there are two most common boundary conditions. These are:

- The 'Dirichlet' boundary condition, meaning the values to be zero on the boundary. Thus, the value left of the boundary is the inverse of the value right of it.

- The 'von Neumann' boundary condition, meaning that the partial derivative of the value is zero on the boundary. This leads to the same value on the right and left side of the boundary.

Both boundary conditions are used in the analysis.

3.3 Reducing the dimension

The minimisation of the cost function, presented up to now, is a high dimensional procedure. The number of grid boxes in the different reconstruction areas varies from about 500 to over 800 and has to be multiplied by the number of variables. Vegetational data, however, are available only at about 40 grid boxes. This requires the reduction of the dimension of the problem by at least one magnitude.
A method, successfully applied for this task by Gebhardt (2003), is the singular value decomposition (SVD), described by Golub and Kahan (1965). The theory states that a $M \times P$ matrix $\mathbf{A}$ can be decomposed into three matrices:

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}^T. \tag{3.54}$$

Here $\mathbf{U}$ is a column-orthonormal $M \times P$ matrix, meaning $\mathbf{U}^T\mathbf{U} = \mathbf{E}$ ($\mathbf{E}$ is the unity matrix). Its columns are called left singular vectors $\vec{u}_p$. The $P \times P$ matrix $\mathbf{V}$ contains the right singular vectors $\vec{v}_p$ and is column- as well as row-orthonormal, meaning $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{E}$. The diagonal matrix Λ finally contains the singular values $\lambda_p \geq 0$.
As the theory of SVD is well documented in the literature and its application to the model operator M is comprehensively described in Gebhardt (2003), it is only briefly described here.
The decomposition of the model operator starts with the definition of the matrix $\mathbf{S}$ with $\mathbf{S} = \Sigma_M^{-1}$. Now the matrix product $\mathbf{S} \cdot \mathbf{W} \cdot \mathbf{M}$ is decomposed separately for the 'Dirichlet' and 'von Neumann' boundary conditions

$$\mathbf{S} \cdot \mathbf{W} \cdot \mathbf{M}_D = \mathbf{U}_D\Lambda_D\mathbf{V}_D^T \tag{3.55}$$

$$\mathbf{S} \cdot \mathbf{W} \cdot \mathbf{M}_N = \mathbf{U}_N\Lambda_N\mathbf{V}_N^T \tag{3.56}$$

where the index D denotes the 'Dirichlet' and N the 'von Neumann' boundary conditions. All matrices share the dimension of $3N \times 3N$.
As Σ_M is set to the unity matrix, it is implicitly assumed that January, July and annual precipitation model errors are uncorrelated. Thus, the three

variables are separable in the SVD, leading to matrices of the shape:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{T_{Jan}} & 0 & 0 \\ 0 & \mathbf{U}_{T_{Jul}} & 0 \\ 0 & 0 & \mathbf{U}_{P_{Ann}} \end{pmatrix}. \tag{3.57}$$

Analogously the matrices Λ and $\mathbf{V}$ are separated. In Fig. (3.3) the first six singular vectors (modes) $\vec{v}_p$ for the January temperature under both boundary conditions are shown. The singular vectors are sorted according to their singular values λ_p. This is done so, that the first mode belongs to the smallest singular value what is tantamount to the largest spatial scale.
Now the cost function can be expressed in terms of projection of the climate field onto the P singular vectors with the smallest singular values. This means the reduction of the problem to a value $P < 3N$, with P_D vectors from the 'Dirichlet' set and P_N vector from the 'v. Neumann' set. Then the climatological fields are approximated as

$$\vec{\mathbf{X}}^{'} \approx \tilde{\vec{\mathbf{X}}}^{'} = \sum_{p=1}^{P_D} c_{D,p}\vec{v}_{D,p} + \sum_{q=1}^{P_N} c_{N,q}\vec{v}_{N,q}. \tag{3.58}$$

The contribution of each singular vector to $\tilde{\vec{\mathbf{X}}}^{'}$ is quantified by the coefficients $c_{D,p}$ and $c_{N,q}$. Some more details about things to take into account when mixing singular vectors from different boundary conditions can be found in Gebhardt (2003). Here it only should be mentioned, that linear independence between the different singular vectors is assured.
When all coefficients are summarised into one vector $\vec{c} = (\vec{c}_D^T \vec{c}_N^T)^T$ and all singular vectors into one matrix $\tilde{\mathbf{V}} = (\tilde{\mathbf{V}}_D \tilde{\mathbf{V}}_N)$, the coefficient vector is given by

$$\vec{c} = (\tilde{\mathbf{V}}^T \tilde{\mathbf{V}})^{-1} \tilde{\mathbf{V}}^T \vec{\mathbf{X}}^{'}. \tag{3.59}$$

Note that $\tilde{\mathbf{V}}^T \tilde{\mathbf{V}} \neq \mathbf{E}$ in general due to the mixing of modes with different boundary conditions. Now the final form of the cost function with full replacement of the climate vector by the coefficient vector is given by

$$\begin{aligned} J(\vec{c}) = &- \sum_{k=1}^{K} ln\left[\prod_{i=1}^{n_k} \frac{f_{\vec{X}} \mid Y_i(\tilde{\mathbf{V}}\vec{c} + \vec{x}_0 \mid y_i)}{m_{\vec{X}}(\vec{x})} \right] \cdot \Delta sin\varphi_k \\ &+ \gamma_M \cdot (\vec{c}^T \tilde{\Lambda} \tilde{\mathbf{U}}^T \tilde{\mathbf{U}} \tilde{\Lambda} \vec{c} - 2\vec{c}^T \tilde{\Lambda} \tilde{\mathbf{U}}^T \Sigma_M^{\frac{1}{2}} \mathbf{W} \vec{\mathbf{Q}}^{'} - \vec{\mathbf{Q}}^{'T} \mathbf{W}^T \Sigma_M^{-1} \mathbf{W} \vec{\mathbf{Q}}^{'}). \end{aligned} \tag{3.60}$$

Now the cost function can be minimised with respect to the coefficients $\vec{c}$. As the term $\vec{\mathbf{Q}}^{'T} \mathbf{W}^T \Sigma_M^{-1} \mathbf{W} \vec{\mathbf{Q}}^{'}$ is a constant, it can be skipped in the minimisation

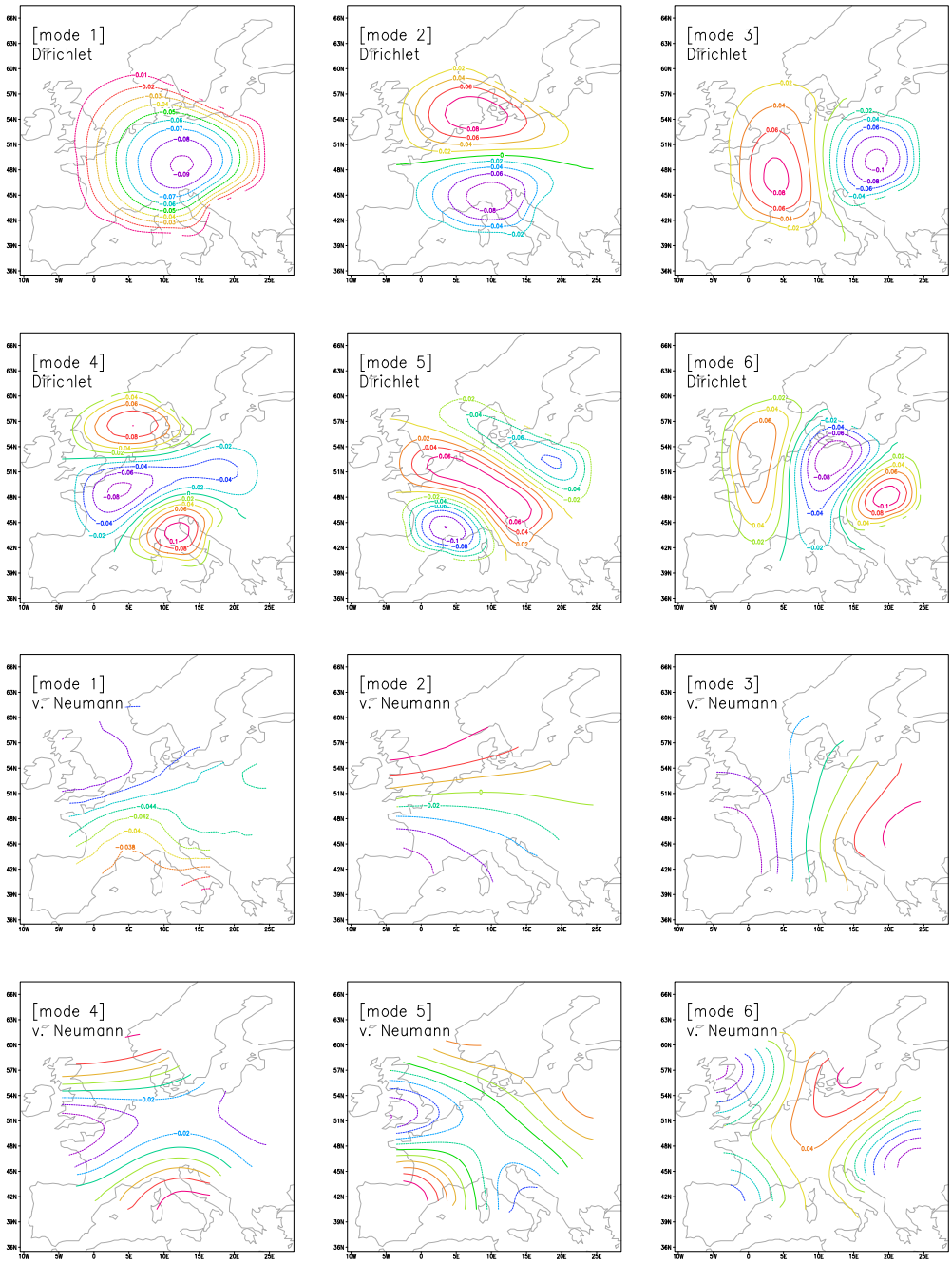

Figure 3.3: First six modes of the advection-diffusion operator for T_{Jan} under different boundary conditions. Displayed analysis area is 13 ky B.P.

procedure. The minimisation is carried out with a modified Powell's method (Press et al., 1992). It belongs to a class of algorithms approaching the minimum of a P-dimensional function with a set of conjugate directions. Gebhardt (2003) found this method not performing worse than alternative approaches.
The contribution of both parts to the costs is controlled by the parameter γ_M, as mentioned in Section (3.1.1).

3.4 The analysis error

A very important aspect of every analysis is not only to present the analysis result but also give information about the analysis error. Without any information about how reliable the analysis result is, it would be nearly worthless. Conveniently the variational analysis provides a way to derive the analysis error from the minimum of the cost function. The analysis error $\vec{\epsilon}_R$ is defined as

$$\vec{\epsilon}_R = \vec{c}_R - \vec{c}_{R,t}, \tag{3.61}$$

where $\vec{c}_R$ is the minimising coefficient vector and $\vec{c}_{R,t}$ is the unknown vector, regarded as true. This means its transformation to the climate space yields the true climate values. When the analysis error is assumed to be unbiased, the analysis error covariance matrix is given by

$$\Sigma_R = E[(\vec{c}_R - \vec{c}_{R,t})(\vec{c}_R - \vec{c}_{R,t})^T]. \tag{3.62}$$

It was shown by Gauthier (1992) and Rabier and Courtier (1992), that Σ_R is related to Hessian matrix $\mathbf{H}$ of J_D at its minimum by

$$\frac{1}{2}\mathbf{H}(\vec{c}_R) = \Sigma_R^{-1}. \tag{3.63}$$

It was pointed out by Courtier et al. (1994), that the relation between these two matrices also holds for non Gaussian error statistics. However, Gebhardt (2003) showed by assuming Gaussian error statistics, that the analysis error covariance matrix is a combination of vegetational and model error:

$$\frac{1}{2}\mathbf{H} = \Sigma_R^{-1} = \Sigma_{veg}^{-1} + \gamma_M \Sigma_{mod}^{-1}. \tag{3.64}$$

He pointed out, that this clarifies the philosophy of reducing the error variances by combination of vegetational and model information.
Despite the dimension reduction, the matrix Σ_R can still be very large and additionally provides error information only in the space of singular modes.

A possibility to extract the essential information is given by the calculation of the eigenvectors $\vec{e}_{R,p}$ of Σ_R with its eigenvalues $\lambda_{R,p}(p = 1, ..., P)$. Analysis of the eigenvectors implies the following aspects:

- Eigenvectors are uncorrelated. The contribution of each eigenvector to the analysis error can be quantified by their eigenvalues (von Storch and Zwiers, 1999).
- The relative contribution of each eigenvector is given by the explained variance, calculated as the ratio of the respective eigenvalue to the sum over all eigenvalues.
- Projection of the analysis result on the eigenvector with the largest eigenvalue gives the impression of the most uncertain pattern. Thus, the analysis of the eigenvectors provides a way to evaluate the quality of the reconstruction.

The projection of the coefficient vector on the p-th eigenvector is given by

$$\gamma_p = \vec{c}_R^T \cdot \vec{e}_{R,p}. \tag{3.65}$$

As the result is projected into the P-dimensional space of singular modes, its interpretation is difficult. More interesting is the transformation of the eigenvectors to the climate space by the singular modes:

$$\vec{e}_{T,p} = \mathbf{V}\vec{e}_{R,p}. \tag{3.66}$$

These eigenvectors in climate space give an impression of patterns with high variability. Thus regions can be identified, where the reconstruction result is more certain or uncertain. Also the analysis error covariance matrix can be transformed into the temperature space by

$$\Sigma_T = \mathbf{V}\Sigma_R\mathbf{V}^T. \tag{3.67}$$

The matrix in temperature space contains information about the variance of every grid box in the analysis area as well as about the covariances between the respective grid boxes. Analysing these covariances or its associated correlation provides a possibility for checking how realistic the correlation structure of the analysed field is. Further, the statistics of error variances show the quality of the reconstruction. It should be noted here, that Σ_T is at most of rank P.

3.4.1 Resampling from the analysis error covariance matrix

It is a challenging task to display the uncertainty of a reconstructed climatological field properly. Most people consider the uncertainty of a certain value to be some plus/minus range around this value. When regarding a two dimensional field this becomes more complicated. Can the uncertainty be represented by the same error interval everywhere in the field? More likely the uncertainty will be large in one area and small in another one.
These aspects of uncertainty are not the only ones when reconstructing large scale temperature anomaly patterns. The analysis error basically is linked to reconstructed coefficient vector in the space of singular modes. As every mode represents a certain spatial pattern, this automatically means an uncertainty in the spatial patterns itself.
Therefore a smart approach is needed to display the uncertainty in the spatial patterns. The idea that came up for this task within this work is to resample alternative fields from Σ_R. Doing that implies to consider the field that is to be reconstructed as a multivariate normal distributed random variable Z with

$$Z \sim N(\vec{c}_R, \Sigma_R). \tag{3.68}$$

This means, that the reconstructed coefficient vector is regarded as conditional expectation of the variety of possible realisations of the coefficient vector given the paleodata and the constraint. The variability of possible alternative coefficient vectors is given by the analysis error covariance matrix. The multivariate problem can be reduced to an univariate one by calculating the principal components of $\vec{c}_R$, given by

$$\alpha_p = \vec{e}_R^{*T} \vec{c}_R. \tag{3.69}$$

For this purpose, the eigenvectors have been sorted in such a way, that $\vec{e}^*_{R,1}$ is the eigenvector belonging to the largest eigenvalue λ_1. The covariance matrix of the principal components Σ_α, calculated as

$$\Sigma_\alpha = \vec{e}_R^{*T} \Sigma_R \vec{e}^*_R, \tag{3.70}$$

is a diagonal matrix containing the eigenvalues λ_p. This is related to the fact that the eigenvectors of Σ_R are also the empirical orthogonal functions (EOFs) of $\vec{c}_R$ (von Storch and Zwiers, 1999). The orthogonality of the eigenvectors leads to uncorrelated principal components.
In the EOF-space there are now given P univariate normal distributed random variables Z_p with

$$Z_p \sim N(\alpha_p, \lambda_p). \tag{3.71}$$

For the resampling M sets of principal components $\vec{\alpha}_m = (\alpha_{1,m}, ..., \alpha_{q,m})$ are drawn randomly from the respective normal distribution. The backtransformation into the space of singular modes and afterwards into climate space is straightforward given by

$$\begin{aligned} \vec{c}_m &= \vec{e}_R \vec{\alpha}_m \\ \vec{X}'_m &= \mathbf{V} \vec{c}_m. \end{aligned} \tag{3.72}$$

Generally, each of the M resampled fields has to be regarded as a possible realisation of the reconstructed climatological field. However, not all fields have to be regarded as equally likely. Their likeliness can be quantified by their distance to the analysed field, which is the estimated expectation of $\vec{X}'_R$. An adequate measure for these distances is provided by the mahalanobis distance (Mahalanobis, 1936). It is calculated in the space of singular modes for the respective resampled coefficient vector $\vec{c}_m$:

$$M_D(\vec{c}_m) = \sqrt{(\vec{c}_m - \vec{c}_R)^T \Sigma_R^{-1} (\vec{c}_m - \vec{c}_R)}. \tag{3.73}$$

Unlike the euclidean distance, the mahalanobis distance incorporates the covariance structure of a multivariate random variable. In the case of a diagonal matrix, the mahalanobis distance is equal to the normalised euclidean distance.

As mentioned, the resampled fields all are possible realisations of the climate of the past. On contrary, the field minimising the cost function, is not really a possible realisation but the expectation of the past climate. This means, that all possible realisations on average lead to this field. Thus the field, minimising the cost function, can look significantly different than most of the realisations and can be expected to be much smoother. This fact emphasises that the variety of resampled fields should be regarded as the real reconstruction result.

Chapter 4

Data

Several datasets have been used within this work. These datasets document the long way from sediment cores to quantitative climate field reconstructions. The datasets can be distinguished between data necessary for estimating the transfer functions and data used for the reconstructions.

4.1 Data for estimating transfer functions

As transfer functions provide the connection between biological proxy data and climatological variables, a proxy dataset and a climate dataset is needed for estimating them. The botanical proxy, used for the reconstructions within this work, is the presence of vegetational taxa. Therefore the proxy dataset has to contain information about the modern distribution of vegetation.

4.1.1 Modern vegetation data

The database, available to Gebhardt (2003), was further expanded during the last years by the paleobotanists in our collaboration. Data sources were vegetation maps by Meusel et al. (1964, 1974) and Meusel and Jäger (1992), as well as Hultén (1964, 1971). Additional data was taken from the Atlas Florae Europaea (AFE): Jalas and Suominen (1973, 1976, 1989). All these data sources, available as printed maps, had to be digitised and still have to be partially digitised. A flexible digitisation software was developed for this purpose by Schölzel et al. (2002). In 2003 the most important 126 European taxa had been digitised and were used for the work of Gebhardt (2003). At the moment the database contains about 230 different taxa, providing a very good basis for climate reconstructions.

Two examples from the digitised taxa dataset can be seen in Fig. (4.1). In

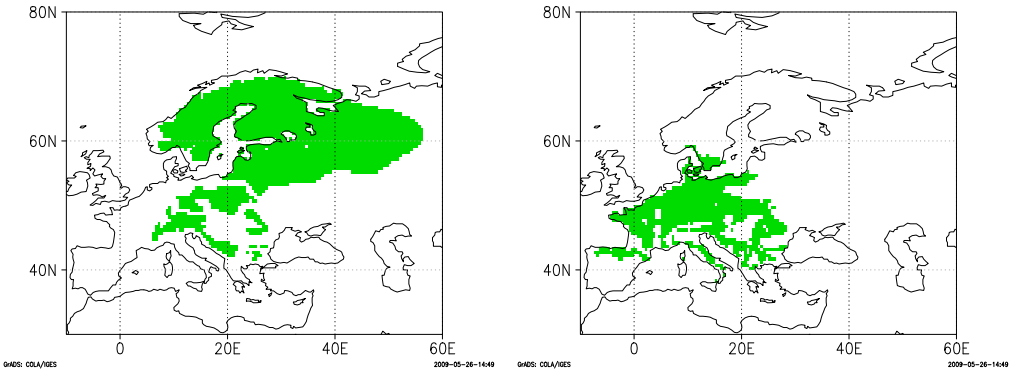

Figure 4.1: Occurrence of *Picea abies* (left) and *Fagus sylvatica* (right) after Meusel et al. (1964).

the datasets all grid boxes are marked, where the respective taxon occurs. The grid spacing for digitisation was set to 0.5°, because this is also the resolution of the climatological data. *Picea abies* occurs very far to the North and therefore is an indicator for a cold climate. In contrary *Fagus sylvatica* occurs in central and western parts of Europe, making it an indicator for mild winters.
The aspects of digitisation of vegetation maps were discussed in Gebhardt (2003) and Kühl (2002). Some important features, however, should also be mentioned here:

- Taxa occurring in geographical disjoint areas can be problematic. The different areas could be related to different subspecies of the same plant and thus represent different climatological conditions.
- Taxa occurring in exclaves could be related to microclimatic conditions that are not representative for the whole grid box. Therefore small exclaves were omitted in the digitisation procedure.
- Anthropogenic influence on the distribution of plants can be a problem. The authors of the vegetation maps tried to leave out these kind of influences.

Especially the first point can be addressed easily by omitting disjoint areas far away from the European area. The European area is regarded as most important as it is the target area for our climate reconstructions. It can be assumed that taxa, being present in Europe today and in the past, represent the same subspecies. Also exclaves can be identified easily and then be excluded.

It is more difficult to avoid human influences in the datasets. For some plants, like *Picea abies*, it is well known where it was planted by humans and where it can occur naturally. The other way round, when trees were deforested, it is more difficult to estimate the area where these trees would occur naturally. This point is an inherent error source when estimating transfer functions on vegetation data. It was minimised as good as possible by the biologists, but its remaining effect is hard to quantify.

4.1.2 Climatological data

The climatological data set from the Climate Research Unit (CRU,New et al. (1999)) was already used in the earlier mentioned previous studies for estimating transfer functions. There are nine variables included in the dataset. These are temperature, precipitation, daily temperature range, cloud cover, frost days, rainy days, sunshine percentage, vapour pressure, wind speed and solar radiation. Generally, transfer functions conditional on all these variables could be estimated. However, only temperature and precipitation are assumed to be limiting factors for the occurrence of plants. The data used from the CRU climatology are presented in Fig. (4.2). As the temperatures of the coldest and warmest month are agreed to be the most important climatological factors for the occurrence of plants, January and July mean temperatures were chosen. Whether precipitation also plays a considerable role, is a question that has to be answered within this work.
All three datasets show different characteristics. First, the January mean temperatures in Eurasia show a greater range with values from -50°C in northern Siberia to 10°C in southern Portugal. The July mean temperatures only range from around 5°C to around 30°C. Furthermore, there are differences in spatial patterns. While the July temperatures are mainly controlled by solar insolation (apart from orography effects) and its isotherms are parallel to the latitudes, the January temperatures are more influenced by the ocean. This becomes clear when the mean temperatures of western Norway and Siberia are compared. While in January western Norway is about 40 K warmer than Siberia, these two regions have approximately the same July mean temperatures. The regions close to the Atlantic ocean benefit from the warm water, transported northwards by the gulf stream. Therefore the winters here are milder than in other regions on the same latitude.
The field of annual precipitation looks completely different to the fields of mean temperatures. A huge area from the Near East to the Pacific coast of Eurasia has very low values of annual precipitation below 500 mm. However, in Europe, the area of interest in this thesis, the values are always above 500 mm. High values above 1000 mm or strong spatial gradients are only found

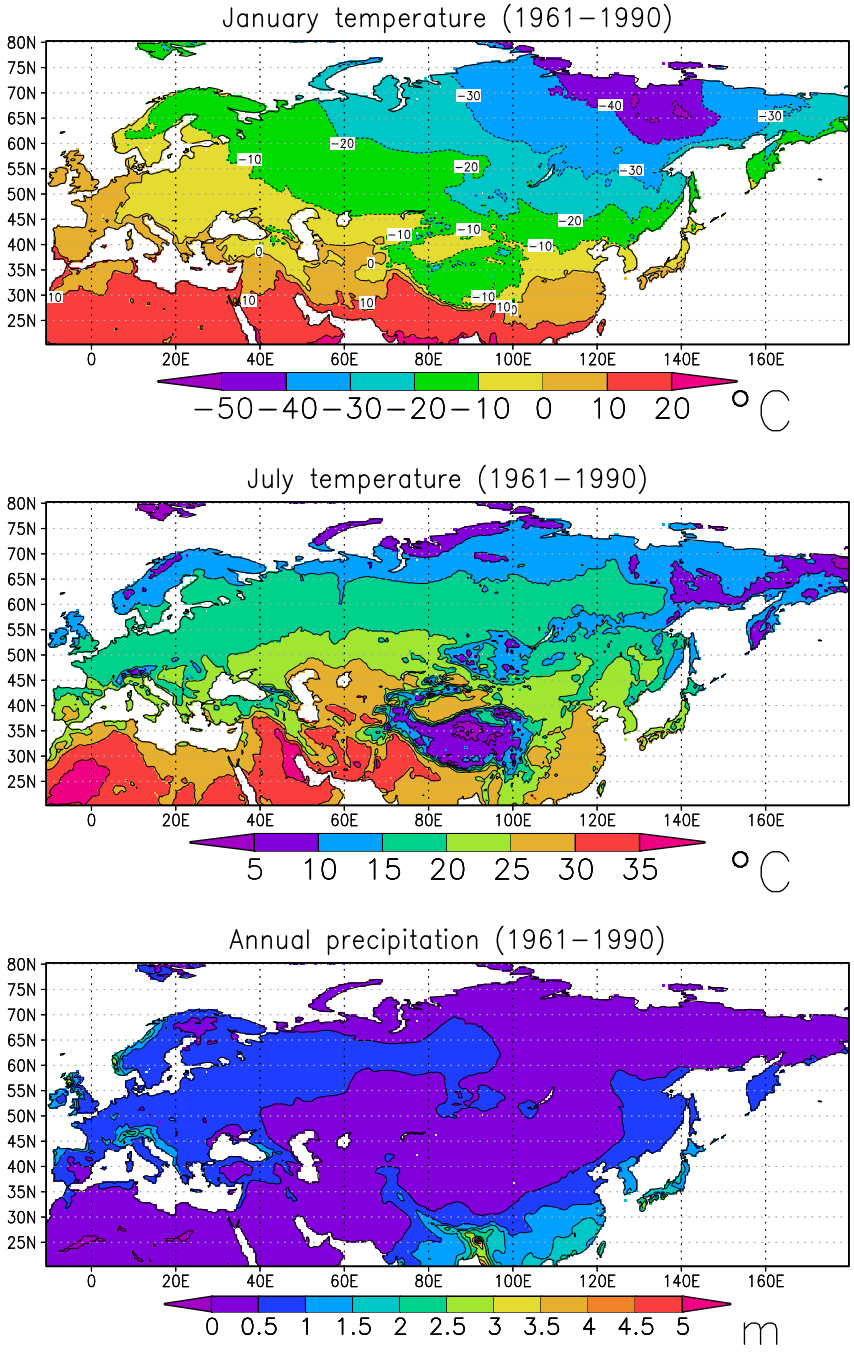

Figure 4.2: CRU 1961-1990 climatology, January temperature (top), July temperature (middle) and annual precipitation (bottom).

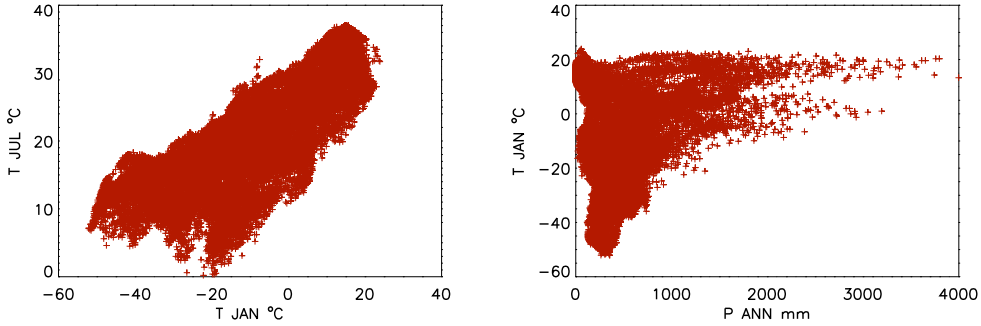

Figure 4.3: Distribution of the CRU climatology values in the climate space. January vs. July temperature (left) and annual precipitation vs. January temperature (right).

in mountainous regions.

Limitations due to the observed climate

The distribution of climate values in the CRU climatology is displayed in Fig. (4.3). It becomes clear that only values from a limited subspace of the climate space are observed in Eurasia. So January values range from about -57°C to 20°C, while July values range from 0°C to about 37°C. Further T_{Jan} and T_{Jul} are highly correlated with a correlation coefficient $R = 0.79$. Thus, a temperature pair with warm July and cold January as well as vice versa is not observed.
Correlation between temperatures and annual precipitation is very low. The values for January and July are $R = 0.18$ and $R = -0.1$, respectively. At least relatively low values of annual precipitation are observed in combination to all observed temperature values.
Higher values of annual precipitation, which are rarely observed, are more problematic. It is plausible that high values of precipitation are observed together with higher temperatures as warm air is able to carry more water. However, the distribution of precipitation vs. January temperature has two branches of high values of annual precipitation, one around 0°C and another close to 20°C. The gap between these branches is surely not explainable by physical laws. Other factors, like the location of mountain ranges or circulation regimes, may play a role here.
The gap between the two branches of annual precipitation illustrates a general problem, when reconstructing climate by transfer functions based on this

dataset. Only climate states, which are contained in the dataset can be theoretically reconstructed. If e.g. a location with modern climate of $T_{Jan} = 0$°C and $P_{ANN} = 2000mm$ would have been slightly warmer in the past, it would automatically be reconstructed dryer.
The limitations due to the observed modern climate lead to the conclusion, that the reconstruction method should not be applied to regions, that are close to the edge of the observed climate today. In these regions variations of the climate out of the modern observed climate would not be reconstructable.

The climate marginal distribution

The non-uniform distribution of the climate observations in the climate space is not only a limiting factor for the possible range of climate reconstructions. It also affects the estimation of the transfer functions itself. This is related to the fact that a taxon occurs more likely in climate that is observed more often. Thus higher pdf-values are generally assigned to climate values that are frequently observed.
The observed climate state vector is included in the Indicator Taxa model (Eq. 2.34) by the climate marginal distribution. Just multiplying the taxon-specific likelihood functions would imply the assumption of an uniformly distributed climate state vector. However, dividing the likelihood functions by the marginal distribution incorporates the problem of dividing by zero or values close to zero at the edges of the distribution. This problem does not arise, when the climate marginal distribution is introduced indirectly by addressing a weight to every observed climate value.
For calculating the weights, the climate space is divided into three-dimensional boxes. The boundaries of these boxes are defined in a way, that every interval contains the same number of values. The weights are calculated as the inverse of values per volume:

$$w_i = \frac{\triangle T_{Jan,i} \cdot \triangle T_{Jul,i} \cdot \triangle P_{Ann,i}}{N_i}. \tag{4.1}$$

These weights are implemented into the M-step of the EM-algorithm (eq. 2.26-2.28) as follows:

$$\hat{\alpha}_k^r = \frac{1}{\sum_{i=1}^{I}(w_i N_i)} \sum_{i=1}^{I} w_i \left(\sum_{n=1}^{N_i} \hat{P}^r(w_k \mid \vec{x}_n) \right), \tag{4.2}$$

$$\hat{\vec{\mu}}_k^r = \frac{\sum_{i=1}^{I} w_i \sum_{n=1}^{N_i} \hat{P}^r(w_k \mid \vec{x}_n)\vec{x}_n}{\sum_{i=1}^{I} w_i \sum_{n=1}^{N_i} \hat{P}^r(w_k \mid \vec{x}_n)}, \tag{4.3}$$

$$\hat{\Sigma}_k^r = \frac{\sum_{i=1}^{I} w_i \sum_{n=1}^{N_i} \hat{P}^r(w_k \mid \vec{x}_n)(\vec{x}_n - \hat{\vec{\mu}}_k^r)(\vec{x}_n - \hat{\vec{\mu}}_k^r)^T}{\sum_{i=1}^{I} w_i \sum_{n=1}^{N_i} \hat{P}^r(w_k \mid \vec{x}_n)}. \tag{4.4}$$

With the EM-algorithm for estimating the transfer function being modified in that way, the climate marginal distribution is already included and can be omitted in the Indicator Taxa model.

4.2 Data for climate reconstructions

4.2.1 Paleobotanical data

During the project, in which framework this work has been carried out, paleobotanical data from several sources had to be collected for providing a satisfactory data basis for climate reconstructions. The data for the time slices, mentioned in the introduction, came from different sources. Especially for the Late Glacial some macrofossil data have been available in the database of the Paleontological Institute. Data from 21 Sites from the European Pollen Database (EPD) were provided by the project member Thomas Giesecke from the University of Liverpool. These datasets were mainly from Northern Europe and cover only the Holocene in most cases. However there exist numerous publications containing pollen or macrofossil data that have not yet been digitised. It was also part of the project to find and digitise data from high quality sites.
A high quality site is defined by several properties. A good temporal resolution of the archive is necessary. In the best case the data comes from a lake that disposes of annual laminated layers (varves). Further, a good dating of the sediment has to be possible. In Central Europe one often benefits from an ash layer (late-glacial Laacher See Tephra(LST)) caused by an eruption of the Laacher See volcano in Western Germany. This eruption was dated to 12880 varve years BP (Brauer et al., 1999) and thus serves as marker for the Alleroed. Wherever the LST is present the identification of the 13000 and 12000 BP time slices is very easy, especially in combination with annual laminated sediments, when the varves can be counted.
An example for a high quality record is given in Fig. (4.4), showing a pollen diagram of the Meerfelder Maar (Litt et al., 2001). In the Lithology on the left side the LST is represented by a black line right above 13000 BP. On the right side the palynostratigraphy of the record is represented by biozones 1 to 4. These zones are determined by certain vegetation signals. So the beginning of the Alleroed for instance is defined by an increase of *Pinus* and a decrease of *Artimisia* pollen ratios. The whole Alleroed is characterised

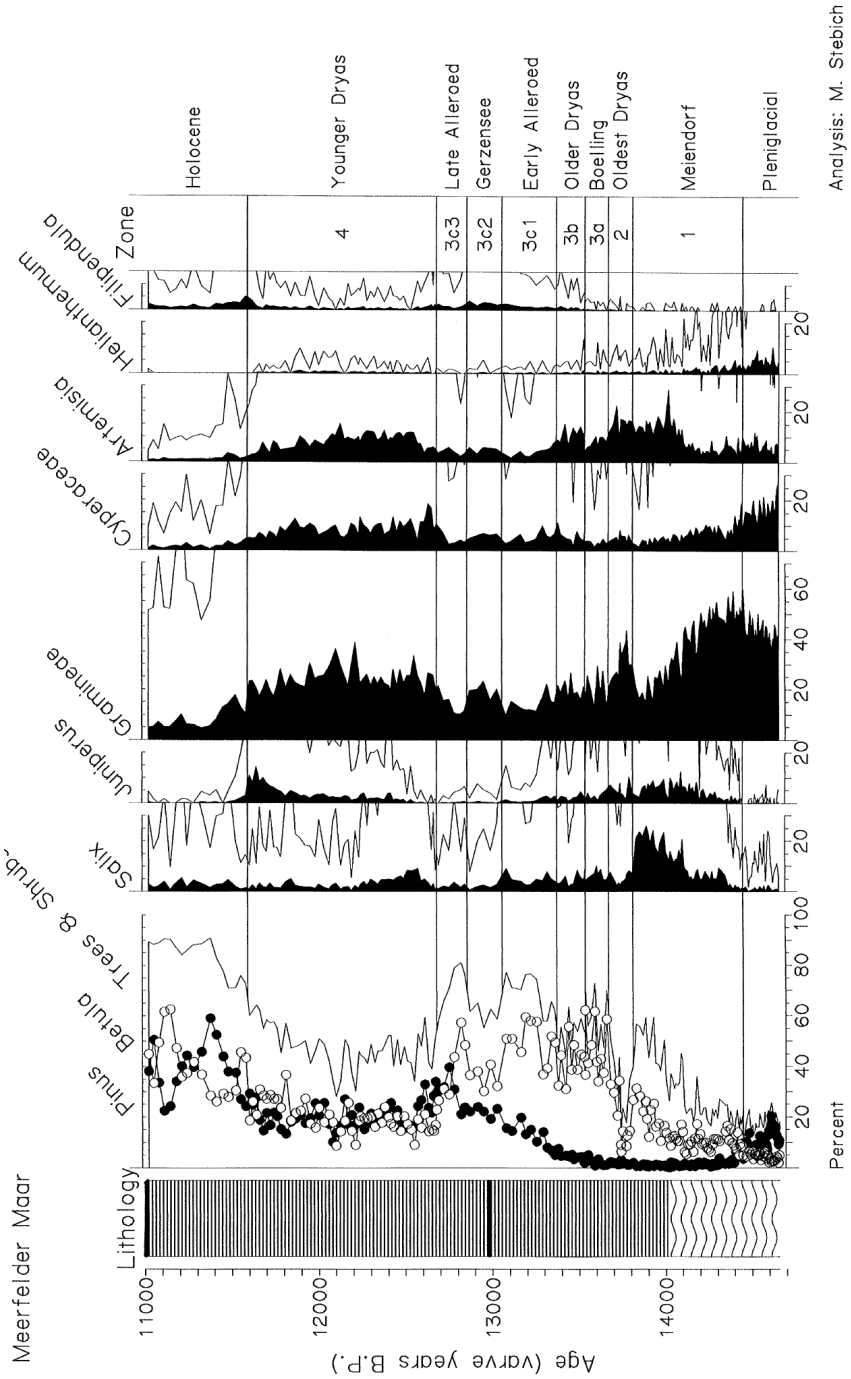

Figure 4.4: Pollen diagram of Meerfelder Maar (Eifel, Germany) for the Late Glacial (Litt et al., 2001).

by a high trees and shrubs ratio. Afterwards the Younger Dryas begins with a rise in non-arboreal pollen.

When LST is absent

In a lot of datasets that were used in this thesis, the LST was not present in the sediment. Then the identification of the Alleroed and the Younger Dryas in most cases is possible with the palynostratigraphy, as the transition between them normally is clearly visible. Another possibility would be using age-depth models based on radiocarbon ^{14}C-dates. However ^{14}C-estimates underlie a relatively large uncertainty which is often underestimated, shown by Telford et al. (2004b,a). They stated that less than 15 dates will result in age-depth model statistics being nothing more than a guess.
Within our project, the mixed effect regression procedure for age-depth models, proposed by Heegard et al. (2005), has been tested. The obtained results partly differ clearly from the identification of the time slices of 13000 and 12000 a BP by the stratigraphy. These differences of a few hundreds of years can have a strong impact when dealing with time slices representing a range of 700 to 1000 years in an era with strong fluctuations in the vegetation records. Thus, for the Late Glacial only sites were used where a clear stratigraphical identification of Alleroed and Younger Dryas was possible.
The time slices of 6000 and 8000 BP are not that clearly visible in the stratigraphy. However in the Holocene the taxa composition, at least in terms of absence and presence, is more stable. Here the age-depth models based on radiocarbon dates are precise enough for being used.
It should be mentioned, that in the literature uncalibrated and calibrated radiocarbon ages can be found. In this thesis generally calibrated radiocarbon dates have been used.

Summary of site selection criteria

The criteria applied for site selection can be summarised as follows: Generally, only sites providing a vegetation record with good resolution were chosen. In these records the Late Glacial time slices must have been identifiable by the LST and/or the palynostratigraphy. For Holocene records there must have been a sufficient number of ^{14}C-dates available for deriving an age-depth model with passable errors.

The question of presence

At first glance the question of whether a taxon is present in the sediment or not seems to be a trivial one. However this is not the case. Only when

dealing with macrofossils it can be stated that a taxon really was present when just one single macrofossil of it has been found in the sediment sample. On contrary pollen are problematic when they are found only with low abundances. Pollen often are transported over long distances, sometimes several hundred kilometers. How far they are transported depends on the pollen itself, the weather conditions and several environmental factors like the landscape. How many pollen are found in the sediment finally depends strongly on the pollen production rate of the respective taxon.
At the moment there exists no objective criterion for how many pollen have to be counted until a taxon can be regarded as present. Thus each debatable taxon was regarded individually by the biologists from the collaboration, who thereafter made a decision whether the taxon was present or not.
A tabulation of all used sites for all time slices with the taxa regarded as present can be found in the appendix.

The question of independence

The statistical formulation of the reconstruction method requires statistical independence of the taxa used for reconstruction at a certain location. This means that the occurrence of a taxon at a certain site is independent from the occurrence of all other taxa at this site. As different taxa are always in competition to each other, it is hard to state if occurrence of taxa really can be regarded as independent.
From the statistical point of view at least two or more taxa with similar distributions should not be considered at one time. This is due to the fact that when two similar pdfs are multiplied, only the variance is reduced, although there is no new information included. As measure for the similarity of two pdfs serves the mahalanobis distance (MD):

$$M_D(\vec{\mu}_1) = \sqrt{(\vec{\mu}_1 - \vec{\mu}_2)^T \Sigma_1 (\vec{\mu}_1 - \vec{\mu}_2)}. \tag{4.5}$$

A taxon was excluded from the analysis by Gebhardt (2003), when the MD was smaller than 0.15. However this preselection rule still left some taxa together with very similar distributions. Thus within this work the minimum MD was set to 0.2.
It should be noted, that the MD cannot be calculated for mixture models. Therefore normal distributions, were used to calculate them. Finding a measure for the similarity of two mixture models for future work would surely enhance the convenience of the preselection. One possibility would be using a measure based on the Kullback-Leibler divergence (relative entropy).

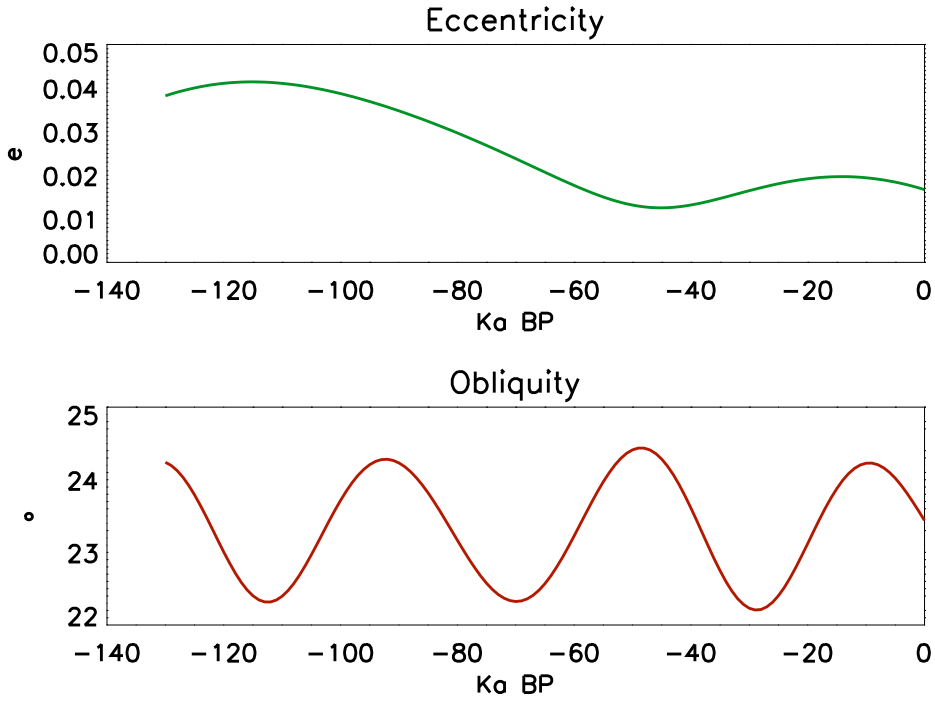

Figure 4.5: Variation of eccentricity and obliquity for the last 125 ka after Berger (1978).

4.2.2 Solar insolation

Solar insolation data is necessary for calculation of the solar insolation S in Eq. (3.52). Even if the solar constant was a real constant, the solar insolation patterns would change over time with the earth's orbital parameters. The two most important orbital parameters, eccentricity and obliquity, are displayed in Fig. (4.5) for the 125 ka covering the time from the Eemian interglacial till today. The obliquity is the angle between the equatorial and the orbital plane of the earth. When the angle between these planes is larger, the point of northern solstice lies further nordwards, meaning stronger differences in summer and winter insolation. The obliquity varies sinusoidal with a period of 40-50 ka. It is noticeable that is reaches a maximum at the beginning of the Eemian (∼125 ka BP) as well as at the beginning of the Holocene. On the other hand it is obvious that not every maximum of the obliquity is followed by an interglacial.
The eccentricity is a measure for how elliptic the orbit of the earth is. With an eccentricity of 0 it would be a perfect circle. This means, the higher

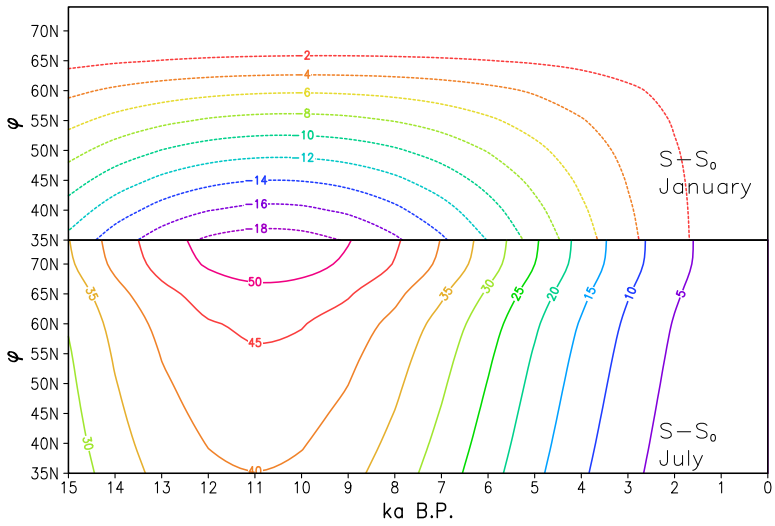

Figure 4.6: Solar insolation anomalies for January and July ($\frac{W}{m^2}$), zonal mean values between 35N and 74N for the last 15000 years.

the eccentricity, the more important is whether the perihelion coincides with northern summer or winter. The eccentricity was relatively high during the Eemian but at a low stable level during the last 60 ka.
The solar insolation anomalies $S-S_0$, where S_0 is the solar insolation of today, can be seen in Fig. (4.6), calculated after Berger (1978). The maximum of summer insolation in northern high latitudes is found at the beginning of the Holocene, around 11000 a BP. This is an important factor for summer snow melting and thus enhancing ice-albedo feedback processes enabling the transition into an interglacial. It is important to note, that northern summer insolation was already high in the Alleroed and Younger Dryas.
The winter insolation does not change very much in high latitudes (it is on a low level there anyway) and has a weak minimum in middle latitudes at the beginning of the Holocene.

Chapter 5

Results

This chapter contains the most important results of this thesis. These involve the successful estimation of three dimensional transfer functions by mixture models, the reconstruction of January and July temperatures and the unsuccessful attempt to reconstruct precipitation fields.

5.1 Transfer functions

As described in the second chapter, the procedure of fitting transfer functions has been changed, compared to Gebhardt (2003). For estimating the optimal number of mixture model components, the BIC was used as criterion, instead of the log-likelihood. Further, the lower limit of the smallest eigenvalue of the mixture models was not fixed, but dependent on the eigenvalues of the initial mixture models.
It had to be assured that these changes lead to good transfer functions in all three dimensions. Further it was examined if an improvement to the transfer functions of Gebhardt (2003) is visible. Additionally a large number of new taxa has been included into the database in the last years. It was checked if transfer functions for these taxa were estimated successfully.
A comparison of some transfer functions is displayed in Fig. (5.1). In the following, the transfer functions from Gebhardt (2003) are referred to as "old", while the ones of this thesis are referred to as "new".
An improvement can be seen for *Hippophae rhamnoides* and *Fagus sylvatica*. The old function for *Hippophae rhamnoides* clearly shows overfitting with two maxima and the density falling down to under 10 percent of the maximum value between them. The new functions does not show these features. However, the region where the secondary maximum of the old function is

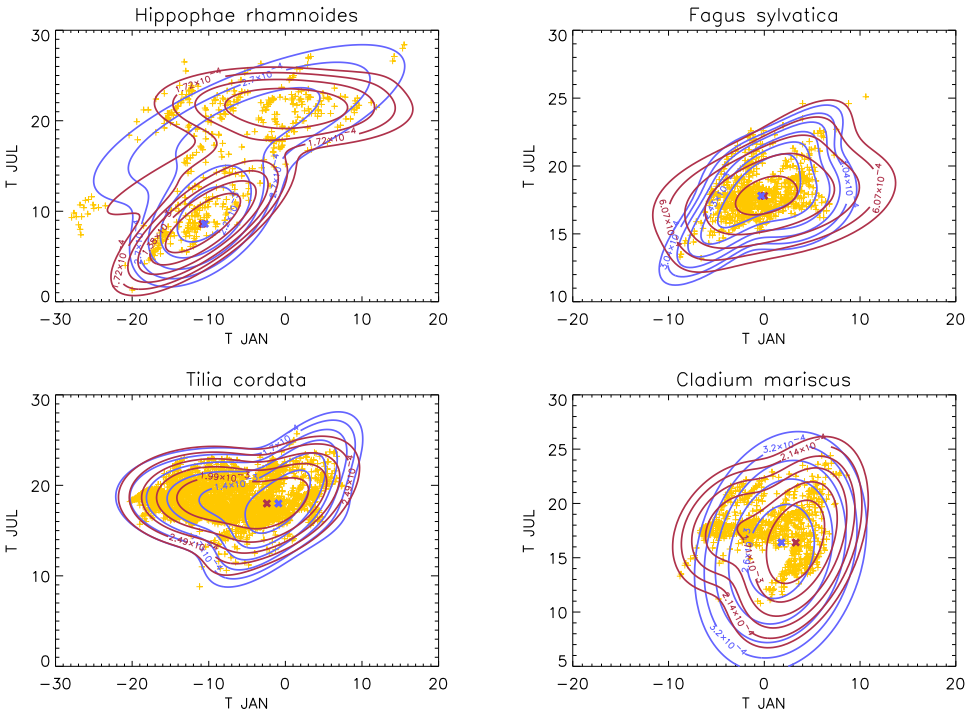

Figure 5.1: Transfer functions for the four taxa *Hippophae rhamnoides*, *Fagus sylvatica*, *Tilia cordata* and *Cladium mariscus*. Old transfer functions by Gebhardt (2003) in red, new transfer functions in blue. The lowest isoline represents 0.025% of the maximum probability density. This value is doubled for the following isolines, respectively.

found might be underrepresented in the new function. It has to be noted, that the data of *Hippophae rhamnoides* is difficult to fit. The data density is thin and data points are sometimes isolated with gaps to other data filled areas.

The new function of *Fagus sylvatica* almost perfectly captures the data points, while the old function is a bit too wide at the cold and warm end of T_{Jan}.

In the most cases the comparison between the old and new transfer functions look like for *Tilia cordata*. Differences between both functions are weak and their maxima are almost identical. Also some cases can be found where the new solution is worse than the old one, like for *Cladium mariscus*. Here the new procedure decided for a single normal distribution to be the best choice. This leads to an overrepresentation of cold July temperatures and an underrepresentation of cold January temperatures.

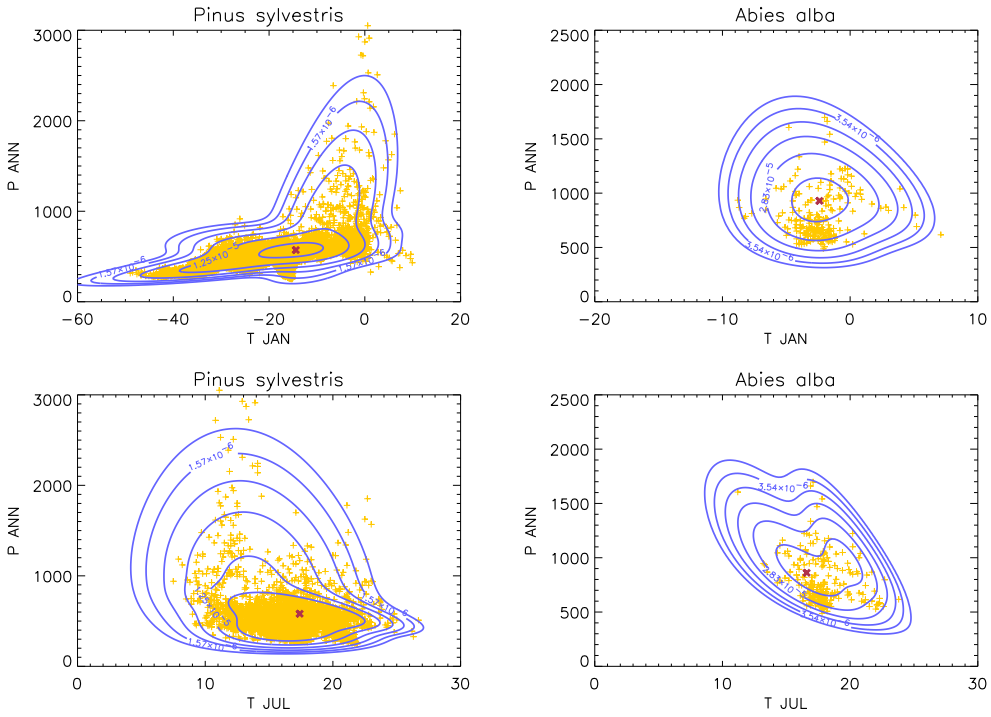

Figure 5.2: Precipitation including marginal distributions of transfer functions for *Pinus sylvestris* and *Abies alba*. Isolines are defined as in Fig. (5.1).

Generally it's hard to decide if the new procedure really leads to much better transfer functions on average. Overfitting is avoided much easier but this has the price of a relatively poor fit for some cases, like e.g. *Cladium mariscus*. However the new procedure is computationally much faster.
An illustration how precipitation is fitted can be found in Fig. (5.2). *Abies alba* is a representative example for taxa with a low number of data points with nevertheless a sufficient density. In these cases also the points with a high value of annual precipitation are well represented. This becomes clear in the T_{Jul}-P_{Ann} marginal distribution, where one value at ($\approx 11°C$, $\approx 1600mm$) is enough to address relatively high probability densities to this area.
Taxa like *Pinus sylvestris*, having a huge amount of points with low values of annual precipitation, are more problematic in this sense. The maximum then lies clearly below 1000mm and there is no chance for a representation of the few points occurring above 2500mm.
As conclusion it can be stated that three dimensional mixture models with normal-normal-gamma marginals could successfully be applied for estimating

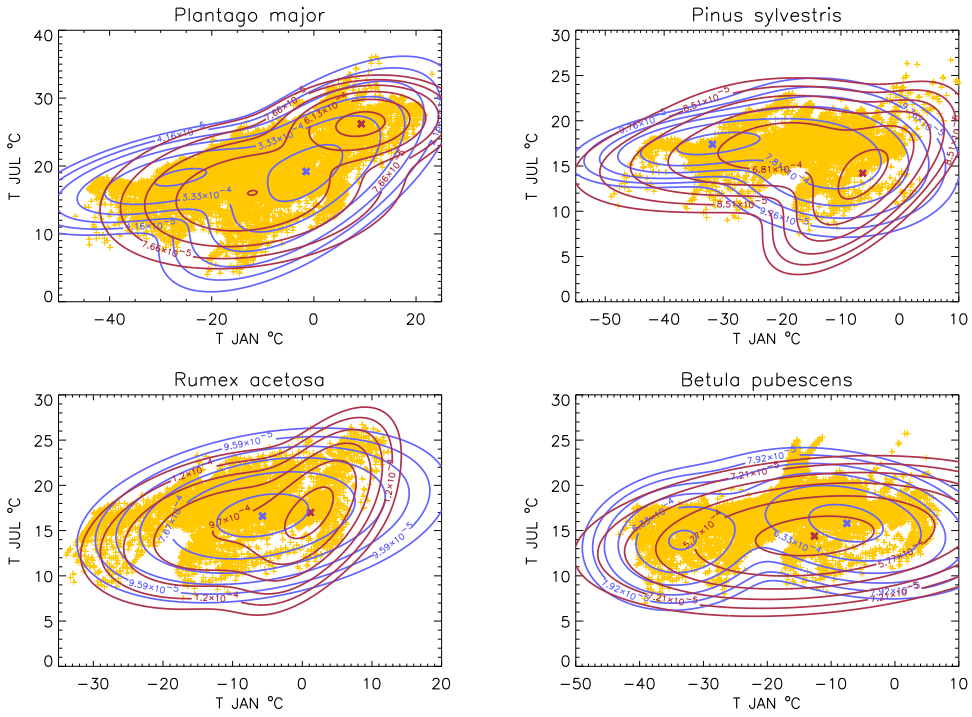

Figure 5.3: 2d T_{Jan}, T_{Jul} marginal distributions of 3d pdfs (red) and 2d T_{Jan}, T_{Jul} pdfs (blue) for the taxa *Plantago major*, *Pinus sylvestris*, *Rumex acetosa* and *Betula pubescens*. Isolines are defined as in Fig. (5.1).

climatological-botanical transfer functions. This means that the important features of the data are captured while avoiding overfitting.

5.1.1 Differences between 3d and 2d

In this section, the question will be answered, if fitting three dimesional transfer functions (pdf3d) results in different two dimensional marginal distributions for T_{Jan}, T_{Jul}. For this purpose the latter are compared to the results of two dimensional mixture models (pdf2d). These also are different to the old ones as M_{opt} is also estimated, using the BIC and initialisation by k-means clustering.

A comparison of pdf3d and pdf2d for four taxa can be found in Fig. (5.3). The taxa *Plantago major* and *Pinus sylvestris* show a typical behaviour. When precipitation is included and three dimensional mixture models are fitted, the maximum of the temperature marginal distribution is shifted to

warm January temperatures. This behaviour is caused by the weighting, applied to the climatological data (Eq. (4.1)). Very high values of annual precipitation are hardly observed in the climate system. Therefore, they get a high weight. As high values of annual precipitation generally are connected with high January temperatures, these automatically get a high weight too, in the 3d analysis.
This feature can have a big impact, especially when the taxon has a wide range of January temperatures. The different T_{Jan} maxima lie about 25 K apart from each other in case of *Pinus sylvestris*.
The examples of *Rumex acetosa* and *Betula pubescens* show that also the number of mixture model components may change between pdf3d and pdf2d. *Rumex acetosa* has only one component in the two-dimensional fit while it has two in the three-dimensional one. With *Betula pubescens* it is the other way round, as it gets two components in pdf2d and only one in pdf3d. Here pdf2d has a bimodal shape with two equally pronounced maxima.
For reconstructions it can be problematic when many taxa, with transfer functions having their maxima shifted to warm Januray temperatures, are combined. Then the reconstruction result could be biased and overestimate January temperatures. What impact the differences between pdf3d and pdf2d have, will become clear when the reconstruction results are discussed.

5.1.2 The smoothing criterion

The procedure of fitting transfer functions can certainly still be improved in the point of smoothing. As already mentioned the application of the BIC does not avoid overfitting in all cases. Overfitting normally can be related to a cluster becoming too small, what in most cases happens at the edge of the distribution.
At the moment, overfitting is avoided by giving a lower limit to the smallest eigenvalue of the mixture model covariance matrices. As limit generally serves the smallest eigenvalue of the clusters used for initialisation of the EM-algorithm. Unfortunately this criterion does not allways produce satisfying results. Sometimes the criterion is not strong enough and overfitting occurs despite the limit. In other cases, the criterion is too strong and the transfer function becomes too wide or imprecise. In both cases the limit has to be adjusted for the respective taxon, e.g. via multiplying by 2 or 0.5.
It would be a desirable improvement if the limit could be coupled to properties of the given dataset in a way that it is automatically detected, which limit is necessary. Doing this would require an identification of the data properties causing overfitting.
Another approach might be to limit the maximum number of mixture model

components to three for taxa with a comparably small number of data points. One reason for overfitting is the choice of $M_{opt} = 4$ on a relative small dataset. Then the initial clusters are small and thus the smallest eigenvalue of the initialisation is also. This could be a reason for the limit being too weak. Anyway overfitting mostly occurs when the optimal number of components is set to four. This is a point that can be addressed in future work.

5.2 Analysis areas

Before the reconstruction results for the respective time slices will be presented, the analysis areas will be described. For the Late Glacial the number of available sites is generally smaller than in the Holocene. Accordingly, the analysis area for the Holocene is larger than for the Late Glacial.
In Fig. (5.4) the analysis area for the Late Glacial time slices is displayed. The area covered by paleodata differs only slightly between 13000 and 12000 BP. Sites occurring only in one time slice are mostly located in the centre of the are and thus do not affect the boundary. There are only two sites in the southwest and one in the north at the edge changing, requiring the boundary to be shifted somewhat. From 13000 to 12000 BP the northern boundary, as well as the southern, is shifted slightly to the north. In the end, the area of 12000 BP is four grid boxes larger than the area of 13000 BP.
The Late Glacial reconstruction area coveres Central and Eastern Europe, parts of Southern Scandinavia and also the Northern Mediterranean regions. A dataless area is caused by the North Sea and most of France is a big region without any available site. Data are also sparse in the Southeast of the area.
The reconstruction area for both Holocene time slices is the same, as only one site does not occur in these both. This site does not lie at the edge of the area. In the Holocene there are a lot more sites available, than in the Late Glacial. These additional sites expand the possible reconstruction area to most parts of Scandinavia and more south in the Mediterranean. This area is visible in Fig. (5.5). Here the grey border lines mark where the additionally available regions can be cut off. For the Holocene time slices, three reconstruction settings where used. One with the whole area, another with Scandinavia and site 70 cut off and finally one with Scandinavia and the southernmost area cut off. The properties of the different settings are summarised in Tab. (5.2).
For all reconstruction settings a 1° grid spacing was chosen, although climate data as well as orography data was available in 0.5°. However, changing the grid spacing to 0.5° would result in a four times larger number of grid boxes of the analysis area, while the number of sites would only be increased by

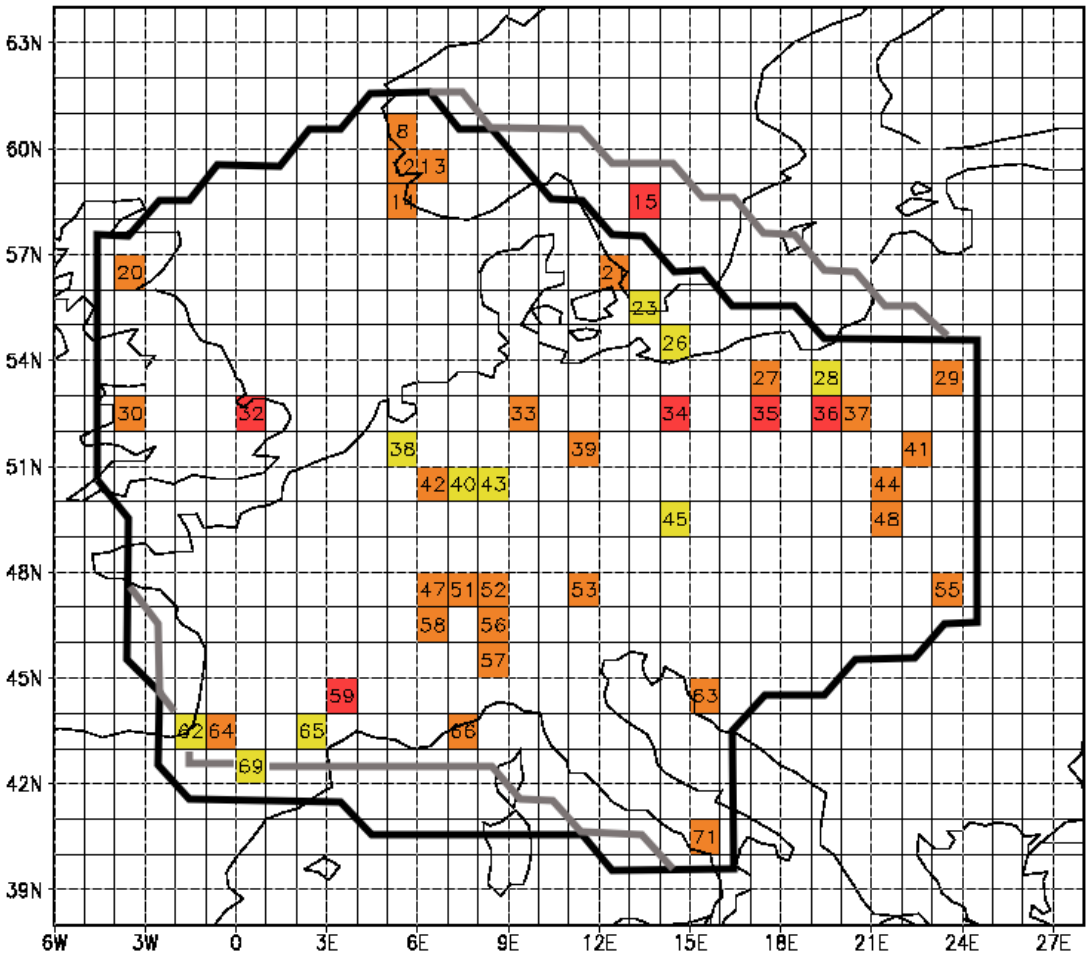

Figure 5.4: Analysis area for the time slices 13000 and 12000 BP. Orange sites occur in both time slices, yellow ones only in 13000 and red ones only in 12000. The black border is the 13000 border. The 12000 border differing from the 13000 is drawn in grey. Site numbers refer to the tables in the appendix.

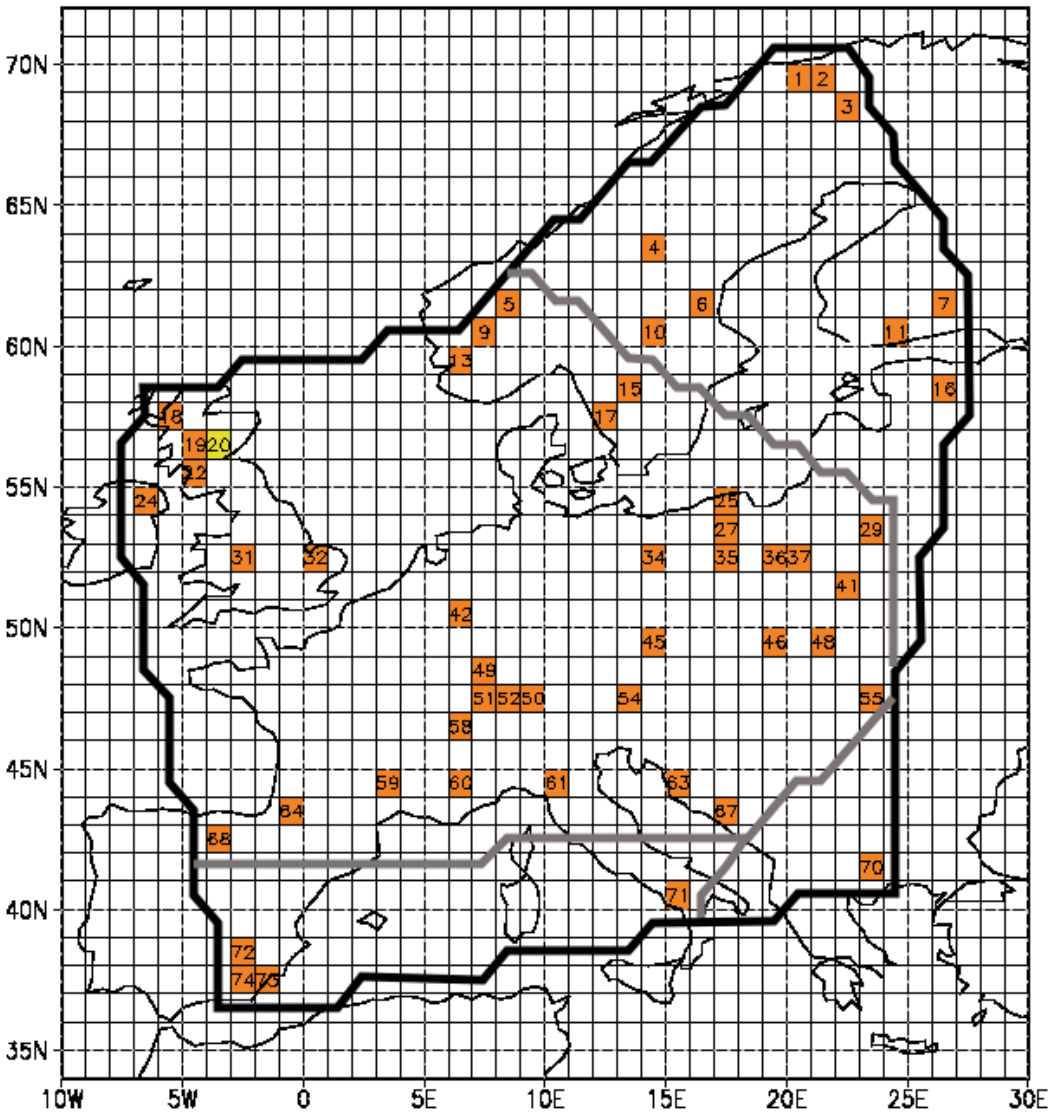

Figure 5.5: Analysis area for the time slices 8000 and 6000 BP. Orange sites occur in both time slices, the yellow one only in 8000. The black border represents the complete reconstruction area for both time slices. The grey borders mark sections that can be cut off from the area. Site numbers refer to the tables in the appendix.

Time slice	min. lon	max. lon	min. lat	max. lat	#Sites	#grid bx.
13000	-3.5	23.5	40.5	60.5	38	493
12000					34	497
8000	-6.5	26.5	37.5	69.5	52	865
8000b					38	552
8000c					48	641
6000					51	865
6000b					37	552
6000c					47	641

Table 5.1: Properties of the different reconstruction areas. The letter "b" means the setting with Scandinavia and southernmost part cut off, "c" means Scandinavia and site 70 cut off.

10-20%. Which grid boxes contain more than one site, is visible from the tables in the appendix. A unique number is addressed to every grid box containing paleovegetation data. The borders are always defined in a way that no extrapolation from sites is necessary.

5.3 Reconstruction results

In this section reconstruction results for the different time slices will be presented. This presentation will emphasise to regard the resampled fields as reconstruction result. The discussion will be held with respect to that fact. Before the results for the paleodata are displayed, it will be shown how the method performs, when it is applied to modern vegetation data. The information, to which point the method is able to reproduce the modern climate, is important in order to interpret the reconstruction results.

5.3.1 Important aspects for the analysis

Several important characteristics of the variational analysis are described by Gebhardt (2003) very detailed. As the general procedure was not changed during this thesis, this should not be repeated. However, some aspects should be mentioned.

Number of singular modes: First, the number of singular modes that are regarded in the analysis, has to be determined. The maximum possible number of modes is constrained by the number of sites which are incorporated in the analysis. This number is the theoretical maximum of the spatial

degrees of freedom. In reality the number of degrees of freedom is smaller, because of spatial correlations within the data.
The most effective number of singular modes for the analysis is defined by the spatial degrees of freedom. This number is found by searching the minimum of the largest eigenvalue of Σ_R. This eigenvalue is the inverse of the smallest eigenvalue of the Hessian at the cost function's minimum. When this eigenvalue is very small it means, that the cost function gets extremely flat in the direction of its associated eigenvector. This makes the minimum hard to find. A small largest eigenvalue of Σ_R and thus a large smallest eigenvalue of the Hessian stands for a better defined minimum of the cost function.
A second criterion for the number of modes is, that it has to be larger than 25. This number is necessary for capturing all important characteristics of the anomaly fields.

Filtering: When the eigenvectors of Σ_R are projected into the climate space, regions of the analysis area can be identified, that lead to high values of the eigenvectors. Often, the eigenvectors with the highest explained variances have their maxima at the boundary of the analysis area or in regions with a low data density. So, they often have a considerable influence on the analysis result, although the data projection on them is low. These patterns are filtered out of the analysis by setting their solution projection to zero.

5.3.2 Reconstruction of the modern climate

The attempt to reconstruct the modern climate was already carried out by Gebhardt (2003). For January, this reconstruction revealed positive anomalies of about 2K in Scandinavia and slightly negative anomalies in Southwestern Europe. Anomalies down to -2K were reconstructed for parts of the Alps. In July, anomalies were close to zero in most of the area.
This experiment has been repeated for this thesis, using a different setting. Gebhardt reconstructed the modern climate by considering all taxa, occurring today. Here this was done by regarding only these taxa, that also were present in the time slice 8000 BP. This limitation produces results that are better comparable. Further it can possibly give an answer to the question if the ability for reconstructing climate is limited by the subset of taxa, that are found in the sediment.
The reconstruction results for modern vegetation are based on the same reconstruction setting as the 8000 BP reconstruction. Especially it should be noted, that vegetation is only included at the same location where data was also available in the past.

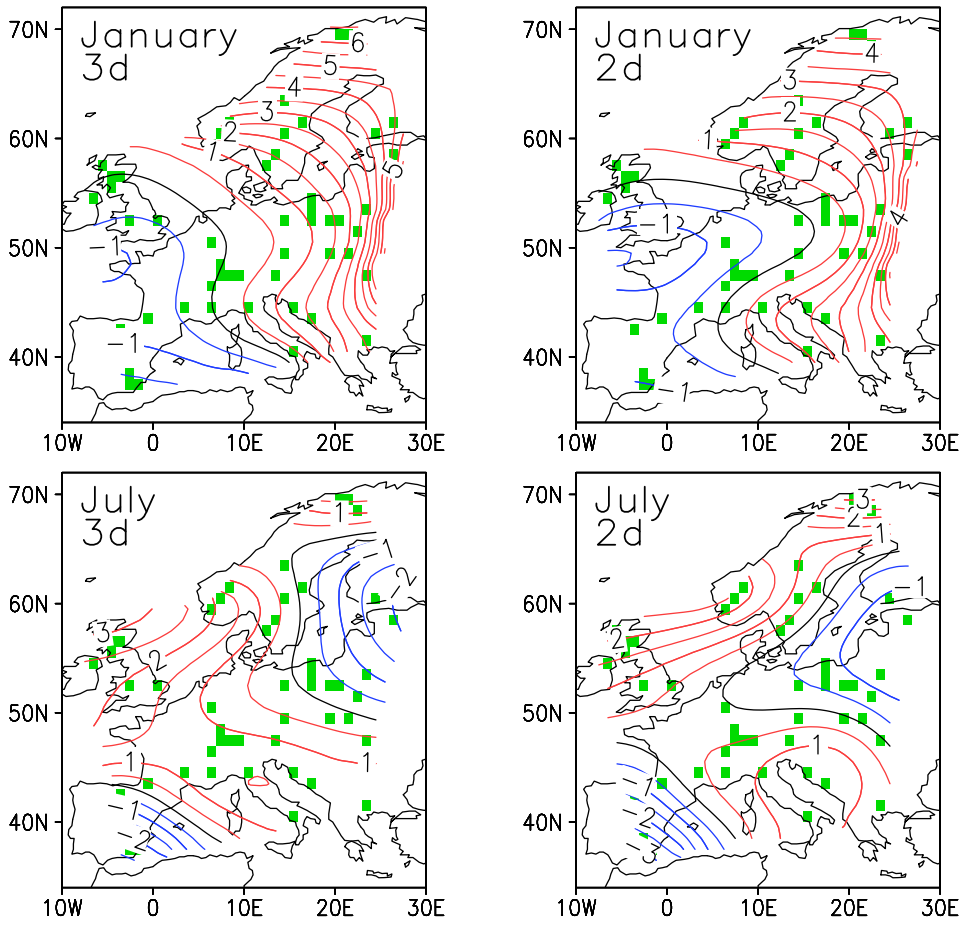

Figure 5.6: Reconstructed January and July temperature anomalies for modern vegetation with pdf3d (left) and pdf2d (right). Green grid boxes indicate grid boxes with vegetation data.

The reconstructed temperature fields for analyses with pdf3d and pdf2d are presented in Fig. (5.6). Here, the temperature fields that minimise the cost function are shown. The results of the resampling are not presented here, because the spatial patterns do not vary a lot.
The solution patterns of the analysis with pdf3d and pdf2d are basically the same. In January, high positive anomalies are retained in the northeast of the area and along the complete eastern boundary. For Central Europe, the British Isles and the Northern Mediterranean, the anomalies are close to zero. Only in the far southwest (pdf3d) of the area or in the west of France (pdf2d), the anomalies are significantly negative.
For July the analysed temperature anomalies are close to zero for most of the

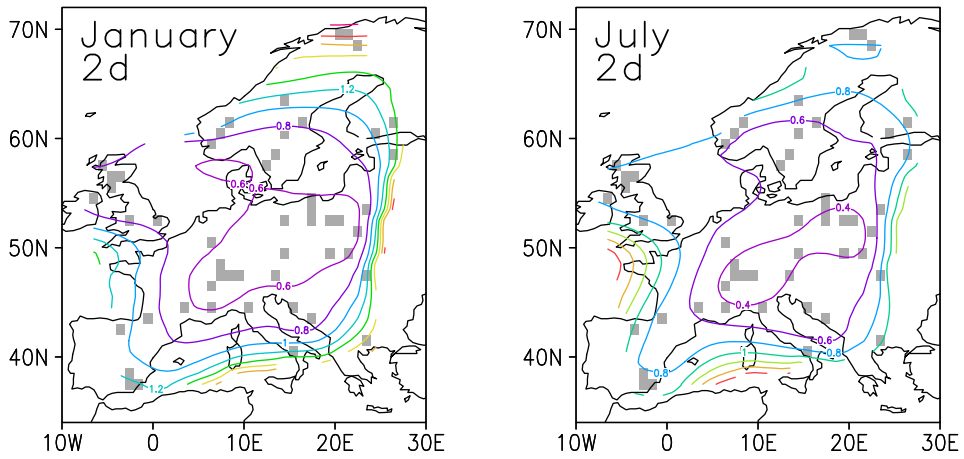

Figure 5.7: Fields of standard deviations from the analysis with pdf2d. Grey grid boxes indicate grid boxes with vegetation data.

area, except the British Isles and the North Sea. A ridge of anomalies above 1 K ranges from France to the Italian peninsula in the pdf3d analysis. The pattern is very different, compared to January, where a strong southwest-northeast gradient was present.
The modern climate is slightly better reconstructed by pdf2d. Positive anomalies in Northern Scandinavia for January and in Britain for July are weaker than with pdf3d. Negative anomalies over the Baltics in July and positive anomalies in Central Europe also are gradually weaker.
The results presented here are similar to those obtained by Gebhardt (2003), at least concerning the positive anomalies in the northeast in January and in northwest in July. However, the negative anomalies over the Alps, analysed by Gebhardt, are not present here.

Fields of standard deviation for the analysis with pdf2d are shown in Fig. (5.7). These are obtained by resampling, as described in Section (3.4.1), with a sample size of 1000. The standard deviations tend to increase near the boundary, as the information there is only available from one direction. High values are also present in Northern Scandinavia in January, where data density is low and taxa with a wide range in T_{Jan} are present. In Central and Eastern Europe, with high data density, standard deviations are very low. Generally, the values are lower in July than in January.

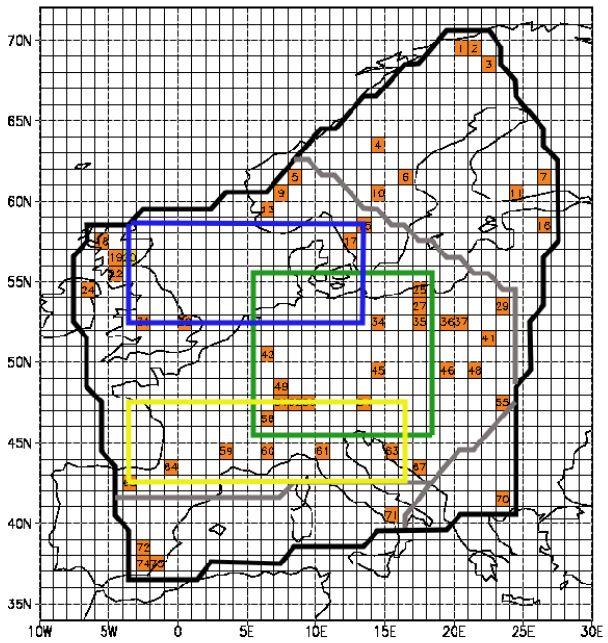

Figure 5.8: Overall reconstruction area for the modern reconstruction with areas, defined for calculating area means. These are NW in blue, CE in green and SW in yellow.

Area means

For making a quantitative comparison of the results for the different time slices, three different areas for calculating area means are defined. First, these are a Central European (CE) area, ranging from 5.5°E to 18.5°E and from 45.5°N to 55.5°N and a northwestern area (NW) ranging from 3.5°W to 13.5°E and from 52.5°N to 58.5°N, covering the North Sea. Finally a southwestern area (SW), representing the northern Mediterranean region, ranges from 3.5°W to 16.5°E and from 42.5°N to 47.5°N. A map showing these areas, can be found in Fig. (5.8). They are defined in a way, that they are placed within the reconstruction areas of all time slices.
An overview over the mean values and uncertainty ranges of the different areas for the pdf2d result is given in Tab. (5.2). The mean anomalies are zero or near zero for all areas in January. A slightly positive anomaly is reconstructed for NW in July, what was expectable from Fig. (5.6). In CE and SW the anomalies for July are not that near to zero as in January but still very small. All area mean values are not significantly different from zero. Thus, it can be stated, that the method succeeded in reconstructing the modern climate.

month		NW	CE	SW
January	0.025-quantile	-0.9	-0.8	-0.8
	mean	0.1	0.0	0.2
	0.975-quantile	1.1	0.9	1.1
July	0.025-quantile	-0.2	-0.3	-0.4
	mean	0.8	0.3	0.4
	0.975-quantile	1.7	0.8	1.1

Table 5.2: Reconstruction of modern temperatures by pdf2d: Mean values and quantiles for the 95% confidence interval of the three areas defined in Fig. (5.8). All values are in K.

Reasons for biases in the reconstruction

In the previous section it became clear that the reconstruction method fails to correctly reconstruct modern January temperatures in northern Scandinavia and July temperatures around Britain. The local reconstruction for the site "Toskaljavri" (Fig. (5.9), top), located in northern Sweden, is a good example for the reason for failing January temperature reconstructions.
First, only some of the seven taxa that passed the preselection have a relative narrow pdf and thus a relative high influence on the reconstruction result. Those are *Salix herbacea, Picea abies and Myriophyllum alterniflorum.* All these three taxa occur at the cold end of January temperatures, possible for them, and not in their optimal climate conditions. Other taxa, like *Juniperus communis* or *Pinus sylvestris*, also occurring in regions with lower winter temperatures have a much too wide range and thus nearly no influence on the reconstruction result.
Also in July most taxa are occurring there at the lower end of their possible ranges. However, as the July range is typically smaller than the January range, the bias is smaller here.
The local reconstruction of the Scottish site "Loch Maree" can be found in Fig. (5.9, bottom). Here, the reconstructed pdf has a maximum that is much more narrow than in "Toskaljavri". The climate of the British isles is maritime with mild winters and relatively cold summers. Taxa, preferring temperate climate, like *Typha latifolia, Hedera helix* or *Fraxinus excelsior* occur there at the cold end of their July temperature range. They exist in Scotland mainly due to the mild winter temperatures, which are normally associated with higher summer temperatures. As there are no taxa present,

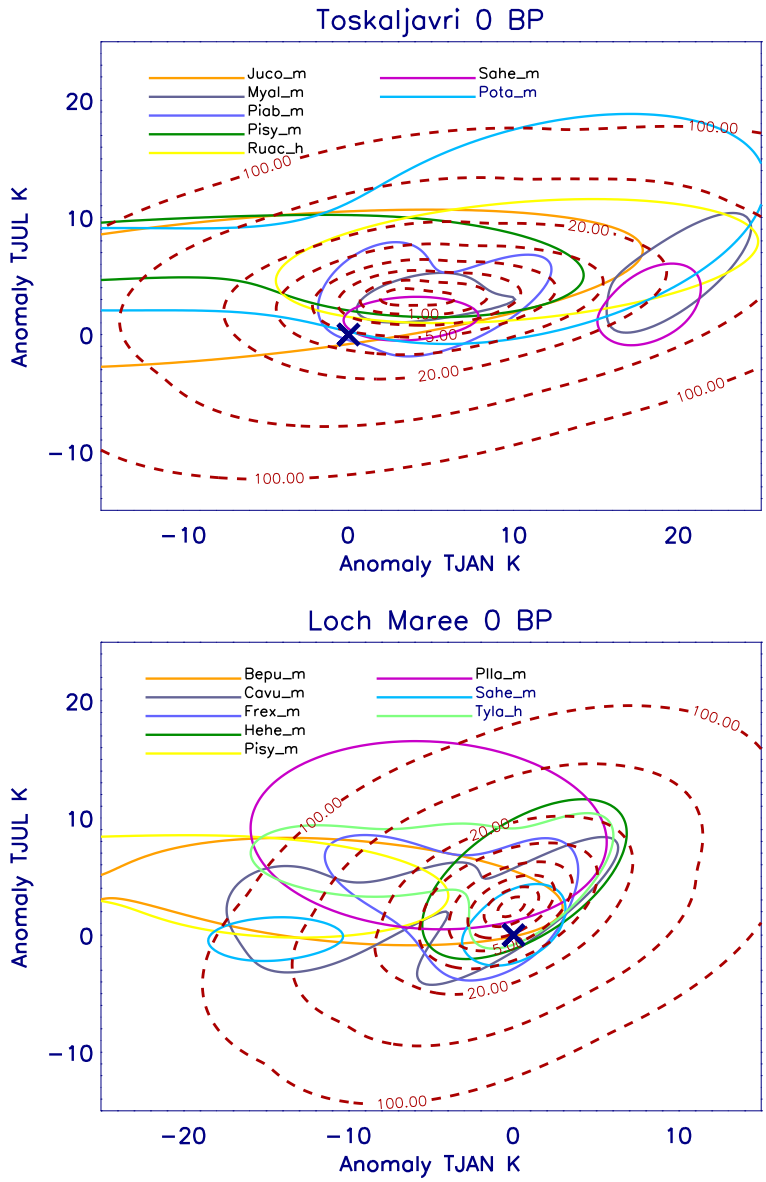

Figure 5.9: Local reconstruction for the sites Toskaljavri and Loch Maree. Pdfs of occurring taxa are represented by their isolines of 33% of the maximum density. The resulting pdf (red dashed lines) is displayed in terms of vegetational costs (ln of the pdf values normalised by the maximum of the pdf). The zero anomaly point of modern climate is marked with a blue X.

which prefer colder summers, the actual July temperatures cannot be reproduced by this reconstruction approach.

At the end of this section, two conclusions for the interpretation of the reconstruction results can be made. First it can be stated, that the reconstruction of July temperatures works well for most of the area. Merely, the results around the British Isles and at the very northern part of Scandinavia should be regarded carefully. Second, the reconstruction of Scandinavian winter temperatures fails.
These problems should not affect the comparison of area means for the different time slices. On the one hand, for January, the areas lie far away from problematic regions. On the other hand, the weak positive July bias in the NW area should not affect the reconstruction of relative changes very much. This region should always have been characterised by maritime climate, at least with the given distribution of land masses.

Paleorecontruction results

In the following section the reconstruction results for the four time slices will be presented and briefly described. A more detailed discussion of the results will follow in the next chapter. There the results of the different time slices will be compared among each other and also with the results of other reconstruction approaches.

5.3.3 Results for 13000 BP

The reconstruction for 13000 BP is based on paleovegetation data from 38 grid boxes. The reconstruction area is shown in Fig. (5.4). An effective number of 29 singular modes was determined for this analysis setting and no eigenvectors had to be filtered. The anomaly fields, minimising the cost function are displayed in Fig. (5.10) for the pdf3d and pdf2d reconstructions. The differences between these two results are very large in January. Over Britain the pdf3d results are up to 5 K warmer. The differences become smaller in all directions from this point, but remain positive for nearly all of the area. These results imply that January temperatures are overestimated, especially in regions where mainly taxa with a wide range in the January temperature are present.
In July the differences are weak. The pdf3d reconstruction is slightly colder in the center of the area and slightly warmer in the southwest. Together with the findings of modern temperature analysis, where pdf2d performed better, it now can be concluded, that regarding temperatures and annual precipitation together has no positive effect, at least for the European region. Therefore, in the following (also for the other time slices) only the results of the pdf2d reconstructions will be discussed.
The mean reconstructed field for January shows very strong negative temperature anomalies below -9 K over Britain. A strong gradient in southeastern direction is present. The anomalies lie around -6 K in Central Europe and reach values around 0 K at the southeastern border of the reconstruction area. For July the anomalies are slightly negative over most of the area with a maximum of values close to zero in Northern Britain. The lowest values are present at the eastern boundary and in the Southwest where anomalies below -3 K are reached.
Fields of standard deviation, representing the analysis error, can be found in the appendix (Fig. (A.1)). They show the typical pattern of lowest values in the centre of the area and largest values at the boundary. The errors are generally smaller in July than in January. This also becomes clear, when looking at the analysis error histograms (Fig. (A.3)). Over the North Sea, there are the clearest differences between the two month. There the July errors have a second miniumum, while January errors increase more continously from the center of the area. The highest July errors are present where it has to be interpolated over large regions without data (Western France and Mediterranean Sea).

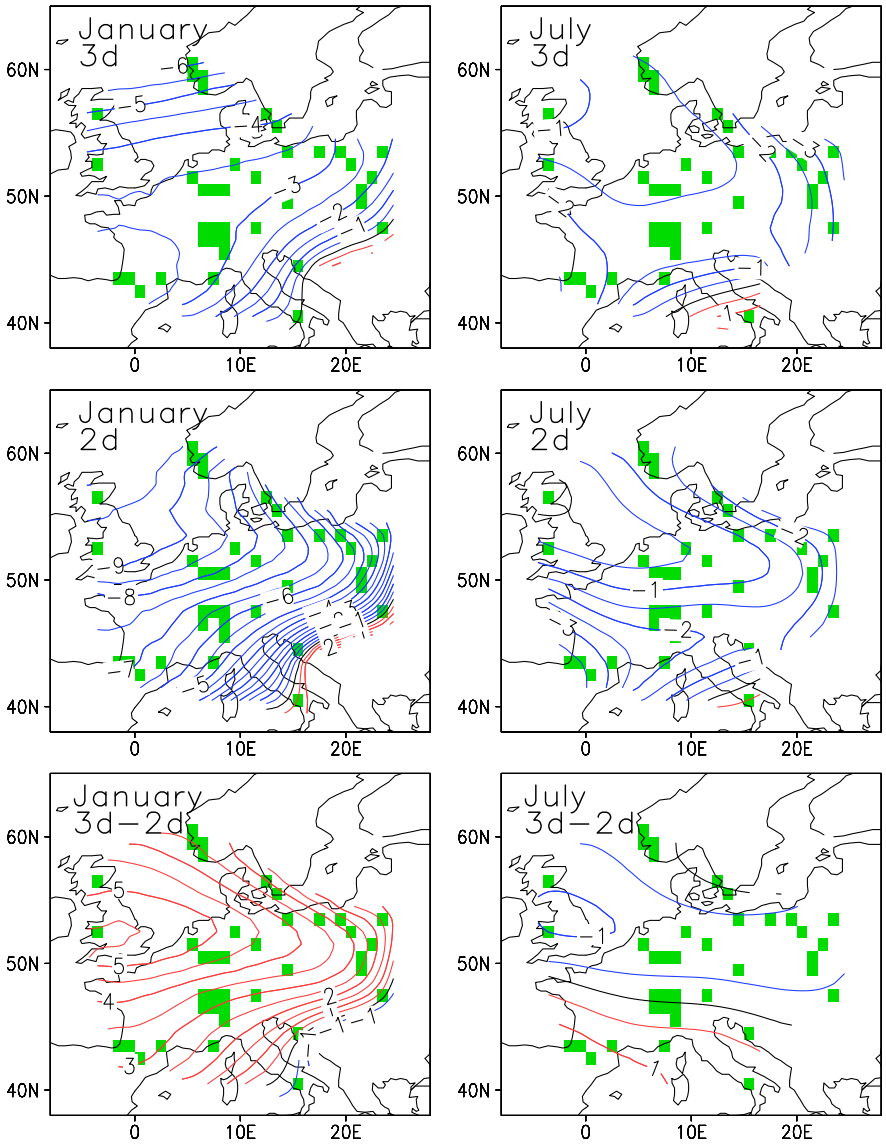

Figure 5.10: Reconstructed January (left) and July (right) temperature anomalies for 13000 BP with pdf3d and pdf2d as well as differences between pdf3d and pdf2d. Green grid boxes indicate grid boxes with vegetation data. Distance of isolines is 0.5 K, negative and positive anomalies are indicated red and blue, respectively.

month		NW	CE	SW
January	0.025-quantile	-11.3	-8.2	-7.2
	mean	-8.6	-6.6	-5.3
	0.975-quantile	-6.1	-4.9	-3.3
July	0.025-quantile	-1.8	-2.1	-3.2
	mean	-0.6	-1.2	-2.1
	0.975-quantile	0.7	-0.3	-0.9

Table 5.3: Reconstruction of 13000 BP: Mean values and quantiles for the 95% confidence interval of the three areas defined in Fig. (5.8). All values are in K.

Resampling results

The results of resampling, presented in Fig. (5.11) and Fig. (5.12), give an impression which realisation of the 13000 BP temperature anomaly fields are possible under the statistics of the analysis error covariance matrix, given the proxy data. Especially the uncertainty in the spatial patterns becomes clear by regarding the results of the resampling procedure. The results of resampling are always represented by 12 fields, randomly chosen from a sample with size 1000. These fields are sorted by the Mahalanobis distance of their respective coefficient vectors (in the space of singular modes) to the coefficient vector minimising the cost function. This is done in a way, that fields 1-3 are nearest and fields 10-12 have the largest distance.
All fields for January show strong negative temperature anomalies in the Northwest and close to zero or slightly positive anomalies in the Southeast. Generally in January, most realisations show a pattern similar to the mean field with a gradient from the North Sea to the Southwest of the area.

Some uncertainty in the spatial patterns can be found over the North Sea, where field 3 shows a north-south gradient with lowest values of -11 K at the northern boundary. In field 4 the lowest values (-13 K) are found over Britain and the gradient over the North Sea is east-west. In field 8 a broad minimum over the North Sea, without any gradients, can be found. Field 5 shows that also a pattern with lowest anomalies in Central Europe is a possible realisation.
For July the resampling shows a variety of different patterns with a maximum of slightly positive anomalies located either in the Northwest or over Central Europe. Large differences are found in Western France, where field 4 shows a realisation with values close below zero and field 8 shows a realisation with deep negative anomalies.

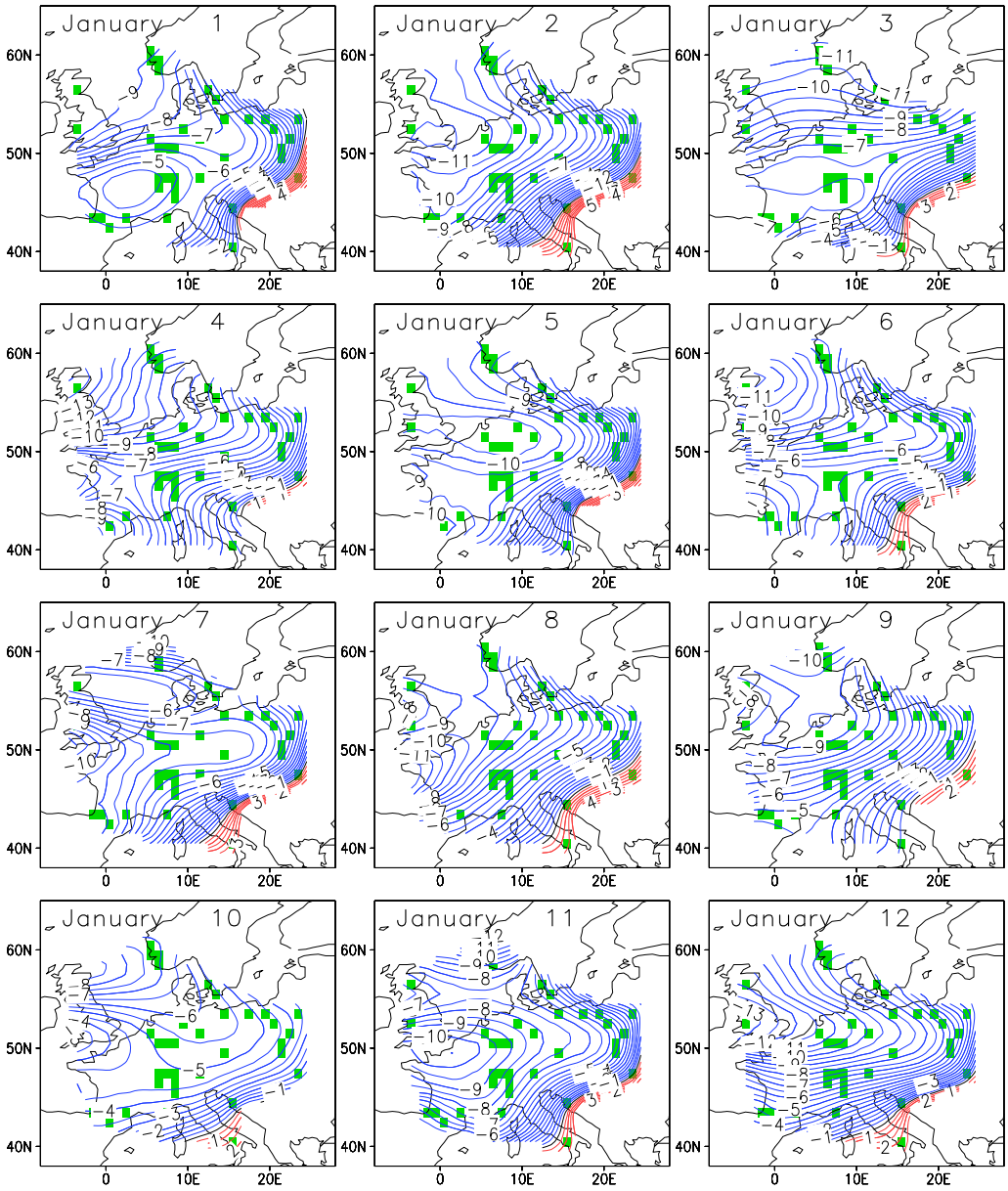

Figure 5.11: Resampling results for the 13000 BP January temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimising coefficient vector (Eq. (3.73)).

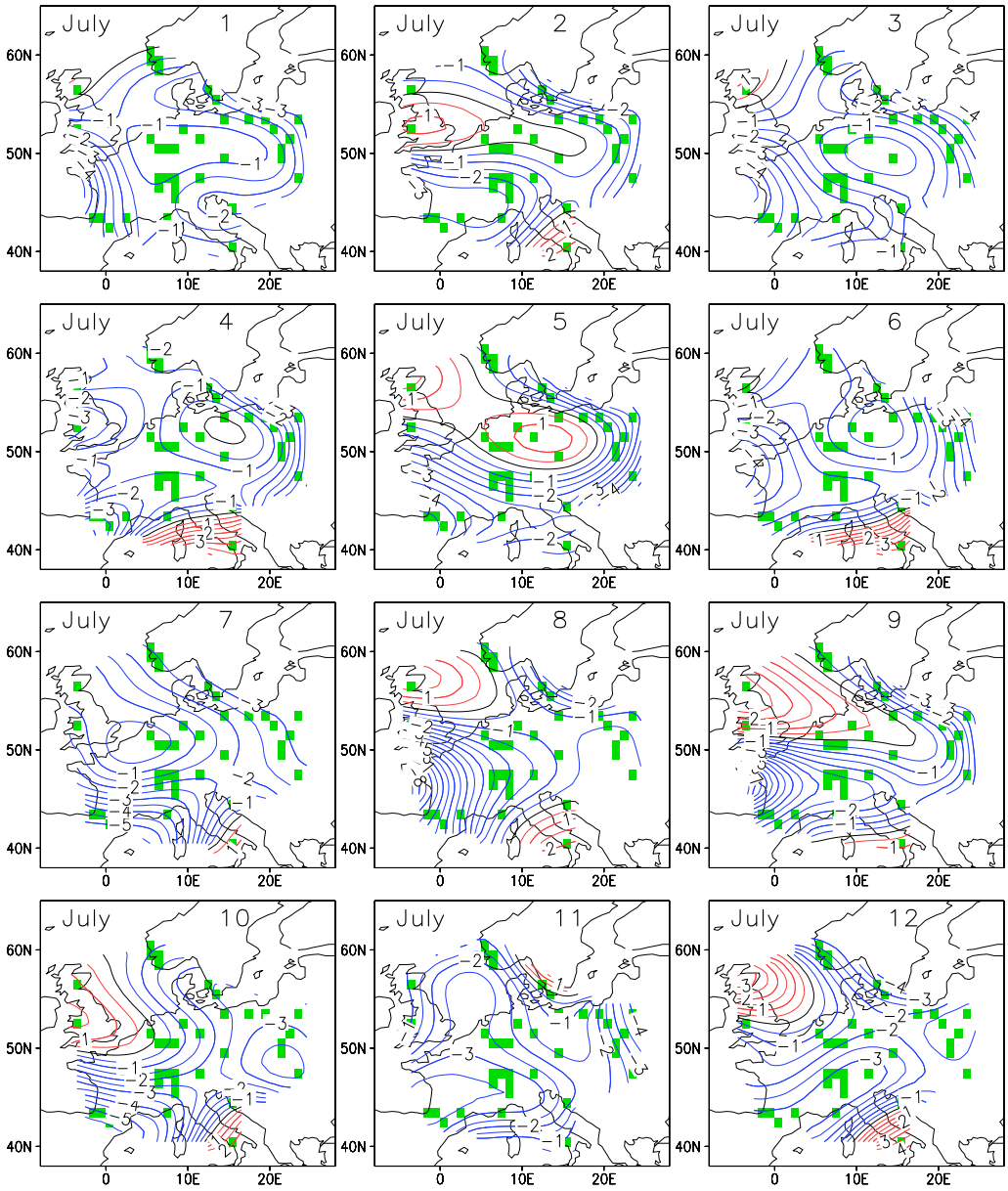

Figure 5.12: Resampling results for the 13000 BP July temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimising coefficient vector (Eq. (3.73)).

The area means, shown in Tab. (5.3), are clearly negative for January in all three areas. Even the 0.975-quantiles have values far below zero. The largest negative anomaly is found in NW, but also the uncertainty is largest here, with a 95%-confidence interval of more than 5 K width, compared to less than 4 K in the other areas.
The July means show an inverse behaviour compared to the January means, with largest values in NW and smallest in SW. They are significantly colder than the 1961-90 values in CE and SW.

Quality control

The sites, contributing to the reconstructed field, can be quality controlled by regarding their contribution to the vegetational costs. When one certain site disagrees substantially with the neighbouring sites, this would be reflected by high vegetational costs at this site. Hence, possible errors in the vegetation dataset or locations being improper can be identified. The costs at each site for the 13000 BP reconstruction is shown in the appendix, in Fig. (A.2). Similar costs arise at most of the sites. One outlier is found in the Alps with site 56 (Gola di Lago). There only three taxa are present and one of them (*Larix*) represents relatively cold winter temperatures, leading to a very cold local reconstruction this far south. However, due to the low number of taxa, the local reconstructed pdf is very wide. Therefore it does not influence the reconstruction very much. As more taxa occur at the neighbouring sites, these constrain the solution much more. The ability to ignore outliers is definitely a strength of this approach.
The situation is similar at the Scandinavian site 13 (Vestre). There, the taxon *Koenigia islanica* occurs at the neighbouring site Utsira and constrains the result to cold anomalies there because it has a narrow range. As the local pdf at Utsira is relatively thin, the solution is pulled to colder January anomalies, resulting in higher costs at Vestre.

5.3.4 Results for 12000 BP

As displayed in Fig. (5.4) the reconstruction area of 12000 BP differs not much from the area of 13000 BP. The reconstruction here is based on data from 34 grid boxes and was carried out by 27 singular modes. The recon-

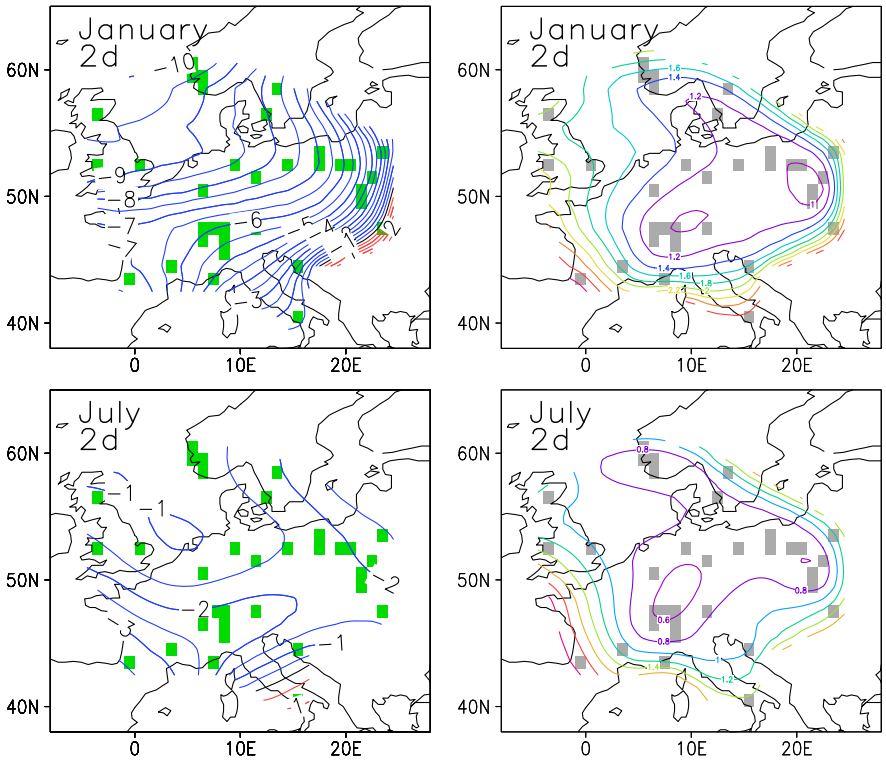

Figure 5.13: Reconstructed January and July temperature anomalies for 12000 BP and associated fields of standard deviation. Green grid boxes (left) and grey grid boxes (right) show grid boxes with vegetation data. Distance of isolines is 0.5 K for the anomalies and 0.2 K for the standard deviation. Negative and positive anomalies are indicated red and blue, respectively.

structed mean fields and the associated standard deviations can be found in Fig. (5.13).
Both, the mean January and the mean July temperature anomaly fields are very similar to the fields, reconstructed for 13000 BP. Again, the January shows a strong gradient of anomalies from the northwest to the southeast. The values, ranging from -10 K to 2 K are similar to the result of 13000 BP, as well as the fields of analysis errors.

month		NW	CE	SW
January	0.025-quantile	-11.4	-8.3	-7.2
	mean	-8.8	-6.7	-5.0
	0.975-quantile	-6.4	-5.2	-2.8
July	0.025-quantile	-2.4	-2.5	-3.4
	mean	-1.2	-1.7	-2.0
	0.975-quantile	0.0	-0.8	-0.7

Table 5.4: Reconstruction of 12000 BP: Mean values and quantiles for the 95% confidence interval of the three areas defined in Fig. (5.8). All values are in K.

Resampling results

Concerning the resampling results for January (Fig. (5.14)), it can be stated that the pattern of strong negative anomalies in the Northwest and a strong gradient to the southeastern boundary is present in all realisations. The position of the minimum varies a bit between Northern Britain (field 3), Norway (field 8) or the North Sea (field 11). It seems so that variability in spatial patterns is not that high here, compared to 13000 BP. However one should be careful in concluding so, as 12 realisations might not give a fully comprehensive impression of the uncertainty.
For July it is interesting to note, that there is only one realisation (field 4) with positive anomalies somewhere in the northwest of the area. This is a clear difference to the resampling results of 13000 BP. Fields 2 and 3 don't show positive anomalies anywhere in the reconstruction area. Maxima and minima seem to be arbitrarily placed somewhere in the area.
The area means for January are nearly the same as for 13000 BP. The July mean value also is the same in SW, but slightly lower than in 13000 BP in CE and NW.

Quality control

The vegetational costs per site (Fig. (A.4)) only show two sites with relatively high costs. These are site 30 (Nant) and site 63 (Lake Vrana). At Nant only two taxa are present, making its local reconstruction very uncertain. Thus this site does not constrain the analysis very much. Nevertheless relatively high local costs occur, as the local cost function minimum lies far away from the analysed field.
At Lake Vrana the local reconstruction results in relatively low July temper-

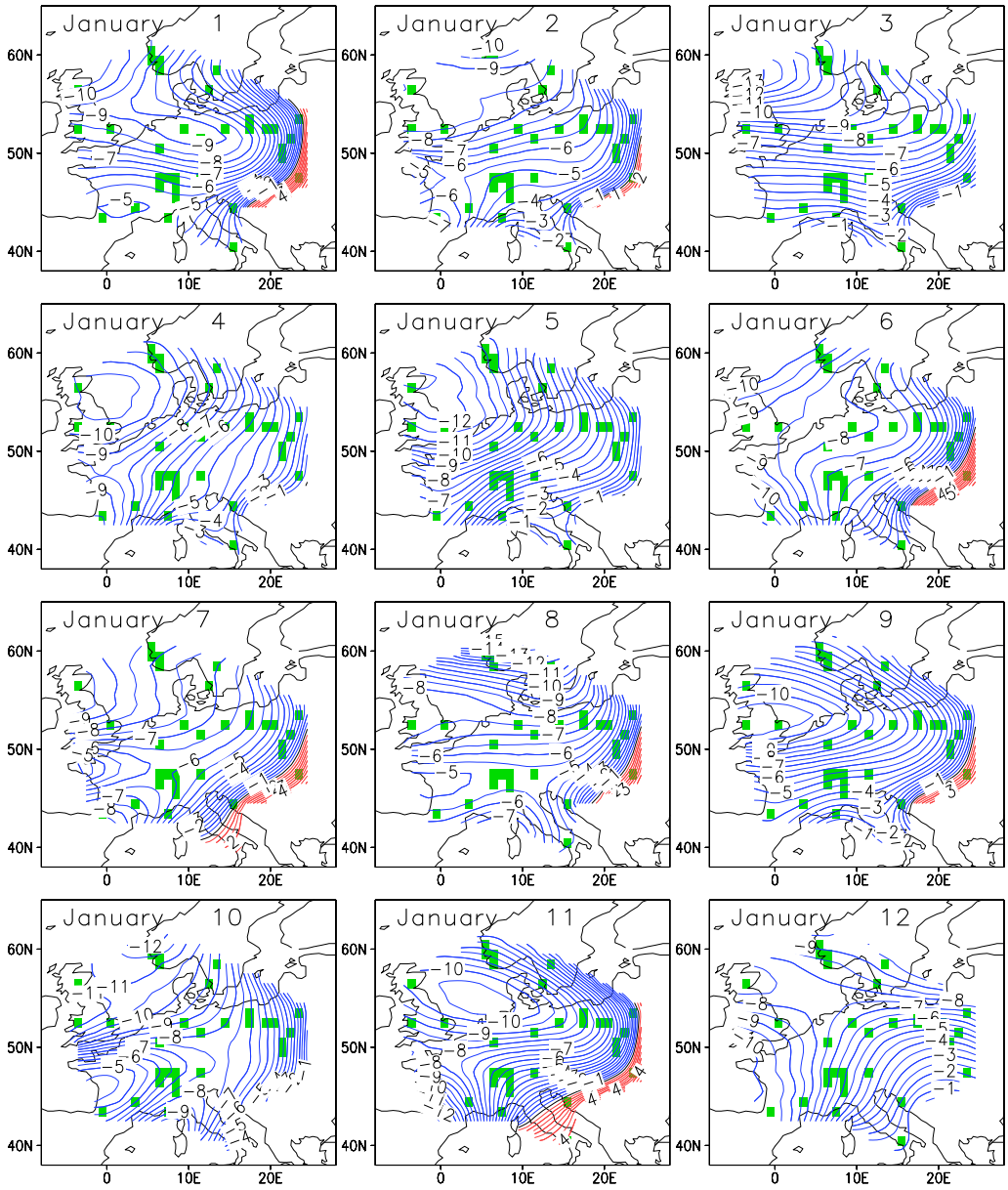

Figure 5.14: Resampling results for the 12000 BP January temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimising coefficient vector (Eq. (3.73)).

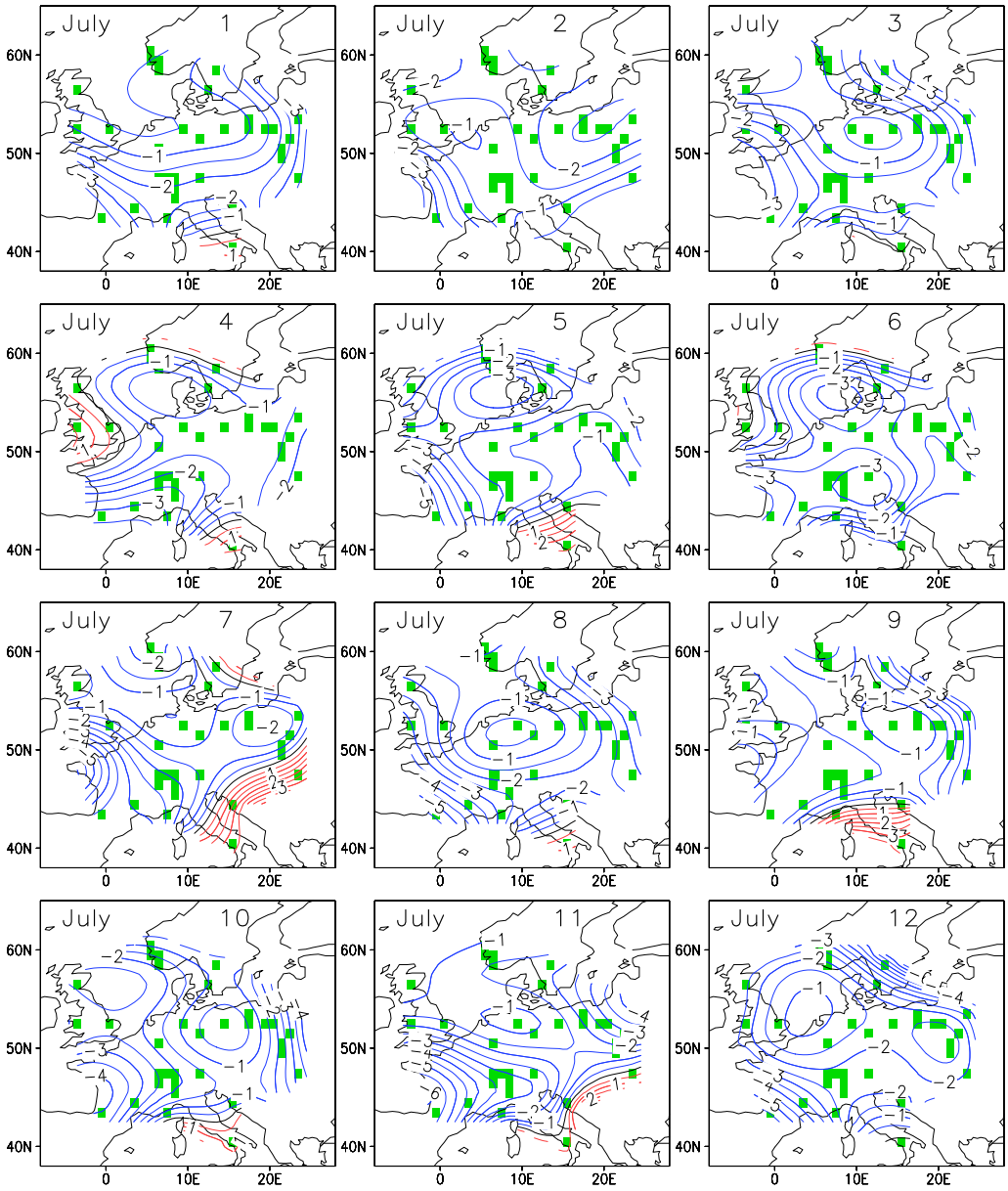

Figure 5.15: Resampling results for the 12000 BP July temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimising coefficient vector (Eq. (3.73)).

atures. As this is in disagreement to the signal from the other sites in the Southeast, this site produces relatively high costs.
Sites 44 and 48 are located in neighbouring grid boxes. As their local reconstruction differs by more than 5 K in January, the analysis results in a trade off between these both sites. When this happens, higher costs are caused by both sites.

5.3.5 Results for 8000 BP - July

The results for the 8000 BP time slice will be presented here separately for January and July. This is done because the reconstruction of modern climate showed, that the reconstruction of July temperatures is possible for the whole area. The reconstruction of January temperatures, however, is not possible for Northern Scandinavia, where high positive anomalies were reconstructed. Hence the January results will be shown later, with Northern Scandinavia omitted.

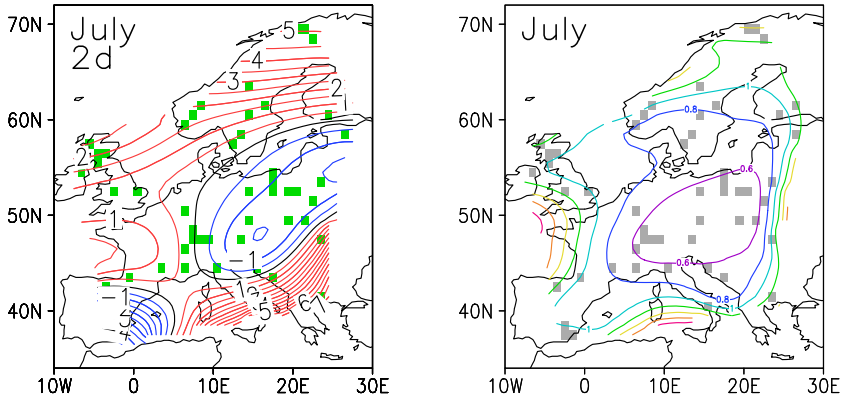

Figure 5.16: Reconstructed July temperature anomalies for 8000 BP (full area) and associated field of standard deviation (right plot). Green grid boxes (left) and grey grid boxes (right) show grid boxes with vegetation data. Distance of isolines is 0.5 K for the anomalies and 0.2 K for the standard deviation. Negative and positive anomalies are indicated red and blue, respectively.

For the reconstruction setting of 8000 BP, 35 modes were determined to be most effective. Two of them had to be filtered after the analysis. Data from 52 grid boxes contribute to the result. The reconstructed mean field is displayed in Fig. (5.16). It shows positive temperature anomalies over a large area, ranging from France over the North Sea up to Northern Scandinavia. In Central Europe the anomalies are close to zero and decrease further eastwards to values below -1 K. Another spot of negative anomalies is found in Southern Spain. Further East, strong gradients to high positive anomalies are found. This feature will be discussed in the following chapter. The pattern is similar to the one of the analysis with modern vegetation, but positive anomalies in Northern Scandinavia as well as negative anomalies in Eastern Europe are stronger here.
The analysis errors show the typical pattern of low values in the centre of

	NW	NWb	NWc	CE	CEb	CEc	SW	SWb	SWc
0.025-qtl.	-0.5	-0.6	-0.6	-1.0	-1.4	-1.4	-0.7	-1.1	-0.7
mean	0.8	0.8	0.9	-0.2	-0.6	-0.5	0.4	0.2	0.3
0.975-qtl.	2.1	2.2	2.3	0.4	0.2	0.4	1.5	1.3	1.3

Table 5.5: July reconstruction of 8000 BP: Mean values and quantiles for the 95% confidence interval of the three areas defined in Fig. (5.8). NW, CE and SW indicate the full area reconstruction setting, letters b and c indicate setting 8000b and 8000c, respectively. All values are in K.

the area and an increase at the boundaries. High errors are found in regions with low data density.
The resampled fields (Fig. (5.17)) indicate a relatively low uncertainty in the spatial patterns. Positive anomalies in the Northwest and Southeast, as well as negative anomalies in the southwestern corner, are present in all realisations. The pattern of negative anomalies in Eastern Europe is shifted slightly in some realisations but is also always present. No realisation shows positive anomalies in this region.

The 8000 BP area means for all three different reconstruction settings, as explained in Tab. (5.2), can be found in Tab. (5.5). The results show nearly no differences between the different settings. Thus, the areas which are omitted in setting 8000b and 8000c have no noticeable influence on the results in the interior of the area, at least for July. The mean anomalies are slightly (not significant) positive in NW and virtually identical to the 1961-90 climate in SW. In CE, the anomalies are slightly, but not significantly, negative.
The quality control (Fig. (A.5)) shows very high costs at many sites in Southern Europe. Also one Scandinavian site shows high costs. This will be examined more closely in the following chapter.

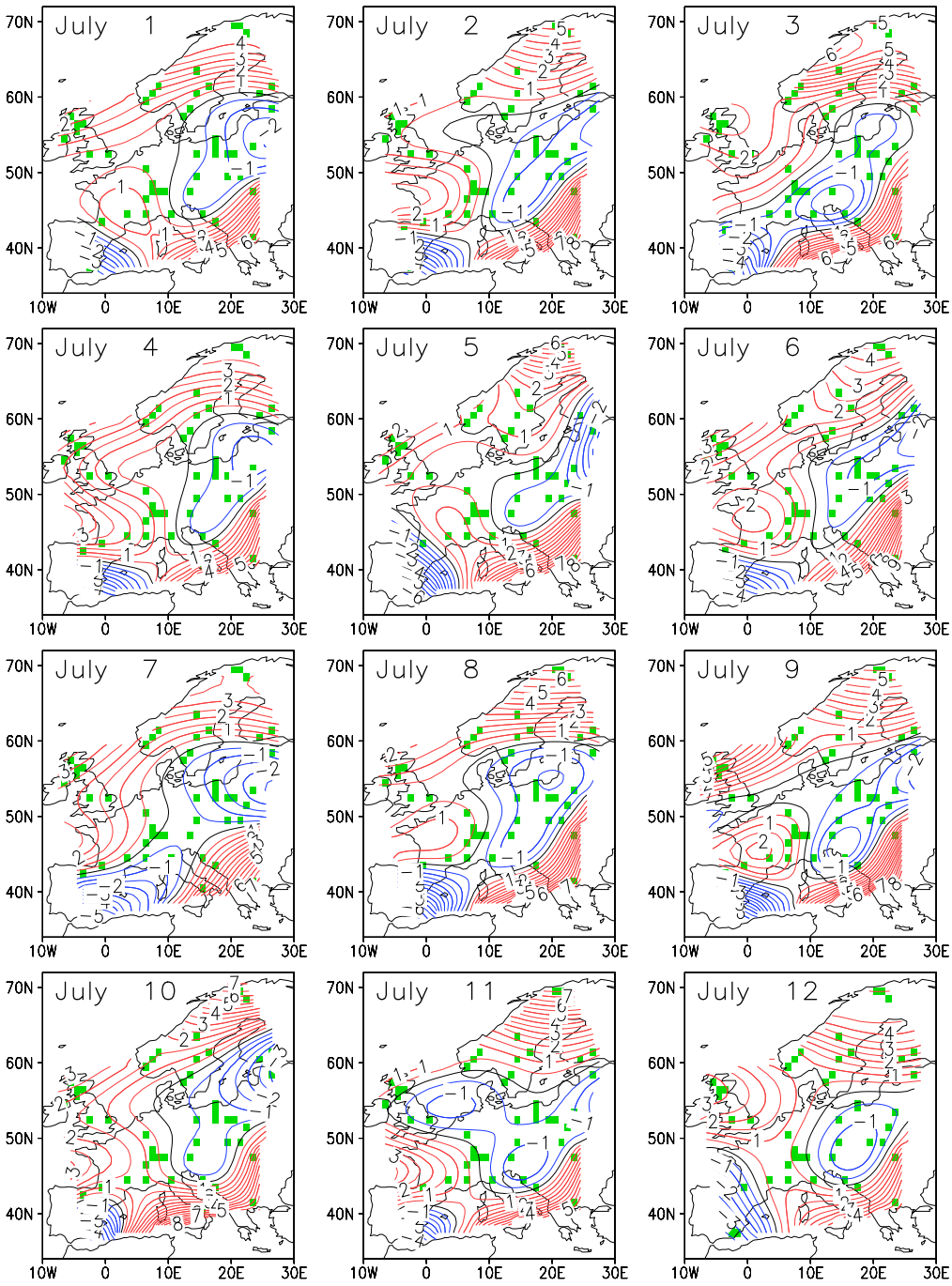

Figure 5.17: Resampling results for the 8000 BP July temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimising coefficient vector (Eq. (3.73)).

5.3.6 Results for 8000 BP - January

The January results for 8000 BP, shown here, are calculated with the 8000c setting, meaning without Northern Scandinavia and the Bulgarian site "Lake Dalgoto". This site produces large positive anomalies that are questionable. Therefore it was omitted in the 8000c setting. Data from 42 grid boxes contribute to the results here. The analysis was carried out with 35 singular modes and two modes had to be filtered.

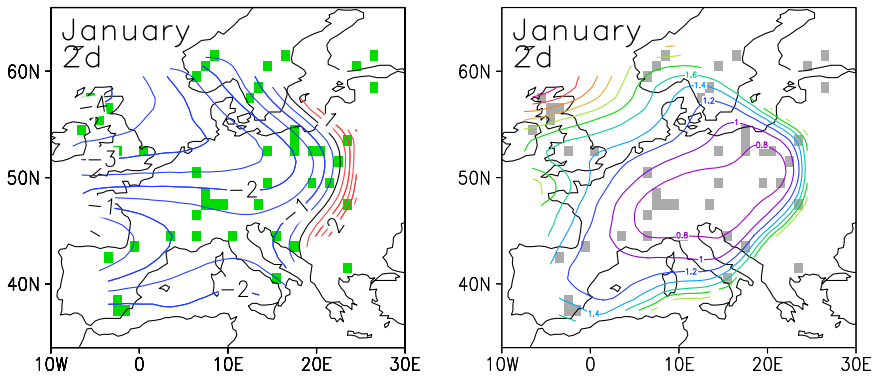

Figure 5.18: Reconstructed January temperature anomalies for 8000 BP (8000c setting) and associated field of standard deviation (right plot). Green grid boxes (left) and grey grid boxes (right) show grid boxes with vegetation data. Distance of isolines is 0.5 K for the anomalies and 0.2 K for the standard deviation. Negative and positive anomalies are indicated red and blue, respectively.

The mean anomaly field (Fig. (5.18)) shows negative temperature anomalies for most of the area, ranging from -1 K in the Mediterranean over -2 K in Central Europe to -4 K in Scotland. A gradient to positive anomalies is located at the eastern boundary.
Standard deviations are again larger than in July, but below 0.8 K where the data density is high. This time the highest errors are found in the northwestern corner of the area.
The negative anomalies over most of the area are also present in all of the resampled fields (Fig. (5.19)). However, there is some uncertainty in the placement of the minimum of the field. It is placed over Scotland in realisation 5 and 6 and a bit more south in realisation 2 and 3. A second minimum over Central Europe can be found in field 4. Also patterns without a clear minimum (like field 8) are possible.

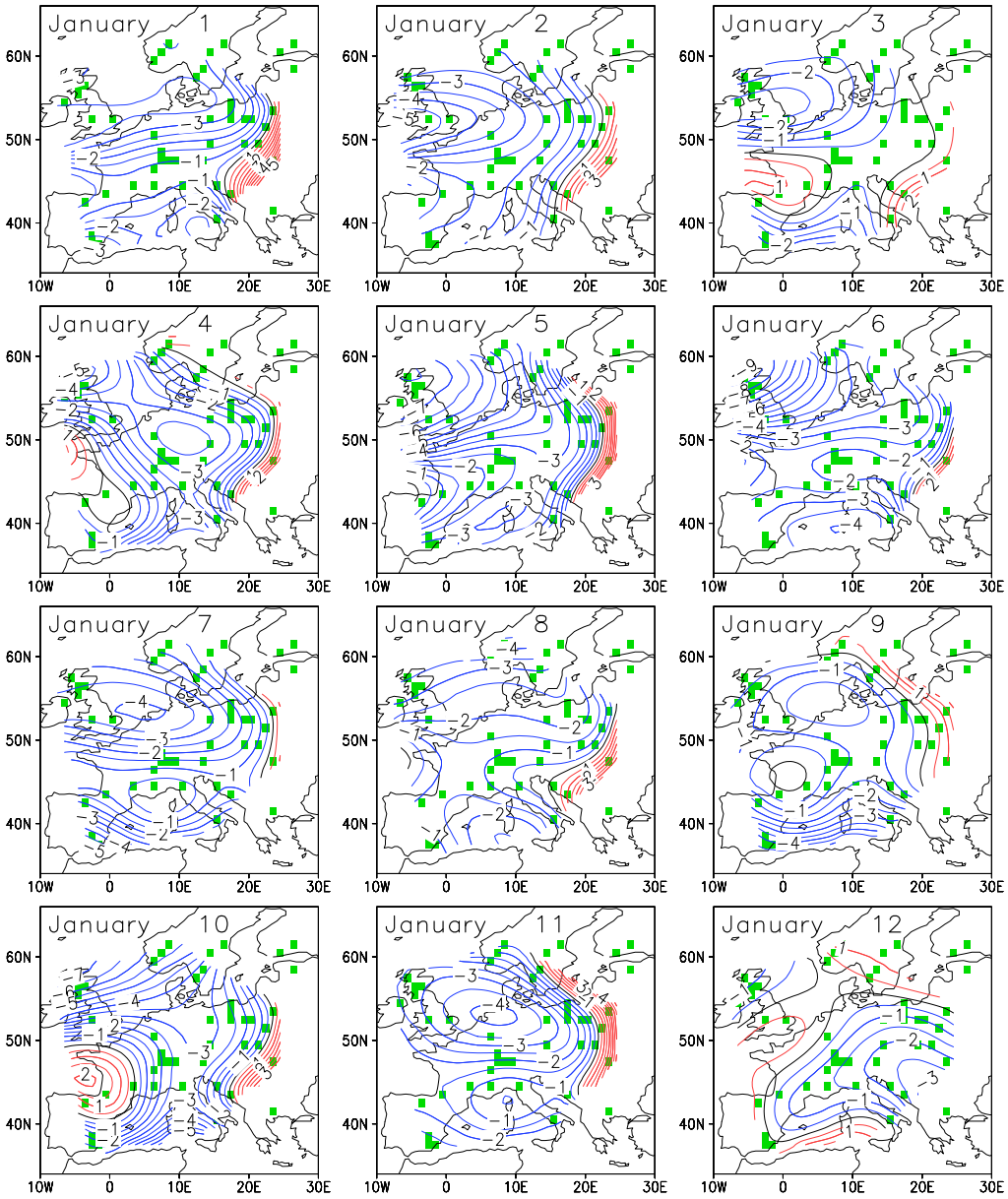

Figure 5.19: Resampling results for the 8000 BP January temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimising coefficient vector (Eq. (3.73)).

	NW	NWb	NWc	CE	CEb	CEc	SW	SWb	SWc
0.025-quantile	-5.9	-5.8	-5.8	-3.3	-3.3	-3.4	-2.4	-2.0	-2.4
mean	-3.4	-3.5	-3.0	-2.1	-2.1	-2.0	-1.1	-0.7	-1.1
0.975-quantile	-0.9	-1.2	-0.4	-1.0	-0.9	-0.8	0.1	0.7	0.2

Table 5.6: January reconstruction of 8000 BP: Mean values and quantiles for the 95% confidence interval of the three areas defined in Fig. (5.8). NW, CE and SW indicate the full area reconstruction setting, letters b and c indicate setting 8000b and 8000c, respectively. All values are in K.

The area means of NW and CE are significantly negative, compared to the modern 1961-90 climate. In SW the negative anomalies are weaker and not significant. The impact of different area settings is low. In NW the anomaly decreases slightly in the 8000c setting. The same happens in the 8000b setting in SW. However, these changes are not very strong.

5.3.7 Results for 6000 BP

The reconstructions area for the 6000 BP time slice is the same as for 8000 BP. Hence, also here the July results are shown for the complete area and the January results without Northern Scandinavia (6000c setting). As the differences to 8000 BP patterns are small, the results of 6000 BP will be presented more briefly. Data from 50 grid boxes are incorporated into the full area setting. The full area analysis was calculated with 35 singular modes. One mode had to be filtered. The analysis for the 6000c setting was calculated with 33 modes (no one filtered) and incorporates data from 40 grid boxes.

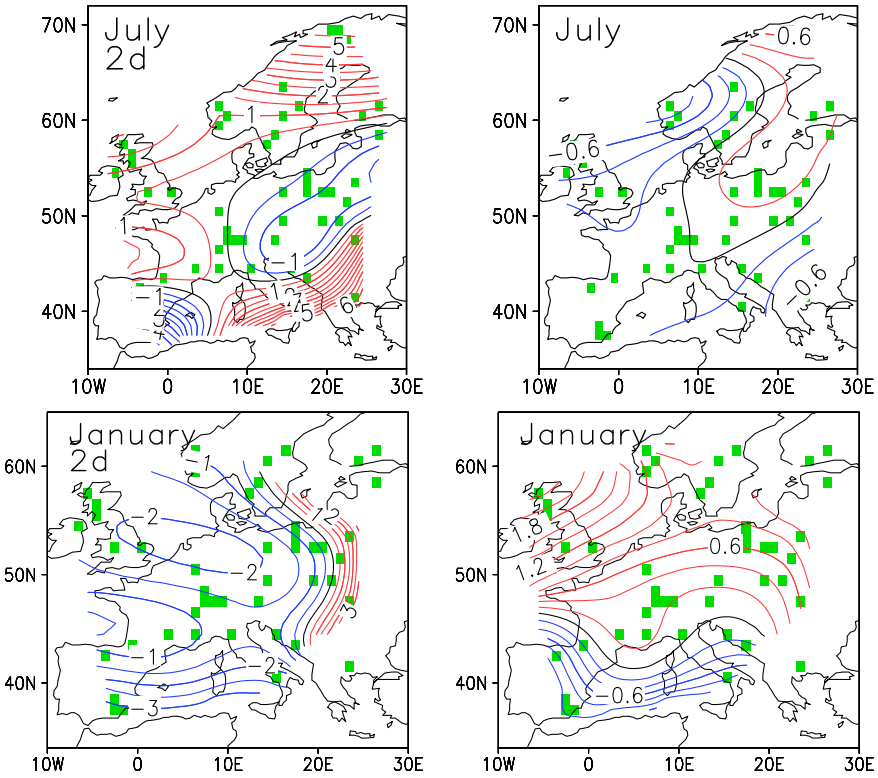

Figure 5.20: Reconstructed January and July temperature anomalies for 6000 BP (left) and difference between 6000 B.P and 8000 BP fields (right). Green grid boxes show grid boxes with vegetation data. Distance of isolines is 0.5 K for the anomalies and 0.2 K for the differences. Negative and positive anomalies are indicated red and blue, respectively.

The mean temperature anomaly fields are displayed in Fig. (5.20) together with the differences between 6000 BP and 8000 BP. The pattern of the mean fields look very similar to them of 8000 BP. The greatest differences for July

		NW	NWb	NWc	CE	CEb	CEc	SW	SWb	SWc
Jan.	0.025-qtl.	-3.8	-3.6	-3.4	-2.6	-2.6	-2.6	-2.5	-2.3	-2.2
	mean	-1.9	-1.9	-1.8	-1.6	-1.6	-1.6	-1.3	-0.9	-0.9
	0.975-qtl.	-0.3	0.0	-0.3	-0.6	-0.5	-0.5	0.0	0.4	0.3
Jul.	0.025-qtl.	-0.8	-0.7	-0.8	-1.0	-1.2	-1.0	-0.9	-0.9	-0.8
	mean	0.5	0.7	0.6	-0.2	-0.4	-0.2	0.2	0.2	0.3
	0.975-qtl.	1.7	2.1	1.9	0.6	0.5	0.6	1.1	1.4	1.3

Table 5.7: Reconstruction of 6000 BP: Mean values and quantiles for the 95% confidence interval of the three areas defined in Fig. (5.8). NW, CE and SW indicate the full area reconstruction setting, letters b and c indicate setting 6000b and 6000c, respectively. All values are in K.

are found at the northwestern boundary where the positive anomalies are up to 0.8 K lower than in 8000 BP. In the rest of the area, the differences between both time slices are around zero. The resampling results, shown in the appendix (Fig. (A.7)), show little variability. The strong positive anomalies in Northern Scandinavia and Southeastern Europe are always present. Over Britain, the Northe Sea and Central Europe anomalies could have been either slightly negative or slightly positive, with a tendency to positive anomalies in northwestern and negative in eastern direction.
For January, the differences are larger. The 6000 BP result is up to 2 K warmer in Britain, slightly warmer in Central Europe and slightly colder in Southern Europe. All realisations of the resampling (Fig. (A.6)) show slightly negative anomalies in Central Europe. The minimum is mostly located over Britain or Northern Germany and highest anomalies are always found at the eastern boundary.
The area means (Tab. (5.7)) for January are significantly negative in NW and CE, but especially in NW clearly higher than in 8000 BP. In SW, the anomalies are also negative, but at least for setting b and c not significant. The July means are slightly positive in NW and around zero in CE and SW. No value is significantly different from the 1961-90 climate.

Chapter 6

Discussion

In this chapter the results presented in the previous chapter will be discussed more in detail. The results of the different time slices will be compared. A closer examination of local reconstructions will give a better impression of the characteristics of the different time slices.
Further, the results of this work will be compared to reconstruction results obtained by other methods. Finally, it will be discussed, why the reconstruction of precipitation fields has not been possible.

6.1 Comparison of the different time slices

The mean values for the areas NW, CE and SW are a good measure for comparing the results for the different time slices quantitatively. These are displayed in Fig. (6.1) and Fig. (6.2). The different quantiles are obtained from the 1000 resampled fields, which were drawn randomly from the distribution provided by the analysis error covariance matrix. Mean values for the respective areas where calculated for each realisation. Median and quantiles then where determined non-parametrically from the resulting distributions of area mean values. The boxes contain half of the values, respectively.
For the Late Glacial, it can be stated that the differences between both time slices, 13000 BP and 12000 BP, are very small. The largest differences are found for July in the areas NW and CE, where the Younger Dryas is slightly colder.
Also the results for the Holocene time slices are nearly identical. The only exception is the January value of area NW. Here 8000 BP is reconstructed about 1 K colder than 6000 BP. However, also the uncertainty is larger in 8000 BP and thus there is nearly no difference in the 0.975-quantile. Compared to the Late Glacial, the Holocene time slices are reconstructed between

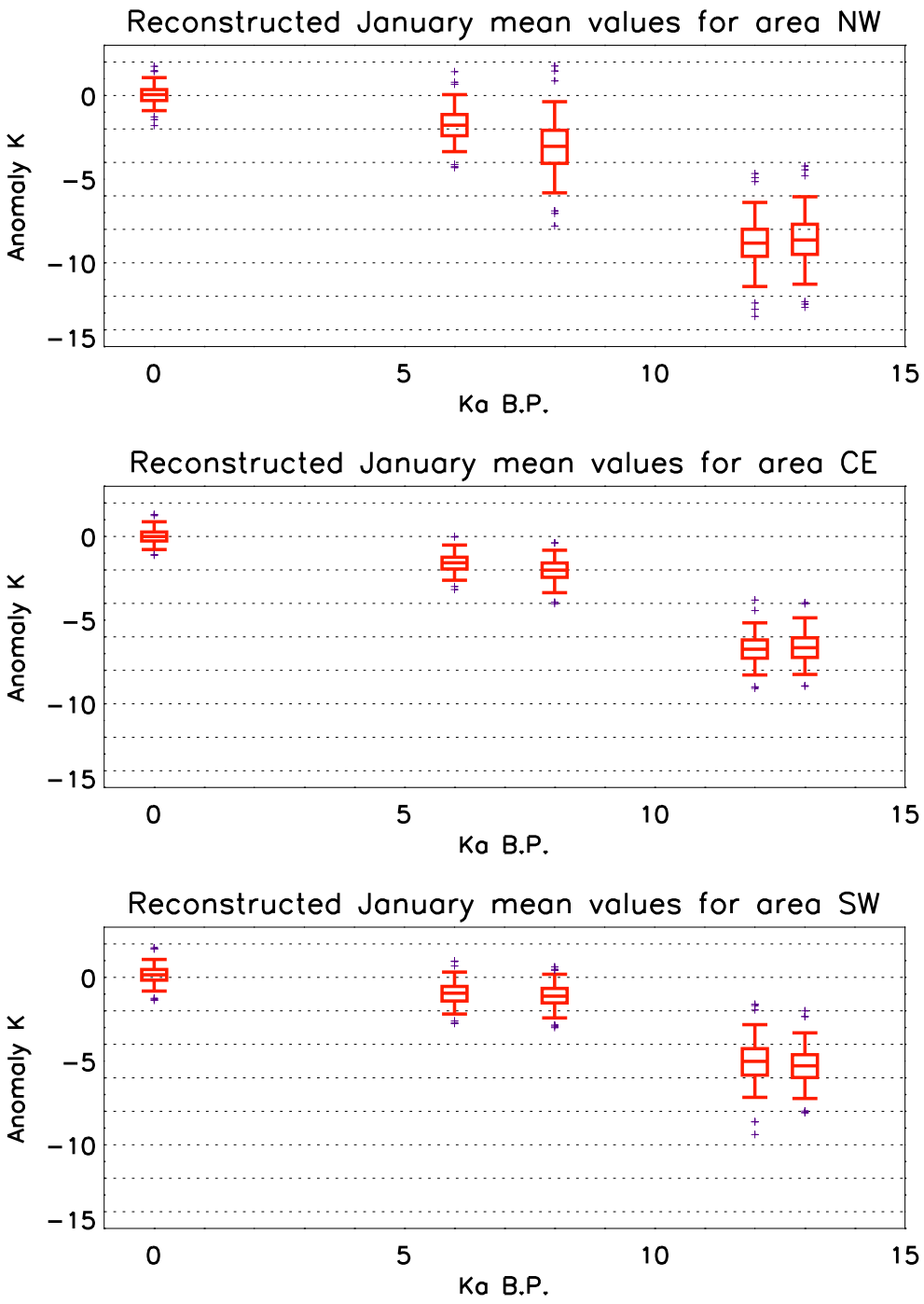

Figure 6.1: Box-Whisker plots for area means of all time slices with quantiles (q025, q25, q50, q75, q975) obtained from the realisations of the resampling procedure. The three most extreme realisations for all area means in both directions are indicated by small crosses, respectively.

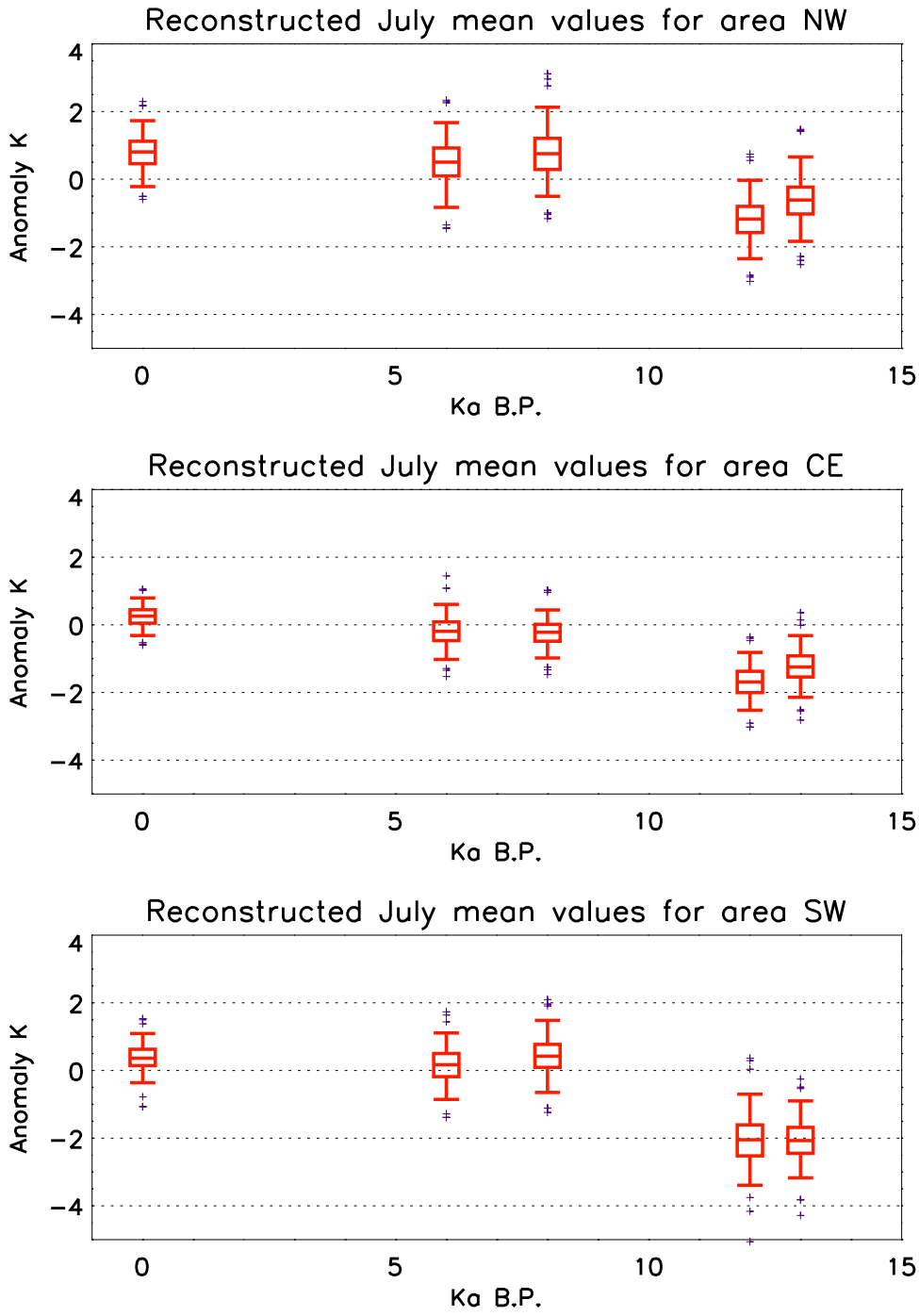

Figure 6.2: Box-Whisker plots for area means of all time slices with quantiles (q025, q25, q50, q75, q975) obtained from the realisations of the resampling procedure. The three most extreme realisations for all area means in both directions are indicated by small crosses, respectively.

4 K (SW) to 6 K (NW) warmer in January and about 2 K warmer in July. Hence it can be stated, that the difference between Late Glacial and Holocene are weaker further south and also weaker in July.
The Holocene values for January are slightly colder than the reconstructed modern values, but these differences are not significant. For July, reconstructed modern, 8000 BP and 6000 BP values are nearly the same.

6.2 Missing difference between Alleroed and Younger Dryas

Data from Greenland ice cores and other North Atlantic records (Fairbanks, 1990; Debret et al., 2007; Bakke et al., 2009) as well as reconstructions based on terrestrial data (Atkinson et al., 1987; Isarin et al., 1998) show clear evidence for a remarkable climate change on the transition between Alleroed and Younger Dryas (YD). This is not the case for the results of this work.
Most reconstructions available for the period of Alleroed and Younger Dryas are local reconstructions and thus difficult to compare to each other. Atkinson et al. (1987) reconstructed a drop of annual mean temperatures by 10 K on the transition between Alleroed and YD for Britain. The annual mean anomalies reconstructed for YD lie around -15 K. On contrary Bakke et al. (2009) reconstructed Sea Surface Temperatures (SST) near the Shetland Isles only about 3 K lower in the YD than in the Alleroed.
Another reconstruction for Luxembourg (Guiot and Couteaux (1992)) showed a temperature drop of 12K in January and 2 K in July between Alleroed and YD. The decrease in January temperatures, reconstructed by Magny et al. (2001) for Switzerland, is more moderate with around 5 K. July temperatures decreased by 2 K there. Lotter et al. (2000) only regarded Summer temperatures at the Switzerland site Gerzensee and reconstructed a decrease by 2 to 4 K.
The most extreme results for the YD were presented by Isarin et al. (1998), who reconstructed January temperature anomalies around -22 K for central Europe, below -25 K for Southern Scandinavia and Northern Britain, based on pollen abundances. They could partially reproduce these results by climate model simulations (Renssen et al. (2001)), at least for Northern Europe. There the model reproduced January temperature anomalies between -30 and -20 K. In Central Europe the anomalies lay between -5 and -15 K, with a strong zonal gradient.
Now, the question should be answered, why the difference between Alleroed and YD is not reconstructed by the approach used in this work. For this

purpose local reconstructions of the sites Abernethy Forest, Krumpa, Lukcze and Meerfelder Maar for 13000 and 12000 BP are shown in the appendix (Fig. A.8 and A.9).
At the site Krumpa the taxa *Typha latifolia* and *Lycopus europaeus*, which represent milder winters, disappear in the YD. Instead, *Betula humilis* is now occurring. This taxon has its maximum at January temperatures around -25°C. Thus, here the local reconstruction of 12000 BP is somewhat colder for January than the one of 13000 BP. The pdf, however, remains very wide and hence the result of the variational reconstruction (blue "X") lies in an area of low costs, although it is about 4 K warmer than the minimum.
At Abernethy Forest especially the occurrence of *Selaginella selaginoides* causes a reconstruction of lower July temperatures in the YD but has no effect on the January results as also *Juniperus communis* and *Carex rostrata*, which prefer a cold climate, vanish.
Taxa occurring in a colder climate, often have a very wide range of winter temperatures. They often are able to exist in both, temperate and very cold climates (see discussion in Kühl et al. (2007)). This leads to wide minima of the cost function at Krumpa and Abernethy Forest. Given these both sites, anomalies of about -15 K are possible in January, but would cause the same costs as anomalies around 0 K.
The local results of Meerfelder Maar and Lukcze show, that the vegetational signal doesn't indicate a cooling, like in Krumpa, at other sites. At Meerfelder Maar, there are some changes in the vegetation, but only *Betula nana* would indicate colder Januaries. As taxa like *Typha latifolia* or *Typha angustifolia*, which prefer milder winters, still are present, the local reconstruction results do not change very much.
At Lukcze in the YD the taxon *Najas marina* occurs. It has a small range and clearly indicates, that January temperatures had not been below -10°C. The presence of this taxon alone is responsible for the reconstruction of mild winter temperatures at this site. The result of the variational analysis, however, does not change a lot, because other sites in the neighbourhood don't show the same signal. Other proxy data, like tree rings, could be helpful for reducing such discrepancies between neighbouring sites.

6.2.1 Would a reconstruction of -20 to -30 K be possible?

The fact, that in the YD mainly taxa with a wide range of T_{Jan} are present, implies that a colder climate is expressed more by absence of mild climate indicators and less by presence of taxa preferring a cold climate. This leads

to the question, if the method would be able to reconstruct anomalies of T_{Jan} of -30 K, as they were modelled by Renssen et al. (2001).
From the taxa, occurring at Abernethy Forest (where anomalies between -15 K and -20 K where modelled), the one limiting January temperature downwards is *Myriophyllum alterniflorum*. This taxon occurs down to T_{Jan} =-20°C, meaning an anomaly of -21.4 K. Hence, the modelled anomalies can not be excluded by the vegetational evidence here.
In southwestern Norway, where Eigebakken is located, January anomalies between -20 and -30 K are the result of climate model runs by Renssen et al. (2001). Here however, the occurrence of *Populus tremula* indicates that January temperatures had not been below -18°C, what means an anomaly of -15 K.
At Lukcze in Poland *Najas marina* occurs and indicates a minimal January temperature of -10°C what would mean an anomaly of -6 K.
All these examples imply that colder January temperatures than reconstructed would have been possible theoretically from the vegetational evidence. A reconstruction of these anomalies would only be possible if one or more taxa, having their optimum in a cold climate, had occurred. Those would be e.g. *Betula humulis*, *Koenigia islandica* or *Larix*. When none of those taxa had been present or conserved in the sediment such a cold climate can not be reconstructed by this approach.
It can be concluded that the general temperature level of the Late Glacial is reproduced by the variational analysis, as the results are clearly colder than those of the Holocene. The temperature change between Alleroed and YD is below the sensitivity of this approach. It is likely, that the sensitivity of an absence/presence approach, based on purely botanical information, is too low in a cold climate. This emphasises the need for additional data sources. Modern analogue methods seem to capture a signal that is contained in the pollen abundances. It is also possible, however, that modern analogue techniques are misleaded by the non-modern analogue situation of the Late Glacial. They could possible mistake the YD pollen compositions for a more continental climate. For the future it would be preferable to include the information of pollen abundances into the probabilistic framework, provided by the Indicator Taxa model.

6.3 The Holocene results

Also for the Holocene most reconstructions that can be found in the literature are local reconstructions. A reconstruction of July temperatures for the Swedish site "Vestre" is given in Bjune et al. (2005). They reconstruct a

temperature level comparable to the 1961-90 climate with similar values for 8000 BP and 6000 BP, what is in good agreement with the results of this work. The July reconstruction for "Tsuolbmajavri" (Korhola et al., 2002) also shows similar values for 0, 6000 and 8000 BP.
Davis et al. (2003) analysed data from 510 fossil sites with a modern analogue approach and obtained temperature fields by a 4D-interpolation with smoothing splines. It should be noted, that this procedure, unlike the approach presented in this work, has not the probabilistic interpretation of conditional sampling.
They also show results that are in good agreement to the results of this work. Especially for northwestern Europe they also reconstruct slightly lower January temperatures for 8000 BP as for 6000 BP with 6000 BP anomalies of -1 K.
The reconstruction of 6000 BP January temperatures by Cheddadi et al. (1997) show a similar pattern with negative anomalies in southwestern Europe and positive anomalies in northeastern Europe.
Wu et al. (2007) show reconstructions for numerous sites in Europe for 6000 BP. Their January results show no clear signal but they reconstruct large July anomalies between 3 and 5 K, especially in Southern Scandinavia and in the Alps. This result is significantly different to the results of this work. Local reconstructions for the Scandinavian sites "Holtjaernen" and "Samboesjoen" and for the Alpine sites "Nussbaumerseen" and "Lobsigensee", presented in Fig. (A.10), show how such anomalies of T_{Jul} would fit into the local reconstructed pdfs.

Scandinavia: At the Scandinavian sites the local pdfs are relatively wide. Hence positive anomalies of 3-5 K would be possible at local costs with value around 5 (actual local costs are 0.52 and 0.07 for Holtjaernen and Samboesjoen, respectively). From the vegetational evidence the occurrence *Picea abies* and *Calluna vulgaris* provide the lower limit for positive anomalies of T_{Jul} with 5 K at Holtjaernen and 3.5 K at Samboesjoen. Thus positive anomalies within that range are not fully excludable for Southern Scandinavia but relatively unlikely based on the local reconstructions.

Alps: The local reconstructions for the Alpine sites result in pdfs with smaller widths. At Nussbaumerseen, the local reconstructed July temperature fits well with the result of the variational analysis. An anomaly of 5 K would produce costs between 15 and 20 here (actual costs 0.65). At

Lobsigensee the local reconstruction is even colder than the result of the variational analysis, mainly due to the occurrence of *Picea abies*. Therefore this site already produces costs of 1.51. An T_{Jul} anomaly of more than 3 K, however, would produce cost of more than 15.
The occurrence of *Picea abies* at both sites would allow for a maximum positive anomaly of 3.5 K. Hence these values can not be fully excluded but are extremely unlikely based on the local reconstructions.

6.3.1 High costs in Southern Europe

The quality control (Fig. (A.5)) for the 8000 BP (as well as 6000 BP) analysis shows high local cost for many Southern European sites. These could indicate errors in the dataset or problems with certain locations. The local reconstructions of the Spanish sites "San Rafael" and "Antas" and of the Pyrenee sites "Las Pardillas" and "Biscaye" are displayed in Fig. (A.11).
The locally reconstructed values at San Rafael are much colder than the results of the variational analysis, leading to high costs here. At the neighbouring grid box, where Antas is located, the local reconstruction with exactly the same taxa fits well with the variational analysis results. The mean altitude of the Antas grid box is 167 m, while the one San Rafael is 690 m. Both sites lie on sea level and both grid boxes have similar modern climate values. Thus the altitude correction for San Rafael (690 m to 0 m) leads to lower reconstructed values there. Possibly the microclimatic conditions are not captured by the large scale climate values here. It is also possible, that the taxa composition here is not really representative for sea level but for higher altitudes as pollen might have come from distant regions. This problem often arises with pollen data in mountainous regions. Here, however, the variational analysis is able to treat the result of San Rafael as outlier.
At Las Pardillas the effect of altitude in the other direction is visible. The values of the local reconstruction are much higher than the result of the variational analysis. This site lies on an altitude of 1850 m, while the mean altitude of its grid box is only 880 m. It is likely that the present taxa are not really representative for the altitude of the site but mainly come from deeper regions. The result of Las Pardillas differs considerably from the results at the near site Biscaye, where slightly negative anomalies are reconstructed. Would the variational analysis result in higher values in this region, the costs would increase more strongly at Biscaye than they would decrease at Las Pardillas. A strong gradient between these two sites would cause too much model costs.
However, the analysis is not always able to treat problematic sites as outlier. This becomes clear, when looking at the local reconstructions of Lake

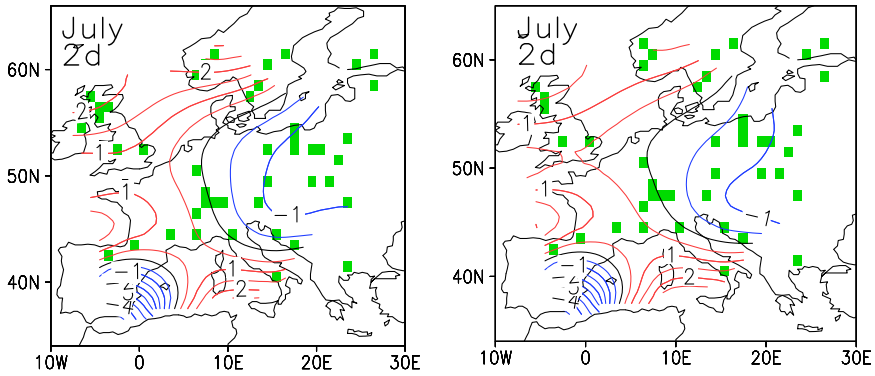

Figure 6.3: Reconstructed July anomalies for the the 8000c (left) and 6000c (right) setting without Lake Dalgoto. Green grid boxes show grid boxes with vegetation data. Distance of isolines is 0.5 K. Negative and positive anomalies are indicated red and blue, respectively.

Vrana, Malo Jezero and Lake Dalgato, presented in Fig. (A.12). Also the reconstructed values for Lake Dalgoto are very high, due to the high altitude of the site. The two near sites Lake Vrana and Malo Jezero show slightly negative anomalies. However, these are double as far away from Lake Dalgoto as Las Pardillas is from Biscaye. This distance seems to be too much for the analysis to be able to compensate this effect. Instead, a strong gradient to positive anomalies between Malo Jezero and Las Pardillas is the result of the variational analysis. The reconstruction result for July temperatures for 8000 BP and 6000 BP, with the site Lake Dalgoto omitted, is shown in Fig. (6.3). It becomes clear, that the gradient to positive anomalies in the southeast was only based on this site. When Lake Dalgoto is omitted, this pattern is completely gone and slightly negative anomalies are present everywhere in the eastern part of the analysis area.

6.4 The problem of reconstructing precipitation

The variational reconstruction for precipitation with a physical constraint based on the omega equation (see Section (3.1.4)) was tested for the 13000 BP time slice. The analysis was carried out with 29 singular modes and no forcing term in the physical constraint. The forcing term, based on the orography, would only result in an anomalous forcing when an assumption of an anomalous wind field would be made. For the first trial it seemed to

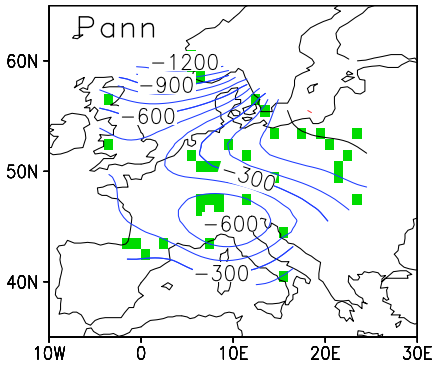

Figure 6.4: Reconstructed anomalies of annual precipitation for 13000 BP. Green grid boxes show grid boxes with vegetation data. Distance of isolines is 100 mm. Negative and positive anomalies are indicated red and blue, respectively.

be sufficient to make the analysis without anomalous forcing.
The resulting field of annual precipitation anomalies is presented in Fig. (6.4). Negative anomalies are found over most of the reconstruction area. The lowest values are present at the western coast of Norway with around -1200 mm. Another minimum of the field is found in the Alps with values around -500 mm. At the northeastern boundary only, the anomalies have values around zero. An experiment with modern vegetation leads to a very similar result. The method is not able to reconstruct the modern precipitation values for most of the area. For understanding that, it is necessary to look at the local reconstruction results. Two local reconstructions for the sites "Utsira" and "Meerfelder Maar" are shown in Fig. (6.5).
In Utsira there are some differences in the maxima of the respective taxa, which range from about 600 mm to about 1400 mm. The reconstruction result, however, is constrained very much by *Koenigia islandica* to a value of 850 mm. The modern climatological value of 2110 mm is far away from the range of all taxa, occurring here. This is a bit different for Meerfelder Maar, where the modern climatological value of 945 mm lies within the range of the different taxa but always above their maximum. Here all occurring taxa have very similar transfer functions. The problem is well known from local precipitation reconstructions at the same site (Litt et al., 2009). Throughout the whole Holocene the reconstructed values are always clear below the modern values. As the signal is weakly different in the Late Glacial and shows some variability in the time series the local reconstruction of changes in annual precipitation has a certain potential.

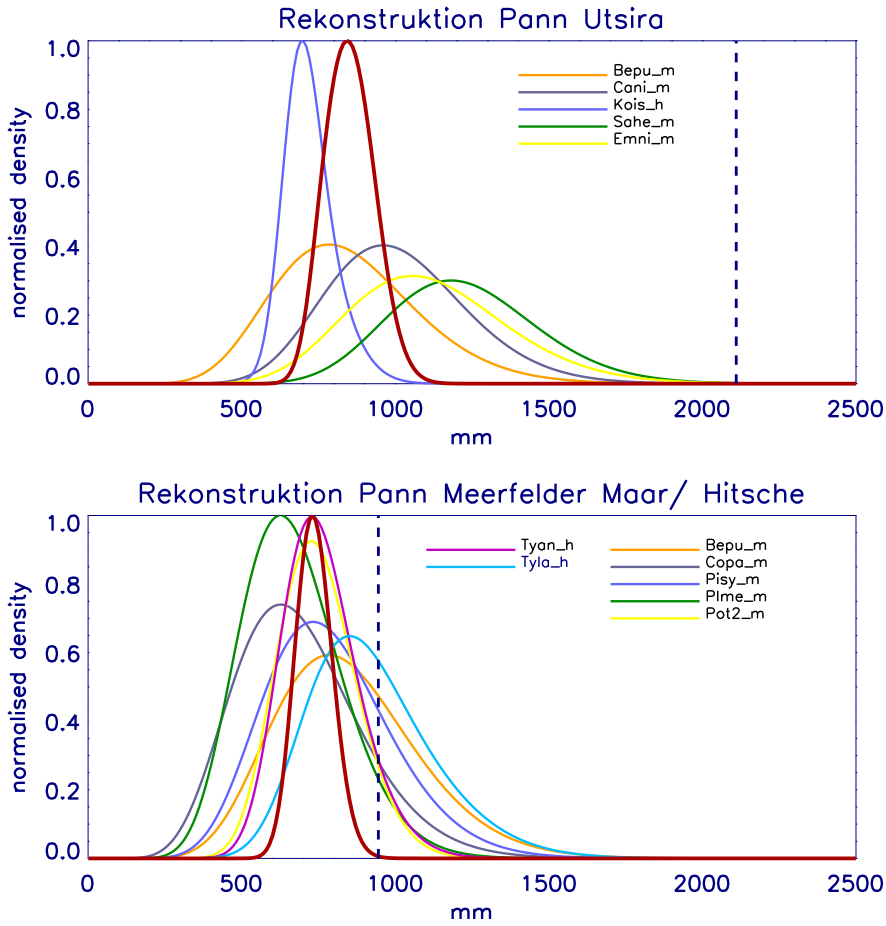

Figure 6.5: Local reconstruction of annual precipitation for the sites Utsira and Meerfelder Maar/Hitsche for 13000 BP. The pdfs of the different taxa are normalised by their overall maximum. These are marginal distributions of the 3D-pdfs. The resulting pdf (dark red) is normalised by its maximum. The modern (1961-90) climate value is indicated by a blue dashed line, respectively.

The attempt to reconstruct physically consistent fields of annual precipitation seems not to be very promising. The locally reconstructed values of annual precipitation always lay between 600 mm and 900 mm throughout the whole analysis area, despite the large variability in mountainous regions in the modern values. Additionally the values are heavily underestimated at most of the sites. The local anomalies are quite well captured by the variational analysis as it results in values around -1200 mm in Norway and around -300 mm in Western Germany, where Meerfelder Maar is located. However this field most likely only shows the level of underestimation in the certain regions. A possible weak climatological signal would not be noticeable in such a reconstruction.
These results lead to the conclusion, that the occurrence of vegetational taxa is not a good proxy for reconstructing annual precipitation. The fact that nearly all taxa have their maximum values below the modern climatological values implies that precipitation in Europe is always more than necessary for the plants to survive. Thus, variations in precipitation (despite a severe drought) will never cause the disappearance (or appearance) of a taxon here. This surely is different in arid regions, where precipitation really is a limiting factor. It is defenitely recommendable to try again the reconstruction of physically consistent fields of annual precipitation in such a region. Finally, finding and using other types of proxy data for precipitation should have a high priority in the near future.

Chapter 7

Concluding remarks

In this final chapter it will be discussed which conclusions can be drawn from this work. For this purpose, first the results will be summarised. Then it will be assessed, which ideas and needs for future research have been arisen from the findings of this work.

7.1 Summary of important results

Mixture models

In the first phase of this work, mixture models have been implemented into the three dimensional copula approach for transfer functions from Schölzel (2005). This was possible without problems from the theoretical point of view. To address the well known problem of overfitting, that can appear when mixture models are used, a different criterion for determining M_{opt} was applied with the BIC. Additionally the em-algorithm was initialised by k-means clustering. These arrangements reduced the cases with overfitting. However, they were not able to avoid overfitting completely from the outset. Hence, it is still necessary to rework the transfer functions of certain taxa by giving a limit to the lowest eigenvalues of the mixture model components.
When the two dimensional T_{Jan}, T_{Jul} mixture models (pdf2d) are compared to the respective T_{Jan}, T_{Jul} marginal distribution of the three dimensional mixture models (pdf3d), considerable differences could be noticed for many taxa. In the case of pdf3d the maximum of the pdf often is shifted to higher January temperatures, compared to pdf2d. The effect of this shifting can be very large, especially when the respective taxon has a wide January temperature range (Fig. 5.3). This behaviour is related to weighting, which is applied to the climatological data as an alternative for dividing by the climate marginal distribution (Eq. 2.34). High values of annual precipitation

get a high weight because they are massively underrepresented in the climatological observations. As it is a three dimensional approach these weights are automatically addressed also to the associated values of January and July temperatures. Because of the fact that high values of annual precipitation are only observed together with relatively high values of January temperature, the maxima of the pdfs are systematically shifted to the warm end of the range by this effect. When multiple taxa with this feature are combined in a reconstruction, this could be the cause for a positive bias in T_{Jan}.

Reconstruction results

Initially, an experiment for reconstructing the modern 1961-90 climate was performed. Only taxa that also occurred in 8000 BP (anywhere in the area) were used at the same grid boxes, where paleodata were available. The results indicated that reconstructing the modern climate works well, except for Northern Scandinavia in January and Britain in July, where large positive anomalies were reconstructed, respectively. These positive anomalies are reduced in the case of pdf2d, what implies that these transfer functions are the better choice.
The results of the variational analysis show nearly no difference between the both Late Glacial time slices. This is in disagreement to other reconstruction approaches as discussed in Section (6.1). Apparently, our proxy data are not sensitive enough for reproducing the differences between Alleroed and YD which are evident from pollen diagrams. The differences, however, are more pronounced in the pollen abundances than in appearance or disappearance of taxa. Additionally, the differences in occurrence of certain taxa between both time slices show more absence of indicators for mild climate in the YD than presence of indicators for very cold climate.
Modern analogue approaches capture the signal in the abundances and thus are able to reconstruct a difference between the two Late Glacial time slices. Eventually one could argue that modern analogue methods mistake a shortening of the growing season (what could cause a similar signal in the abundances) for a drop of winter temperatures. However, the difference between Alleroed and YD was also reconstructed by Atkinson et al. (1987) who used an indicator taxa approach with beetle species, which show a certain sensitivity to winter temperatures. This is not the case for most of the taxa, occurring in the YD, because they can occur in a wide range of January temperatures. Additionally the climate of the Late Glacial often lies in the centre of this range. Therefore it can be stated, that the low sensitivity of the method here is caused by insensitive proxies.
Despite this lack of sensitivity the variational analysis has proven its abil-

ity to reconstruct temperature fields for a cold climate phase, which show anomalies below -13 K for northwestern Europe in the most extreme realisations.

The results for the Holocene time slices are in much better agreement to the results of other reconstruction approaches. Major differences are found only in comparison to the July temperature results of Wu et al. (2007), who reconstructed high anomalies of T_{Jul} for a lot of European sites. However, these results can be judged as unlikely, given the local reconstruction results with the indicator taxa method and the vegetation evidence (see discussion in Section 6.3).
For the calculated area means it can be stated that the temperatures at 8000 and 6000 BP have not been significantly different from the modern values, although reconstruction of January temperature results in slightly lower values for both time slices. Especially the mean value for area NW shows a continuous increase, as 6000 BP is warmer than 8000 BP and 0 BP is warmer than 6000 BP. These differences are not significant, due to larger errors in 8000 BP.

Displaying the uncertainty

The resampling from the analysis error covariance matrix, presented in Section (3.4.1), turned out to be a valuable tool for displaying the spatial uncertainty of reconstructed fields. As already emphasised, the resulting field of the variational analysis represents the expectation value (mean field) of the reconstructed climate. The resampled fields all are possible realisations of the reconstructed climate. Regarding only the mean field could lead to the wrong impression of relatively smooth patterns. The different realisations, obtained by resampling, often show more heterogeneous patterns and also often considerable differences in the patterns among each other. Therefore, regarding different realisations gives a more realistic impression of the uncertainty and indicates which patterns (like maxima, minima or gradients) are robust and which are uncertain.
It should be pointed out, that regarding the analysis result as conditional expectation of a probability distribution is a new point of view, different to the common one. In most palaeoclimate reconstruction studies the problem is reduced to finding the "one and only" solution. The results of the conditional sampling from the analysis error covarinace matrix make clear, however, that a variety of solutions is possible. This should be always kept in mind, as well as the fact that climate reconstruction is a task that should be addressed by a probabilistic approach.

Annual precipitation

The attempt to reconstruct annual precipitation clarified some problems, which are more related to the proxy data that to the reconstruction approach. Investigation of local reconstruction results showed that the reconstruction underestimates annual precipitation at most of the localities. This problem is worse, the higher the true values of annual precipitation are and is related to the transfer functions, which all have their maxima at relatively low values (see discussion in Section 6.4).
It became clear that any climatological signal wound vanish in the much larger signal of underestimation. This signal, however, is well reproduced by the variational analysis. This indicates the general capability of the variational analysis, to reconstruct even extreme anomaly fields of annual precipitation under the constraint of the omega equation. Hence, it can be concluded, that there is a good chance for the method to produce realistic results, when it is applied to proxy data which is more sensitive to annual precipitation.

7.2 Suggestions for future research

Concerning the consequences for future research, some things that were already mentioned by Gebhardt (2003) remained untouched in this work and some new came to mind. All these points will be mentioned in the following and it is left to my successors, which relevance they will assign them.

Transfer functions

The differences between pdf2d and pdf3d transfer functions imply that it might be necessary to regard temperatures and precipitation separately in future approaches. Before doing that it can be tested, if a modification of the weighting procedure could solve this problem. It would be a possibility, for instance, to exclude very high values of annual precipitation from the weighting.
Regarding the problem of overfitting, still no "perfect" approach has been found, which avoids overfitting from the outset. The earlier suggestion of Gebhardt (2003), to avoid overfitting by "ecologically consistent interpolation of pdf-values" in some cases, has not been addressed, because it is yet unclear how this should be done exactly. When overfitting is not avoidable in any mixture model fit, regardless of number of components or smoothing criterion, it also seems to be practical to go back to the normal distribution. This possibility was chosen by the BIC for a lot of taxa, anyway.

Sensitivity

Several aspects of sensitivity of the method are unclear at the moment. One thing that should be checked is the sensitivity to uncertainties in the paleodata. In the current approach the paleodata, which is found in the sediment is considered to represent the "true" vegetation of the past. It is absolutely sure, however, that not all taxa that have been present in the past are represented in the sediment. The effect of this source of uncertainty to the reconstruction result should be quantified. A possible experiment for addressing this point would be an analysis with modern vegetation, with randomly leaving out a certain number of taxa at every site. This type of cross-validation approach, by repeating this experiment sufficiently often, could give an estimate of the effect that not all present taxa are found in the sediment.
Further, reconstruction approaches for more time slices are necessary to evaluate the general capabilities of the method to resolve differences between similar climate states. It was a result of this work, that the method is not able to resolve the differences between Alleroed and YD. Now it would be interesting, if differences between the Late Glacial and the Last Glacial Maximum (LGM) can be resolved. For the Holocene, two interesting time slices to compare to each other would be 1000 BP (Medieval Climate Anomaly) and 300 BP (Little Ice Age).
To address the problem of low sensitivity to changes in the January temperature in cold climate states, a new idea came up within our collaboration recently. This touches the question how the information, which is contained in the pollen abundances, could be incorporated into the Bayesian indicator taxa model. The different mixture models components generally represent different clusters in the data. If certain abundances could be related to certain clusters, it would be possible only to use one cluster for the reconstruction. This could reduce the January range and thus enhance the sensitivity considerably. However, it has not been investigated up to now, how promising this idea is.
Finally, it would be desirable to fully use the potential of the Bayesian approach by incorporation of prior probability distributions. For instance it would be possible to use GCM-output from the PMIP-II project as a prior for the 6000 BP time slice. It would be interesting to see, if and how this would change the reconstruction result and if this would influence the sensitivity of the method. When the idea of selecting one component of the mixture model is followed, it could be a possibility to use reconstruction result of the full mixture models as prior pdf.

Other proxy data

It was already recommended by Gebhardt (2003) to incorporate other proxy data into the variational analysis. This could be done relatively easy for every proxy, which has a quantitative relation to temperature or precipitation. Especially the incorporation of marine proxy data would be extremely helpful as the method currently is limited to continental areas. Reconstruction results for oceanic areas would surely help to clarify the role of the oceans in climate variability. Another promising proxy would be the winter temperature sensitive beetle species from Atkinson et al. (1987). For the Holocene, also the incorporation of data from tree rings could be promising.
Not really other proxy data but other input data would be the local reconstructions of Wu et al. (2007). Their dataset would allow for an application of the variational analysis in a much larger area, covering Europe, Asia and parts of Africa. The output of inverse vegetation modelling could be implemented into the variational analysis without problems. A reconstruction of physically consistent fields with that size would be a real breakthrough for paleoclimate research.

Further open issues

It was part of the motivation for reconstructing large scale physically consistent fields, to possess reconstruction results, which can be quantitatively compared to climate model output. For the 6000 BP time slice this would be possible very quickly, as this time slice was addressed by the PMIP-II project and the resulting data is available to our collaboration.
Two things, which were proposed by Gebhardt (2003) have not been worked on during this thesis. The first was to reduce the positive bias of January temperatures in Northern Scandinavia by introducing a north-south gradient as weak constraint into the variational analysis. On the other hand the idea of selecting only one mixture model component for reconstruction could also reduce this problem.
Finally the model error covariance matrix in the variational analysis has still not been estimated till now and remained set to unity. For a full representation of the analysis error it would be necessary to have an estimate for the model error covariance matrix. It should be referred to the proposals of Gebhardt (2003) to address this problem.

Appendix A

Additional figures

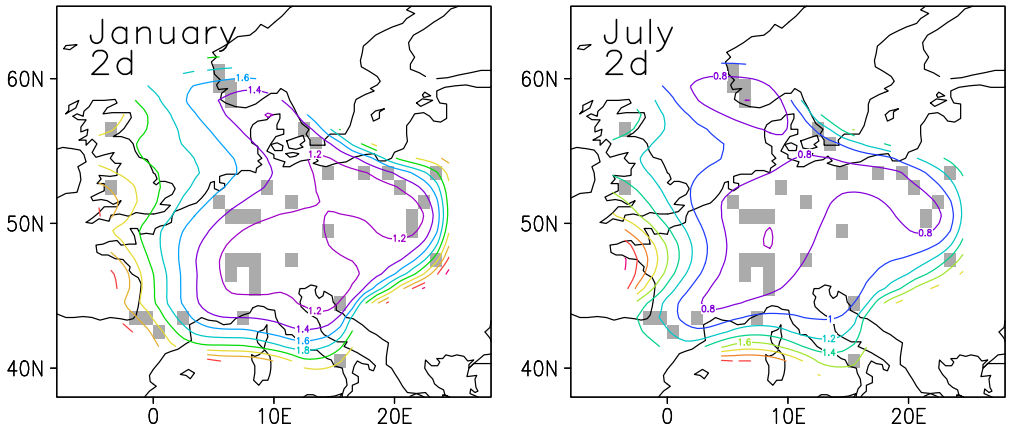

Figure A.1: Fields of standard deviations from the analysis with pdf2d for 13000 BP. Grey grid boxes indicate grid boxes with vegetation data.

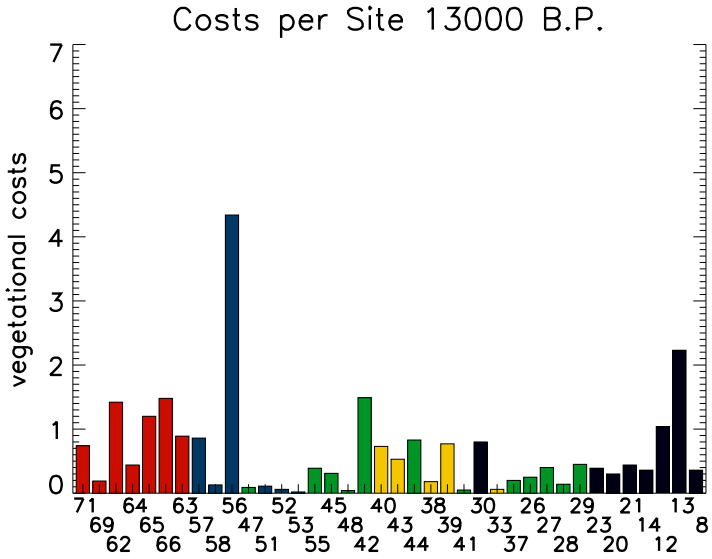

Figure A.2: Vegetational costs at the different sites of the 13000 B.P. reconstruction. Site numbers are written below the bars. Black = Scandinavia and Britain, green = Eastern Europe, yellow = Central Europe, blue = Alps, red = Southern Europe.

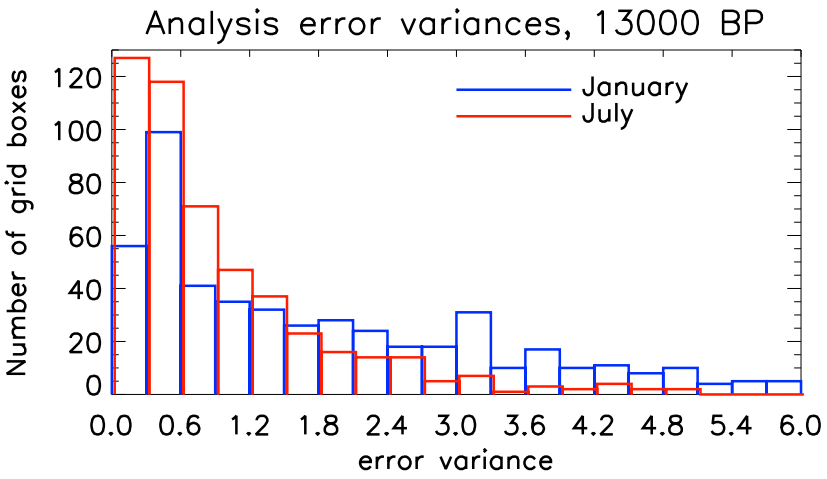

Figure A.3: Histograms of analysis error variances for the 13000 BP mean fields, consisting of 20 bins with binwidth 0.3. Bins for July error variances are shifted slightly for better visibility.

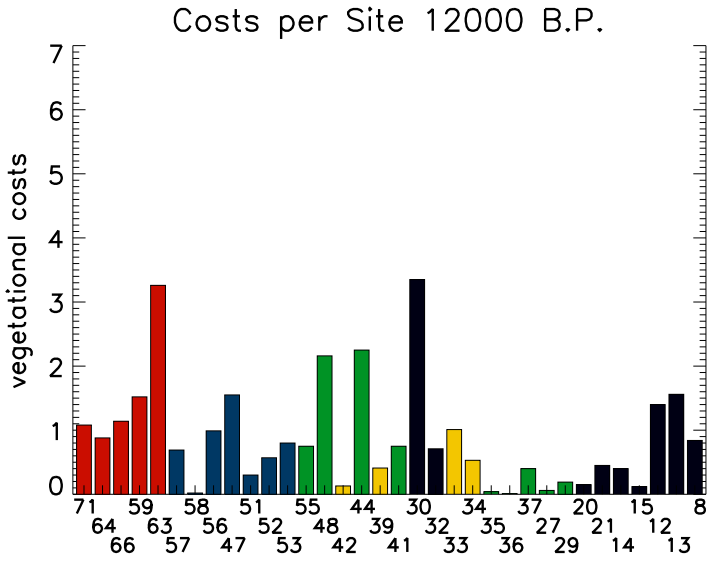

Figure A.4: Vegetational costs at the different sites of the 12000 B.P. reconstruction. Site numbers are written below the bars. Black = Scandinavia and Britain, green = Eastern Europe, yellow = Central Europe, blue = Alps, red = Southern Europe.

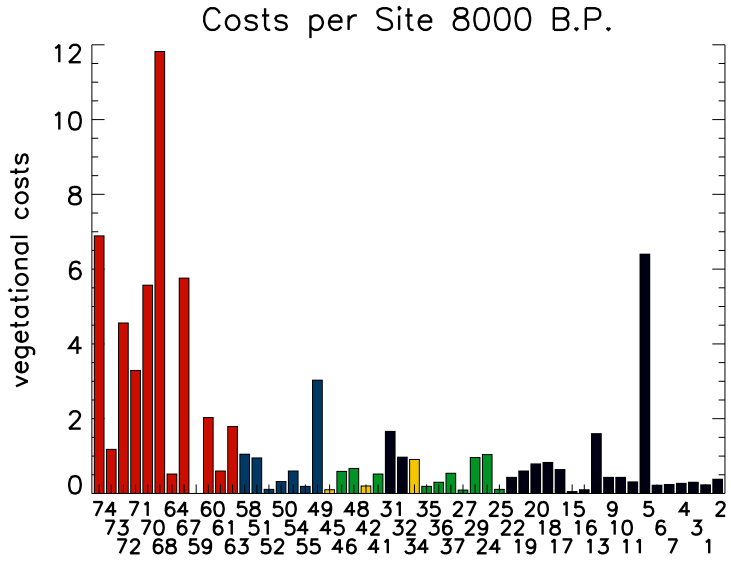

Figure A.5: Vegetational costs at the different sites of the 8000 B.P. reconstruction. Site numbers are written below the bars. Black = Scandinavia and Britain, green = Eastern Europe, yellow = Central Europe, blue = Alps, red = Southern Europe.

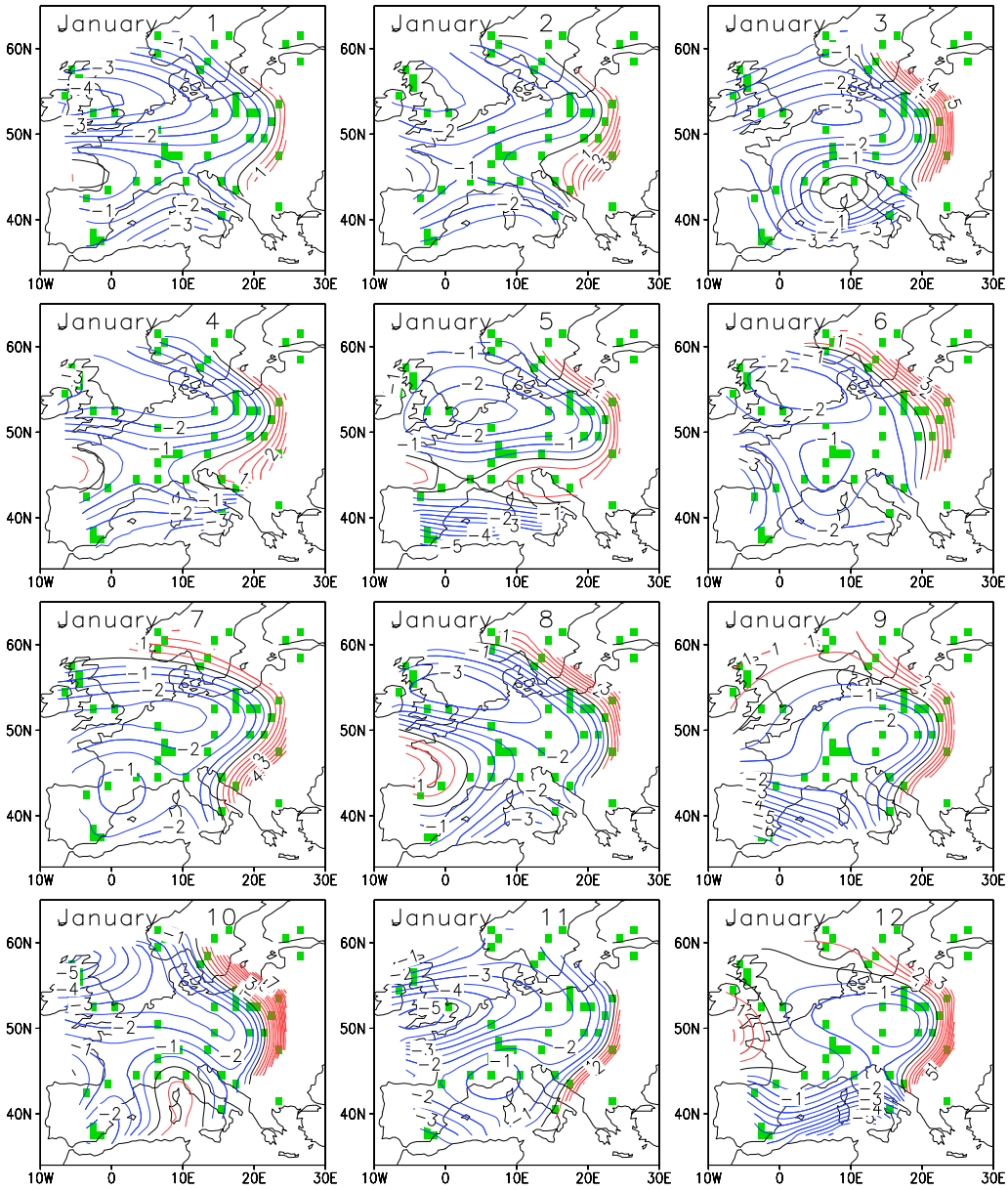

Figure A.6: Resampling results for the 6000 B.P. January temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimizing coefficient vector (Eq. (3.73)).

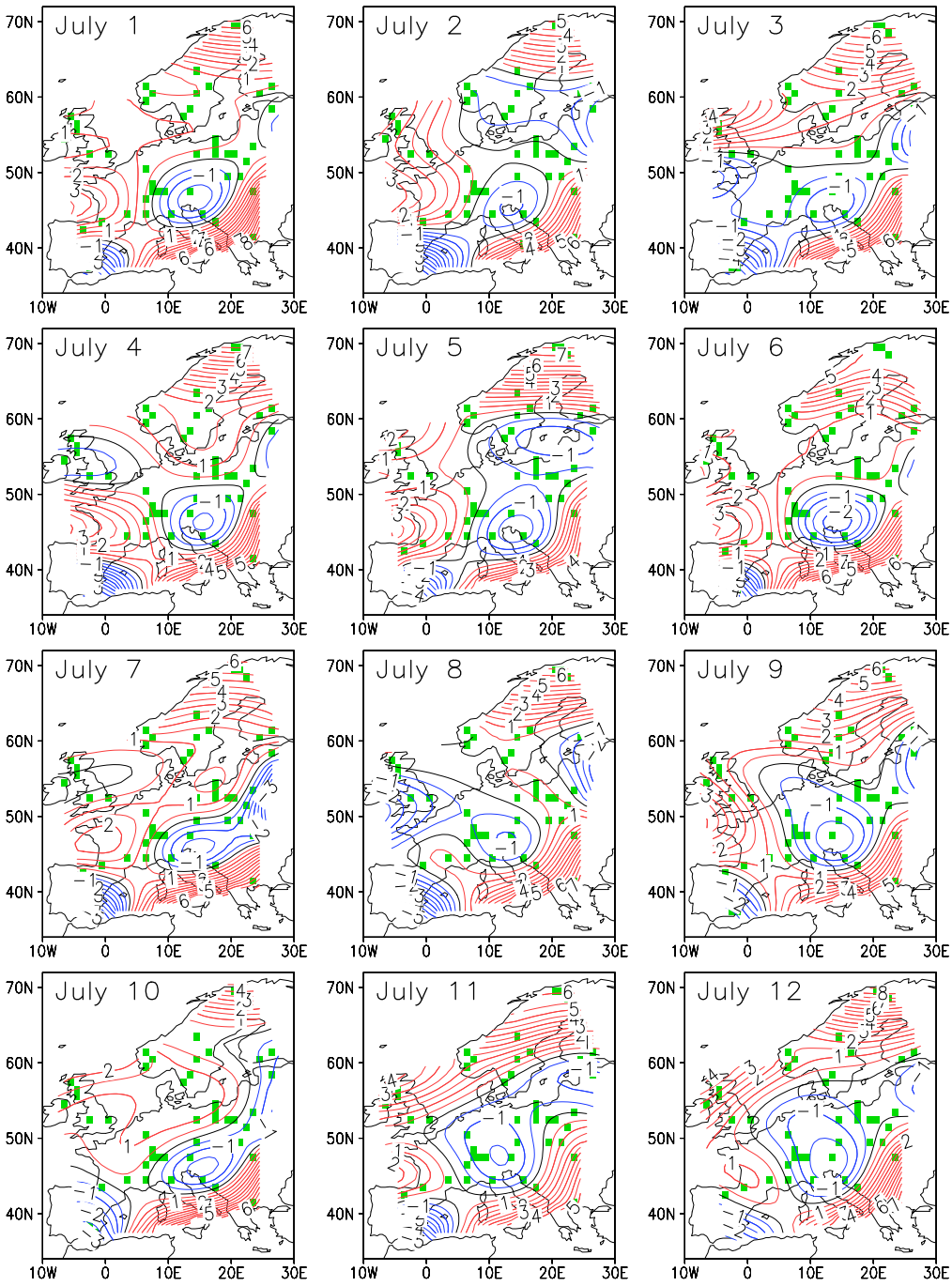

Figure A.7: Resampling results for the 6000 B.P. July temperature anomalies. Isoline distance is 0.5 K, negative anomalies are indicated blue, positive anomalies are indicated red. Fields are sorted from top to bottom, according to their Mahalanobis Distance to the minimizing coefficient vector (Eq. (3.73)).

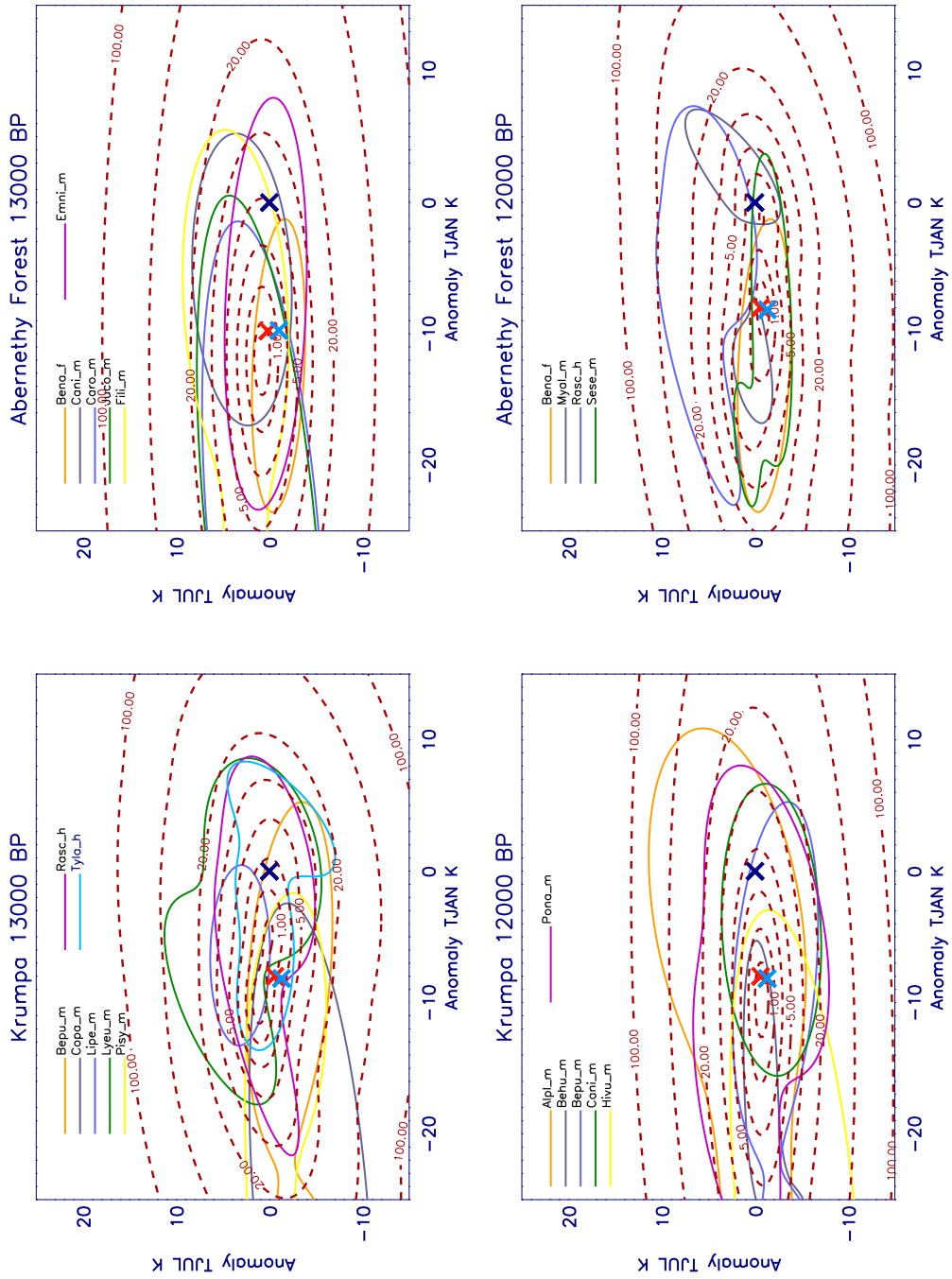

Figure A.8: Local reconstruction for the sites Abernethy Forest and Krumpa for the Alleroed (left) and YD (right). Pdfs of occurring taxa are represented by their isolines of 33% of the maximum density. The resulting pdf (red dashed lines) is displayed in terms of vegetational costs (ln of the pdf values normalised by the maximum of the pdf). Modern climate is marked by a dark X. The reconstruction results of the variational analysis are marked by a red and blue X for Alleroed and Younger Dryas, respectively.

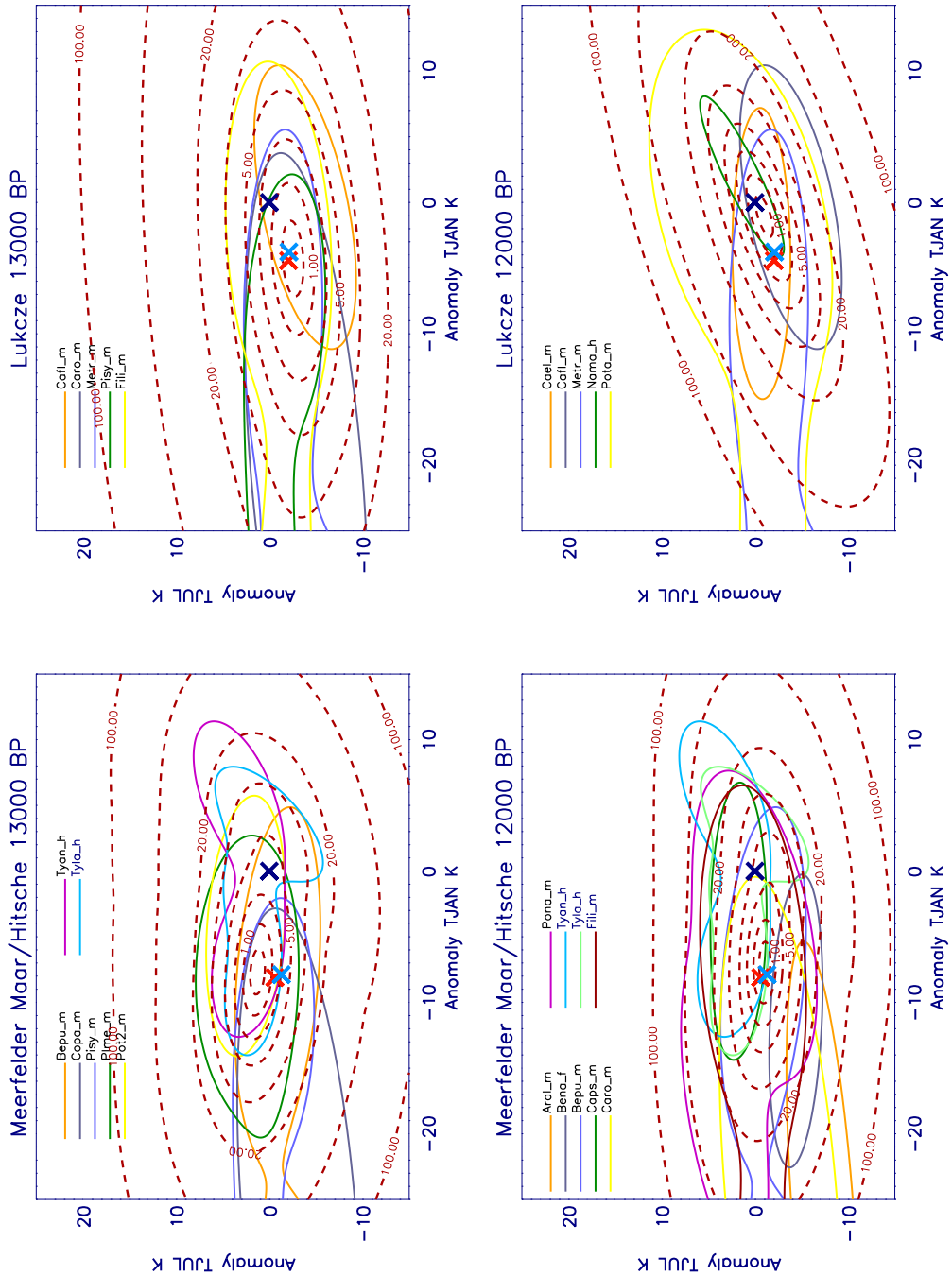

Figure A.9: Local reconstruction for the sites Lukcze and Meerfelder Maar for the Alleroed (left) and YD (right). Pdfs of occurring taxa are represented by their isolines of 33% of the maximum density. The resulting pdf (red dashed lines) is displayed in terms of vegetational costs (ln of the pdf values normalised by the maximum of the pdf). Modern climate is marked by a dark X. The reconstruction results of the variational analysis are marked by a red and blue X for Alleroed and Younger Dryas, respectively.

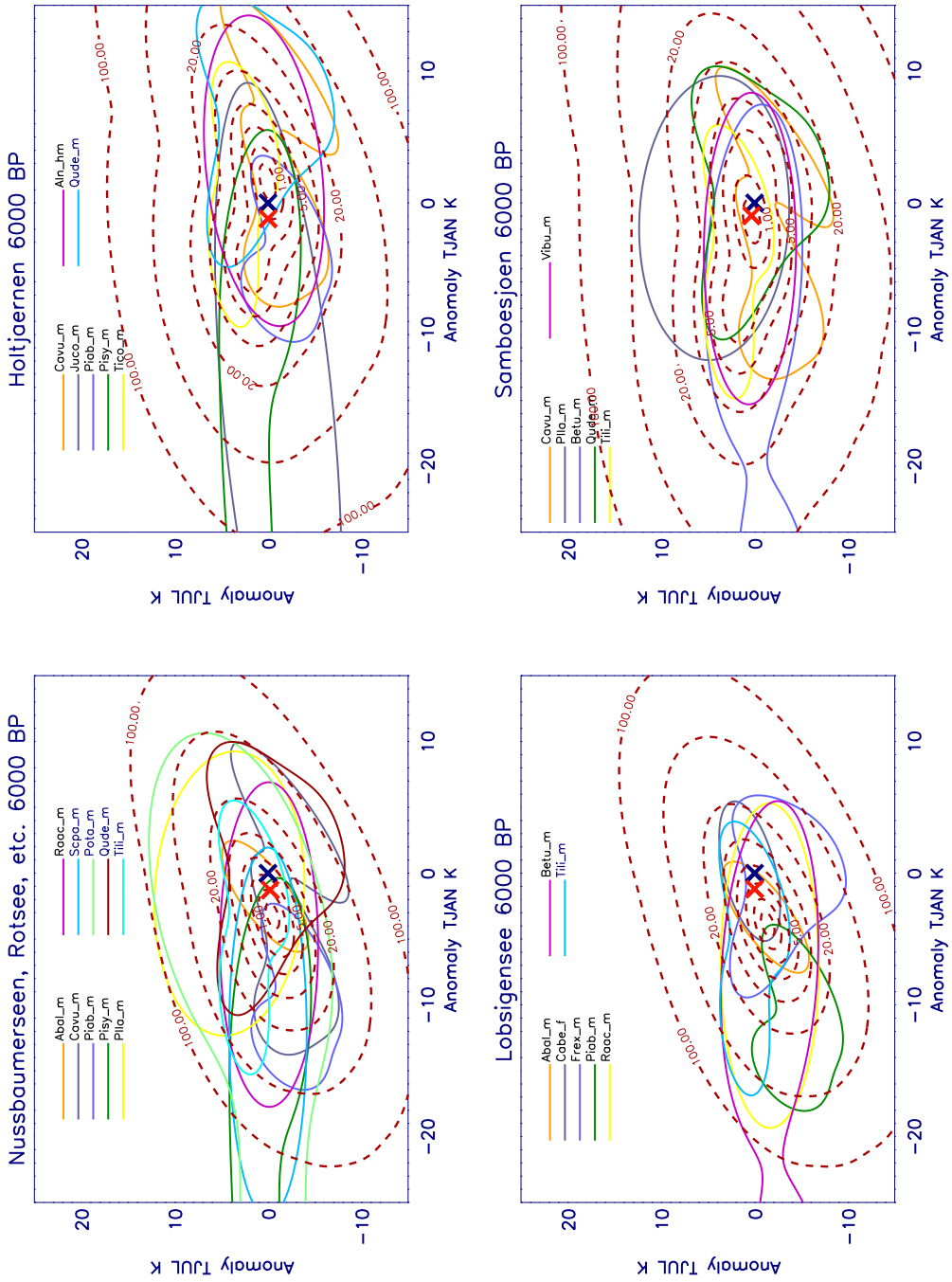

Figure A.10: Local reconstruction for the sites Holtjaernen, Samboesjoen and the combined sites Lobsigensee/Loermoss and Nussbaumerseen/Rotsee/Breitnau/Soppensee/Steerenmoos for 6000 BP. Pdfs of occurring taxa are represented by their isolines of 33% of the maximum density. The resulting pdf (red dashed lines) is displayed in terms of vegetational costs (ln of the pdf values normalised by the maximum of the pdf). Modern climate is marked by a dark X while the reconstruction results of the variational analysis are marked by a red X.

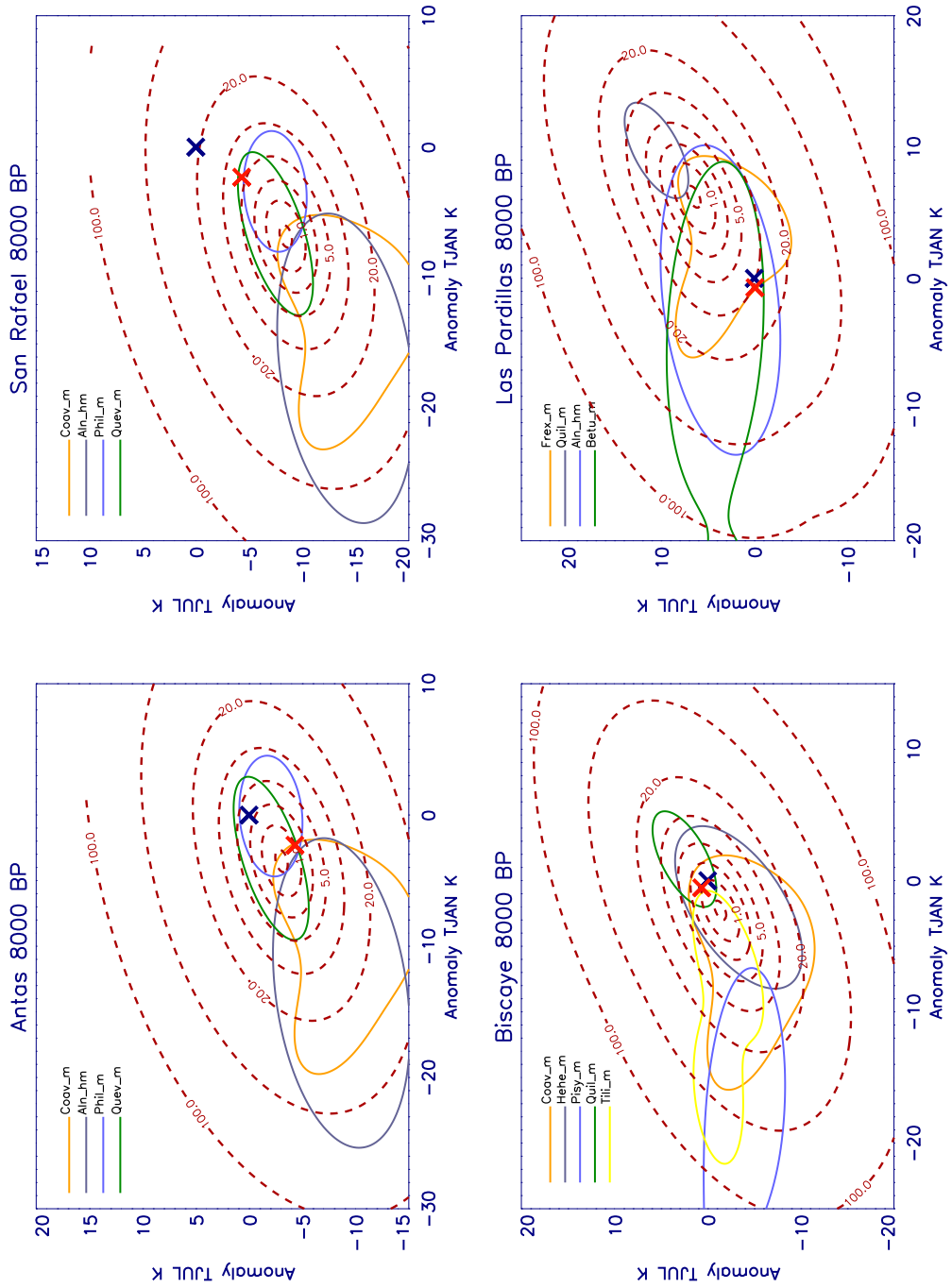

Figure A.11: Local reconstruction for the sites San Rafael, Las Pardillas, Antas and Biscaye for 8000 BP. Pdfs of occurring taxa are represented by their isolines of 33% of the maximum density. The resulting pdf (red dashed lines) is displayed in terms of vegetational costs (ln of the pdf values normalised by the maximum of the pdf). Modern climate is marked by a dark X while the reconstruction results of the variational analysis are marked by a red X.

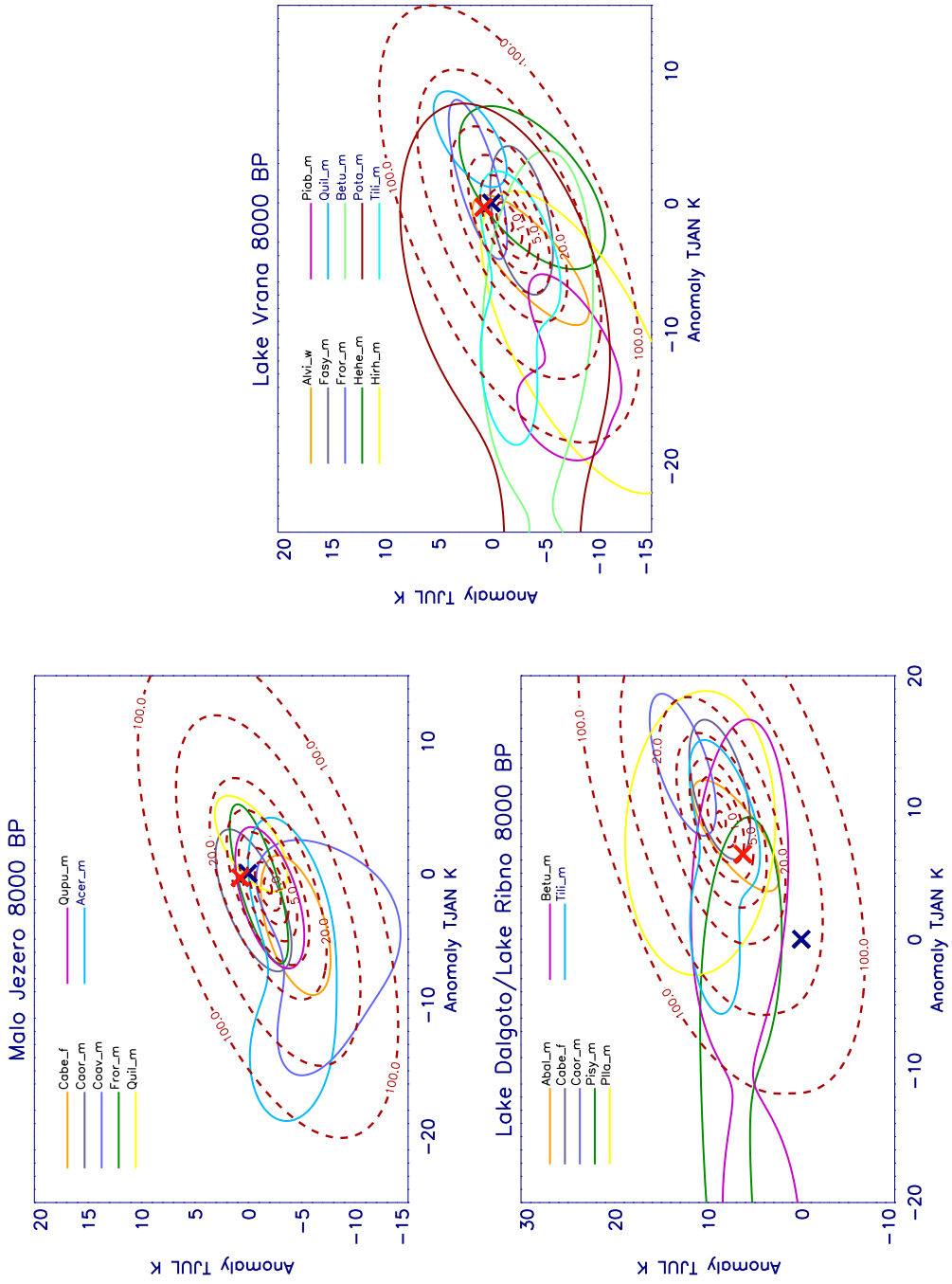

Figure A.12: Local reconstruction for the sites Lake Vrana, Malo Jezero and Lake Dalgoto for 8000 BP. Pdfs of occurring taxa are represented by their isolines of 33% of the maximum density. The resulting pdf (red dashed lines) is displayed in terms of vegetational costs (ln of the pdf values normalised by the maximum of the pdf). Modern climate is marked by a dark X while the reconstruction results of the variational analysis are marked by a red X.

Appendix B

Fossil sites and present taxa

In the following pages all fossil sites, from which data have been use in this work are listed. The first tables contain site number and names, together with coordinates, site altitudes and references. These are followed by tables containing all present taxa for the respective time slices. Taxa which were sorted out by the preselection procedure are left black, while taxa which have been used for the reconstructions are indicated green. In the second column of these tables is show, what kind of data has been used: macrofossil (m), pollen (p) or both (m,p).

Nr.	Site	Lon	Lat	Alt	References
1	Dalmutladdo	20.5	69.5	355m	Bjune et al. (2004)
2	Toskaljavri	21.5	69.5	704m	Seppä and Birks (2002)
3	Lake Tsuolbmajavri	22.5	68.5	526m	Seppä and Birks (2001)
4	Abbotjaernen	14.5	63.5	250m	Gisecke (2005b)
5	Brurskardtjoerni	14.5	63.5	1310m	Bjune (2005)
6	Klotjaernen	16.5	61.5	160m	Gisecke (2005a)
7	Laihalampi	26.5	61.5	137m	Heikkilä and Seppä (2003)
8	Dale	5.5	60.5	36m	Mangerud (1970) Birks (1993, 1994)
9	Trettetjoern	7.5	60.5	819m	Bjune et al. (2005) Bjune (2005)
10	Holtjaernen	14.5	61.5	110m	Gisecke (2005a)
11	Hirvilampi	24.5	61.5	114m	Rankama and Vuorela (1988)
12	Utsira	5.5	59.5	35m	Mangerud (1970) Paus (1990) Birks (1993, 1994)
13	Vestre	6.5	59.5	570m	Bjune et al. (2005) Bjune (2005)
14	Eigebakken	6.5	58.5	27m	Paus (1989)
15	Flarken	13.5	58.5	110m	Digerfeldt (1977) Seppä et al. (2005)
16	Raigastvere	26.5	58.5	52m	Pirrus et al. (1987)
17	Samboesjoen	12.5	57.5	35m	Digerfeldt (1982)
18	Loch Maree	-5.5	57.5	107m	Birks (1972)
19	Dubh Lochan	-4.5	56.5	75m	Stewart et al. (1984)
20	Abernethy Forest	-3.5	56.5	221m	Birks (1970) Birks and Matthews (1978)
21	Bjoerkeroeds Mosse Haekulls Mosse	12.5	56.5	100m	Liedberg (1988)
22	Machrie Moor	-4.5	55.5	50m	Robinson and Dickson (1988)
23	Toppeladugaerd	13.5	55.5	35m	Liedberg (1988)
24	Sluggan Moss	-6.5	54.5	52m	Smith and Goddard (1991)
25	Zarnowiec	17.5	54.5	5m	Latalowa (1982)
26	Swietoujsc	14.5	53.5	4m	Latalowa and Borowka (2006)
27	Maly Suszek	17.5	53.5	115m	Miotk-Szpiganowicz (1992)
28	Strazym	19.5	53.5	4m	Boinska (1987)

Nr.	Site	Lon	Lat	Alt	References
29	Stare Biele	23.5	53.5	143m	Kuprijanowicz (2000) Marek (2000)
30	Nant Francon	-3.5	52.4	182m	Borrows (1974, 1975) Switsur and West (1973)
31	King's Pool	-2.5	52.5	100m	Bartley and Morgan (1990)
32	Hockam Mere	0.5	52.5	33m	Godwin and Tallantire (1951) Bennett (1983)
33	Haemelsee	9.5	52.5	20m	Merkt and Müller (1999)
34	Treppelsee	14.5	52.5	52m	Gisecke (2000)
35	Skrzetuszewskie	17.5	52.5	109m	Tobolski (1990)
36	Lake Gosciaz	19.5	52.5	64m	Ralska-Jasiewiczowa et al. (1998)
37	Bledowo	20.5	52.5	78m	Binka et al. (1991)
38	Milheeze	5.5	51.5	27m	Bos (1998) Bos et al. (2006)
39	Krumpa	11.5	52.5	115m	Früchtl (1998)
40	Central Rhineland	7.5	50.5	60m	Baales et al. (2002)
41	Lukcze	22.5	51.5	163m	Balaga (1990)
42	Meerfelder Maar Hitsche	6.5	50.5	335m	Brauer et al. (1999) Kubitz (2000) Stebich (1999) Früchtl (1998)
43	Wetterau	8.5	50.5	145m	Bos (1998)
44	Czajkow	21.5	50.5	205m	Szczepanek (1971) Srodon (1983)
45	Svarcenberk Rezabinec	14.5	49.5	400m	Jankovska (1980) Pokorny and Jankovská (2000) Pokorny (2002) Rybnickova and Rybnicek (1985)
46	Malopolskie	19.5	49.5	656m	Obidowicz (1988, 1989, 1990, 1993)
47	Sewensee	6.5	47.5	501m	Juvigné (1977) Schloss (1979) Woillard (1975, 1978)
48	Tarnowiec Besko Roztoki	21.5	49.5	240m	Koperowa (1970) Harmata (1987) Srodon (1990)
49	Feigne d'Artimont	7.5	48.5	1100m	Janssen et al. (1975)
50	Ried bei Oberschan	9.5	47.5	640m	Wegmüiller (1976)

Nr.	Site	Lon	Lat	Alt	References
51	Lobsigensee Loermoss	7.5	47.5	550m	Ammann and Tobolski (1983) Ammann et al. (1985) Ammann (1989) Zwahlen (1985)
52	Nussbaumerseen Rotsee Breitnau Soppensee Steerenmoss	8.5	47.5	610m	Rösch (1983, 1985) Lotter (1988, 1991) Lotter et al. (1992) Lotter (1997, 1999) Rösch (1989, 2000)
53	Lanser See	11.5	47.5	840m	Bortenschlager (1984) Oeggl (1992)
54	Fuschlsee	13.5	47.5	663m	Voigt (1996)
55	Steregoiu	23.5	47.5	1300m	Björkman et al. (2002)
56	Gola di Lago	8.5	46.5	970m	Zoller and Kleiber (1971)
57	Lago di Ganna	8.5	45.5	452m	Schneider and Tobolski (1983)
58	Le Tronchet Marais du Rosey	6.5	46.5	600m	Gaillard (1984, 1985)
59	Lac du Bouchet	3.5	44.5	1200m	Reille and de Beaulieu (1988)
60	Lac de Saint	6.5	44.5	1308m	Digerfeldt et al. (1997)
61	Lago Padula	10.5	44.5	1187m	Watson (1996)
62	Le Moura	-1.5	43.5	40m	Reille (1993)
63	Lake Vrana	14.5	44.5	13m	Jahns and van den Bogaard (1998) Schmidt et al. (2000)
64	Biscaye	-0.5	43.5	410m	Reille and Andrieu (1995)
65	Balma	2.5	43.5	560m	Leroyer and Heinz (1992)
66	Biot	7.5	43.5	50m	Nicol-Pichard and Dubar (1998)
67	Malo Jezero	17.5	43.5	0m	Jahns and van den Bogaard (1998)
68	Las Pardillas	-3.5	42.5	1850m	Sánchez and Hannon (1999)
69	Troubat	0.5	42.5	541m	Heinz and Barbaza (1998)
70	Lake Dalgoto Lake Ribno	23.5	41.5	2310m	I. and Ammann (2003) Tonkov et al. (2002)
71	Lago di Monticchio	15.5	40.5	1326m	Watts et al. (1996b,a)
72	Canada de la Cruz	-2.5	38.5	1350m	Carrión et al. (2002)
73	Antas	-1.5	37.5	0m	Pantaléon-Cano et al. (2003)
74	San Rafael	-2.5	37.5	0m	Pantaléon-Cano et al. (2003)

Nr.	Dt.	Site	Time slice	Occuring taxa
1	m,p	Dalmutladdo	8000	*Betula, Juniperus communis, Pinus sylvestris, Filipendula Potamogeton, Rumex acetosa, Rumex acetosella, Picea abies*
			6000	*Alnus, Betula, Corylus avellana, Fraxinus excelsior Pinus sylvestris, Ulmus, Filipendula, Thalictrum aquilegifolium Selaginella selaginoides, Juniperus communis, Picea abies*
2	p	Toskaljavri	8000	*Betula, Juniperus communis, Pinus sylvestris*
			6000	*Betula, Juniperus communis, Pinus sylvestris, Calluna vulgaris*
3	p	Lake Tsuolbmajavri	8000	*Alnus, Betula, Juniperus communis, Pinus sylvestris*
			6000	*Betula, Pinus sylvestris, Picea abies*
4	p	Abbortjaernen	8000	*Alnus, Juniperus communis, Pinus sylvestris Betula, Filipendula*
			6000	*Alnus, Fraxinus excelsior, Juniperus communis, Pinus sylvestris Betula, Ulmus, Filipendula, Juniperus communis*
5	m,p	Brurskardtjoerni	8000	*Alnus, Juniperus communis, Pinus sylvestris ,Potamogeton Rumex acetosa, Rumex acetosella*
			6000	*Betula, Ranunculus acris, Rumex acetosella Alnus, Pinus sylvestris, Rumex acetosa*
6	p	Klotjaernen	8000	*Juniperus communis, Picea abies, Pinus sylvestris, Tilia cordata Calluna vulgaris, Nymphaea alba, Alnus, Betula Fraxinus excelsior, Filipendula, Potamogeton*
			6000	*Betula,Fraxinus excelsior,Hippophae rhamnoides Calluna vulgaris, Potamogeton, Rumex acetosa, Alnus, Ulmus Corylus avellana, Pinus sylvestris, Menyanthes trifoliata Juniperus communis, Nymphaea alba*

Nr.	Dt.	Site	Time slice	Occuring taxa
7	p	Laihalampi	8000	*Juniperus communis, Pinus sylvestris, Populus tremula, Picea abies Alnus, Betula, Corylus avellana, Fraxinus excelsior, Filipendula, Ulmus*
			6000	*Juniperus communis, Pinus sylvestris, Populus tremula, Tilia Ulmus, Alnus, Betula, Corylus avellana, Picea abies*
8	m,p	Dale	13000	*Betula pubescens, Salix herbacea, Menyanthes trifoliata Juniperus communis*
			12000	*Betula nana, Juniperus communis, Salix herbacea*
9	p	Trettetjoern	8000	*Dryas octopetala, Filipendula, Rumex acetosa, Betula Pinus sylvestris, Rumex acetosella*
			6000	*Juniperus communis, Pinus sylvestris, Quercus deciduous Ranunculus acris, Alnus, Betula, Ulmus, Calluna vulgaris Viburnum opulus, Filipendula, Rumex acetosa*
10	p	Holtjaernen	8000	*Juniperus communis, Pinus sylvestris, Quercus deciduous, Tilia cordata Calluna vulgaris, Rumex acetosa, Alnus, Betula, Rumex acetosella Corylus avellana, Fraxinus excelsior, Ulmus, Filipendula, Potamogeton*
			6000	*Alnus, Pinus sylvestris, Quercus deciduous, Tilia cordata Calluna vulgaris, Corylus avellana, Juniperus communis Picea abies, Betula, Fraxinus excelsior*
11	p	Hirvilampi	8000	*Alnus, Populus tremula, Filipendula, Myriophyllum alterniflorum Betula, Corylus avellana, Pinus sylvestris, Ulmus Cornus mas, Corylus avellana, Sambucus, Tilia, Picea abies*
			6000	*Calluna vulgaris, Filipendula, Alnus, Betula Pinus sylvestris, Ulmus, Picea abies*
12	m,p	Utsira	13000	*Salix herbacea, Carex nigra, Betula pubescens, Empetrum nigrum Koenigia islandica, Juniperus communis*
			12000	*Betula nana, Juniperus communis, Koenigia islandica, Carex nigra Salix herbacea, Empetrum nigrum, Rumex acetosa, Rumex acetosella*

Nr.	Dt.	Site	Time slice	Occuring taxa
13	m,p	Vestre	13000	*Juniperus communis, Pinus sylvestris, Salix herbacea, Plantago major Rumex acetosa, Empetrum nigrum, Betula, Filipendula*
			12000	*Betula nana, Juniperus communis, Pinus sylvestris, Rumex acetosa Salix herbacea, Filipendula*
			8000	*Betula pubescens, Quercus deciduous, Salix herbacea, Tilia Calluna vulgaris, Filipendula,Rumex acetosa, Juniperus communis Alnus, Corylus avellana, Fraxinus excelsior, Pinus sylvestris, Ulmus*
			6000	*Alnus, Juniperus communis, Myrica gale, Pinus sylvestris Populus tremula, Quercus deciduous, Betula, Corylus avellana, Ulmus*
14	p	Eigebakken	13000	*Betula pubescens, Corylus avellana, Populus tremula, Dryas octopetala Koenigia islandica, Rumex acetosa, Sanguisorba officinalis, Ulmus, Alnus Juniperus communis, Pinus sylvestris, Quercus deciduous, Filipendula*
			12000	*Betula nana, Populus tremula, Dryas octopetala, Sanguisorba officinalis Plantago major, Rumex acetosa, Empetrum nigrum Juniperus communis, Filipendula*
15	p	Flarken	12000	*Hippophae rhamnoides, Juniperus communis, Myriophyllum alterniflorum Pinus sylvestris, Rumex acetosa, Betula nana, Filipendula*
			8000	*Pinus sylvestris, Quercus deciduous, Tilia, Calluna vulgaris Rumex acetosa, Alnus, Betula, Corylus avellana Juniperus communis, Ulmus, Filipendula*
			6000	*Betula, Corylus avellana, Tilia, Calluna vulgaris Potamogeton, Rumex acetosa, Alnus, Pinus sylvestris Quercus deciduous, Ulmus, Filipendula*
16	p	Raigastvere	8000	*Betula, Quercus deciduous, Tilia, Alnus, Picea abies Corylus avellana, Pinus sylvestris, Ulmus*
			6000	*Betula, Quercus deciduous, Tilia, Ulmus Corylus avellana, Pinus sylvestris, Picea abies*

Nr.	Dt.	Site	Time slice	Occuring taxa
17	p	Samboesjoen	8000	*Quercus deciduous, Tilia, Calluna vulgaris, Cladium mariscus* *Rumex acetosa, Alnus, Betula, Corylus avellana* *Pinus sylvestris, Ulmus, Myriophyllum alterniflorum*
			6000	*Betula, Quercus deciduous, Tilia, Viburnum* *Calluna vulgaris, Plantago lanceolata, Alnus, Ulmus* *Corylus avellana, Pinus sylvestris, Rumex acetosella*
18	m,p	Loch Maree	8000	*Juniperus communis, Pinus sylvestris, Quercus deciduous* *Betula, Corylus avellana, Fraxinus excelsior, Filipendula* *Calluna vulgaris, Rumex acetosella*
			6000	*Alnus glutinosa, Fraxinus excelsior, Juniperus communis* *Calluna vulgaris, Plantago major, Rumex acetosa, Betula* *Quercus deciduous, Ulmus, Pinus sylvestris*
19	m,p	Dubh Lochan	8000	*Betula, Quercus deciduous, Calluna vulgaris, Potamogeton* *Pinus sylvestris, Ulmus, Nymphaea alba*
			6000	*Betula, Calluna vulgaris, Nymphaea alba, Potamogeton natans* *Alnus glutinosa, Corylus avellana, Fraxinus excelsior, Pinus sylvestris* *Ulmus, Quercus deciduous*
20	m,p	Abernethy Forest	13000	*Betula nana, Carex nigra, Carex rostrata, Filipendula* *Juniperus communis, Rumex acetosella, Selaginella selaginoides* *Calluna vulgaris, Empetrum nigrum*
			12000	*Betula nana, Ranunculus sceleratus, Selaginella selaginoides* *Myriophyllum alterniflorum, Rumex acetosella*
			8000	*Betula, Corylus avellana,Calluna vulgaris, Potamogeton* *Juniperus communis*

Nr.	Dt.	Site	Time slice	Occuring taxa
21	m	Bjoerkeroeds Mosse Haekulls Mosse	13000	*Arctostaphylos alpinus, Arctostaphylos uva-ursi, Betula nana* *Dryas octopetala, Selaginella selaginoides*
			12000	*Salix herbacea, Betula pubescens, Arctostaphylos alpinus* *Dryas octopetala, Rumex acetosella, Betula nana* *Urtica dioica, Selaginella selaginoides*
22	m,p	Machrie Moor	8000	*Alnus, Betula, Cladium mariscus, Quercus deciduous, Potamogeton* *Plantago lanceolata, Corylus avellana, Ulmus, Filipendula*
			6000	*Cladium mariscus, Alnus, Betula, Quercus deciduous* *Corylus avellana, Ulmus*
23	m	Toppeladugaerd	13000	*Arctostaphylos alpinus, Dryas octopetala, Potamogeton* *Juniperus communis, Prunus spinosa, Betula nana*
24	p	Sluggan Moss	8000	*Alnus, Betula, Quercus deciduous, Menyanthes trifoliata* *Corylus avellana, Ulmus, Pinus sylvestris*
			6000	*Alnus, Betula, Quercus deciduous, Ulmus* *Corylus avellana, Pinus sylvestris*
27	p	Maly Suszek	13000	*Betula, Dryas octopetala, Typha latifolia*
			12000	*Betula, Juniperus communis, Myriophyllum alterniflorum* *Myriopyhyllum spicatum, Rumex acetosa, Pinus sylvestris*
			8000	*Juniperus communis, Pinus sylvestris, Tilia, Picea abies* *Calluna vulgaris, Rumex acetosa, Quercus deciduous, Picea abies* *Corylus avellana, Ulmus, Filipendula, Alnus, Betula*
			6000	*Acer, Alnus, Carpinus betulus, Quercus deciduous* *Betula, Calluna vulgaris, Corylus avellana, Ulmus* *Frangula alnus, Fraxinus excelsior, Pinus sylvestris*
28	m	Strazym	13000	*Pinus sylvestris, Menyanthes trifoliata, Potamogeton*

Nr.	Dt.	Site	Time slice	Occuring taxa
25	m	Zarnowiec	8000	*Alnus, Quercus deciduous, Tilia, Pinus sylvestris* *Betula, Corylus avellana, Ulmus*
			6000	*Alnus, Quercus deciduous, Tilia, Potamogeton gramineus, Betula* *Sparganium minimum, Urtica dioica, Corylus avellana, Ulmus*
26	m	Swietoujsc	13000	*Alnus, Pinus sylvestris, Populus tremula* *Arctostaphylos uva-ursi*
29	p	Stare Biele	13000	*Betula pendula, Pinus sylvestris, Betula nana, Filipendula*
			12000	*Betula nana, Juniperus communis, Filipendula* *Nymphaea alba, Selaginella selaginoides*
			8000	*Betula, Fagus sylvatica, Tilia, Calluna vulgaris, Alnus* *Corylus avellana, Ulmus, Quercus deciduous, Nymphaea alba*
			6000	*Alnus, Carpinus betulus, Juniperus communis, Picea abies* *Quercus deciduous, Tilia cordata, Calluna vulgaris, Filipendula* *Betula, Corylus avellana, Ulmus, Menyanthes trifoliata*
30	m	Nant Francon	13000	*Betula nana, Juniperus communis, Filipendula ulmaria* *Ranunculus flammula, Rumex acetosa, Betula pubescens*
			12000	*Salix herbacea, Luzula campestris*
32	p	Hockam Mere	12000	*Betula, Juniperus communis, Potamogeton, Myriophyllum alterniflorum* *Rumex acetosa, Myriopyhyllum spicatum, Filipendula*
			8000	*Corylus avellana, Pinus sylvestris, Tilia cordata, Typha latifolia* *Sambucus nigra, Alnus glutinosa, Betula, Hedera helix* *Quercus deciduous, Ulmus, Nymphaea alba*
			6000	*Alnus glutinosa, Betula, Fraxinus excelsior, Hedera helix* *Sambucus nigra, Taxus baccata, Tilia cordata, Plantago lanceolata* *Potamogeton, Corylus avellana, Myrica gale, Pinus sylvestris* *Quercus deciduous, Ulmus, Nymphaea alba*

Nr.	Dt.	Site	Time slice	Occuring taxa
31	p	King's Pool	8000	*Pinus sylvestris, Quercus deciduous, Tilia, Plantago lanceolata Potamogeton, Rumex acetosa, Alnus, Corylus avellana Betula, Ulmus, Filipendula, Plantago major*
			6000	*Pinus sylvestris, Quercus deciduous, Tilia, Plantago lanceolata Potamogeton, Rumex acetosa, Alnus, Corylus avellana Betula, Ulmus, Filipendula, Plantago major*
35	p	Skrzetuszewskie	12000	*Betula, Juniperus communis, Plantago major, Pinus sylvestris*
			8000	*Acer, Picea abies, Pinus sylvestris, Quercus deciduous, Alnus Tilia platyphyllos, Calluna vulgaris, Betula, Corylus avellana, Ulmus*
			6000	*Alnus, Betula, Quercus deciduous, Tilia, Ulmus, Picea abies Calluna vulgaris, Acer, Corylus avellana, Pinus sylvestris*
36	m,p	Lake Gosciaz	12000	*Betula nana, Pinus sylvestris, Salix pentandra, Rumex acetosa Typha latifolia, Myriophyllum verticalum, Empetrum nigrum Menyanthes trifoliata, Plantago media, Juniperus communis Filipendula, Sanguisorba officinalis*
			8000	*Betula, Picea abies, Pinus sylvestris, Quercus deciduous Tilia cordata, Tilia platyphyllos, Calluna vulgaris, Rumex acetosa Alnus, Corylus avellana, Fraxinus excelsior, Filipendula Ulmus, Plantago media, Potamogeton, Typha latifolia*
			6000	*Betula, Fagus sylvatica, Fraxinus excelsior, Hedera helix Picea abies, Salix pentandra, Taxus baccata, Tilia cordata Viburnum opulus, Calluna vulgaris, Cladium mariscus, Typha latifolia Alnus, Carpinus betulus, Corylus avellana, Frangula alnus Pinus sylvestris, Quercus deciduous, Ulmus, Filipendula, Acer Rumex acetosa, Thalictrum aquilegifolium, Typha angustifolia*

Nr.	Dt.	Site	Time slice	Occuring taxa
34	p	Treppelsee	12000	*Betula, Hippophae rhamnoides, Juniperus communis Typha angustifolia, Pinus sylvestris, Filipendula, Rumex acetosa*
			8000	*Alnus, Pinus sylvestris, Quercus deciduous, Tilia, Betula, Picea abies Corylus avellana, Populus tremula, Ulmus, Rumex acetosa*
			6000	*Alnus, Betula, Quercus deciduous, Calluna vulgaris, Picea abies Thalictrum aquilegifolium, Corylus avellana, Tilia, Ulmus*
37	m,p	Bledowo	13000	*Pinus sylvestris, Carex rostrata, Potamogeton pusillus Menyanthes trifoliata, Zannichellia palustris Urtica dioica, Potamogeton praelongus*
			12000	*Betula nana, Pinus sylvestris, Carex rostrata, Potamogeton Zannichellia palustris, Selaginella selaginoides*
			8000	*Betula, Hedera helix, Picea abies, Quercus deciduous, Alnus Tilia, Calluna vulgaris, Corylus avellana, Ulmus, Nymphaea alba*
			6000	*Corylus avellana, Fagus sylvatica, Hedera helix, Picea abies Tilia, Calluna vulgaris, Filipendula, Plantago lanceolata, Ulmus Rumex acetosa, Typha latifolia, Alnus, Betula, Carpinus betulus*
38	m	Milheeze	13000	*Menyanthes trifoliata, Nuphar lutea, Potamogeton Scirpus lacustris, Sparganium minimum, Carex rostrata Nymphaea alba, Typha latifolia, Ranunculus flammula*
40	m,p	Central Rhineland	13000	*Betula pendula, Betula pubescens, Carex flava, Lychnis flos-cucul Typha latifolia, Urtica dioica, Populus tremula, Pinus sylvestris Hippuris vulgaris, Betula humilis, Alisma plantago-aquatica Carex pseudocyperus, Carex riparia, Carex rostrata Potamogeton gramineus, Ranunculus sceleratus, Rubus caesius Salix pentandra, Galium palustre, Plantago major Salix caprea, Filipendula ulmaria, Solanum dulcamara*

Nr.	Dt.	Site	Time slice	Occuring taxa
39	m,p	Krumpa	13000	*Betula pubescens, Pinus sylvestris, Comarum palustre* *Lycopus europaeus, Ranunculus sceleratus, Typha latifolia* *Ceratophyllum demersum, Chenopodium glaucum* *Potamogeton frisii, Potamogeton pectinatus, Potentilla anserina* *Juniperus communis, Betula humilis, Hippuris vulgaris* *Linum perenne, Carex rostrata, Filipendula ulmaria*
			12000	*Betula pubescens, Betula humilis, Alisma plantago-aquatica* *Carex nigra, Hippuris vulgaris, Potamogeton natans* *Potentilla anserina, Pinus sylvestris, Filipendula* *Carex rostrata, Myriopyhyllum spicatum*
42	m,p	Meerfelder Maar Hitsche	13000	*Betula pubescens, Comarum palustre, Populus tremula* *Typha angustifolia, Typha latifolia, Pinus sylvestris, Filipendula* *Menyanthes trifoliata, Eleocharis palustris, Betula* *Myriopyhyllum spicatum, Potamogeton, Urtica dioica* *Plantago media, Juniperus communis, Plantago major*
			12000	*Filipendula, Typha angustifolia, Typha latifolia, Betula nana* *Betula pubescens, Arctostaphylos alpinus, Carex pseudocyperus* *Carex rostrata, Potamogeton natans, Juniperus communis* *Myriopyhyllum spicatum, Eleocharis palustris, Selaginella selaginoides*
			8000	*Alnus, Betula, Hedera helix, Quercus deciduous, Acer* *Tilia, Myriopyhyllum spicatum, Corylus avellana, Pinus sylvestris* *Populus tremula, Ulmus, Filipendula, Potamogeton*
			6000	*Acer, Betula, Hedera helix, Quercus deciduous* *Alnus, Myriophyllum alterniflorum, Corylus avellana, Ulmus* *Pinus sylvestris, Ranunculus acris, Typha angustifolia*
43	m	Wetterau	13000	*Menyanthes trifoliata*

Nr.	Dt.	Site	Time slice	Occuring taxa
33	p	Haemelsee	13000	*Myriopyhyllum spicatum, Populus tremula, Empetrum nigrum Pinus sylvestris, Betula pubescens, Pinus sylvestris, Filipendula*
			12000	*Juniperus communis, Filipendula, Empetrum nigrum Selaginella selaginoides, Betula*
41	m,p	Lukcze	13000	*Carex flava, Carex rostrata, Pinus sylvestris Filipendula, Menyanthes trifoliata, Betula*
			12000	*Carex elongata, Carex flava, Menyanthes trifoliata Najas marina, Potamogeton, Carex acutiformis, Carex rostrata Pinus sylvestris, Betula, Myriopyhyllum spicatum*
			8000	*Alnus, Picea abies, Quercus deciduous, Tilia Filipendula, Menyanthes trifoliata, Corylus avellana, Ulmus Fraxinus excelsior, Myriopyhyllum spicatum, Typha latifolia*
			6000	*Corylus avellana, Fagus sylvatica, Picea abies, Tilia Juniperus communis, Calluna vulgaris, Rumex acetosa Rumex acetosella, Alnus, Carpinus betulus, Fraxinus excelsior Quercus deciduous, Typha latifolia, Ulmus*
45	m,p	Rezabinec Svarcenberk	13000	*Betula pubescens, Potamogeton gramineus*
			8000	*Betula pubescens, Corylus avellana, Fagus sylvatica, Tilia Pinus sylvestris, Picea abies, Calluna vulgaris, Ulmus Najas marina, Lycopus europaeus, Menyanthes trifoliata Rumex acetosa, Alnus glutinosa, Quercus deciduous, Filipendula Ceratophyllum demersum, Nuphar lutea, Nymphaea alba Myriopyhyllum spicatum, Potamogeton natans, Typha latifolia*
			6000	*Alnus, Abies alba, Tilia, Filipendula Quercus deciduous, Corylus avellana, Ulmus*

Nr.	Dt.	Site	Time slice	Occuring taxa
44	m,p	Czajkow	13000	*Pinus sylvestris, Carex riparia, Menyanthes trifoliata, Potamogeton Ranunculus sceleratus, Sparganium minimum, Larix, Comarum palustre Hippuris vulgaris, Filipendula, Betula pendula, Plantago major*
			12000	*Larix, Typha latifolia, Carex riparia, Menyanthes trifoliata Potamogeton, Ranunculus sceleratus, Sparganium minimum, Filipendula Pinus sylvestris, Plantago major, Hippuris vulgaris, Comarum palustre*
47	m	Sewensee	13000	*Betula pubescens, Potamogeton gramineus, Potamogeton pusillus*
			12000	*Alnus viridis, Betula nana*
48	m,p	Tarnowiec Besko Roztoki	13000	*Lycopus europaeus, Betula nana, Picea abies, Pinus sylvestris Filipendula, Plantago media, Alnus, Typha angustifolia Typha latifolia, Carex rostrata, Juniperus communis, Larix Sanguisorba officinalis, Nymphaea alba, Betula pubescens*
			12000	*Abies alba, Carex elongata, Cladium mariscus, Hippuris vulgaris Potamogeton filiformis, Nymphaea alba, Filipendula, Betula nana Pinus cembra, Plantago media, Potamogeton natans, Urtica dioica Carex rostrata, Rubus idaeus, Rumex acetosella Juniperus communis, Typha latifolia, Pinus sylvestris Larix, Menyanthes trifoliata, Rumex acetosa*
			8000	*Acer, Alnus, Betula, Carpinus betulus, Corylus avellana, Ulmus Picea abies, Quercus deciduous, Potamogeton natans Fagus sylvatica, Sambucus nigra, Tilia platyphyllos, Filipendula Rubus idaeus,*
			6000	*Abies alba, Acer, Alnus, Fraxinus excelsior, Betula Picea abies, Rubus idaeus, Tilia platyphyllos, Calluna vulgaris Filipendula, Potamogeton natans, Thalictrum aquilegifolium Corylus avellana, Fagus sylvatica, Quercus deciduous Carpinus betulus, Ulmus, Tilia cordata*

Nr.	Dt.	Site	Time slice	Occuring taxa
46	p	Malopolskie	8000	*Abies alba, Betula nana, Corylus avellana, Tilia Fagus sylvatica, Picea abies, Filipendula, Plantago lanceolata Rumex acetosa, Acer, Alnus viridis, Carpinus betulus Fraxinus excelsior, Pinus sylvestris, Quercus deciduous, Ulmus*
			6000	*Abies alba, Betula nana, Fagus sylvatica, Fraxinus excelsior Hedera helix, Pivea abies, Pinus sylvestris, Calluna vulgaris Plantago lanceolata, Plantago major, Rumex acetosa, Acer Thalictrum aquilegifolium, Alnus viridis, Ulmus, Filipendula Carpinus betulus, Corylus avellana, Quercus deciduous, Tilia*
50	p	Ried bei Oberschan	8000	*Alnus, Betula, Carpinus betulus, Hedera helix Quercus deciduous, Tilia, Potamogeton, Acer Corylus avellana, Ulmus, Nymphaea alba*
			6000	*Abies alba, Betula, Corylus avellana, Fagus sylvatica Tilia, Potamogeton, Acer, Alnus, Ulmus*
51	m,p	Loermoos Lobsigensee	13000	*Betula pendula, Betula pubescens, Pinus sylvestris, Nuphar lutea Menyanthes trifoliata, Hippophae rhamnoides, Lycopus europaeus Potamogeton pusillus, Comarum palustre, Juniperus communis Ranunculus sceleratus, Potamogeton perfoliatus, Filipendula*
			12000	*Betula pendula, Pinus sylvestris, Populus tremula, Nuphar lutea Menyanthes trifoliata, Hippophae rhamnoides, Betula pubescens Potamogeton pusillus, Comarum palustre, Juniperus communis Filipendula, Plantago major, Potamogeton perfoliatus*
			8000	*Abies alba, Betula, Fagus sylvatica, Tilia, Acer, Picea abies Hedera helix, Pinus sylvestris, Corylus avellana, Ulmus Frangula alnus, Quercus deciduous, Filipendula, Typha latifolia*
			6000	*Abies alba, Betula, Carpinus betulus, Fraxinus excelsior Tilia, Ranunculus acris, Acer, Alnus, Corylus avellana, Picea abies Frangula alnus, Pinus sylvestris, Quercus deciduous, Ulmus*

Nr.	Dt.	Site	Time slice	Occuring taxa
49	p	Feigne d'Artimont	8000	*Abies alba, Betula, Corylus avellana, Tilia, Acer, Alnus* *Sanguisorba officinalis, Quercus deciduous, Ulmus, Filipendula*
			6000	*Abies alba, Alnus, Betula, Hedera helix, Quercus deciduous* *Tilia, Plantago lanceolata, Acer, Ulmus*
52	m,p	Nussbaumerseen Rotsee Breitnau Soppensee Steerenmoss	13000	*Betula pendula, Pinus sylvestris, Najas marina, Hippophae rhamnoides* *Rumex acetosa, Sanguisorba officinalis, Potamogeton, Plantago major* *Filipendula, Betula nana, Populus tremula, Juniperus communis*
			12000	*Juniperus communis, Hippophae rhamnoides, Betula pendula* *Menyanthes trifoliata, Pinus sylvestris, Populus tremula* *Betula pubescens, Sanguisorba officinalis, Potamogeton natans* *Najas marina, Filipendula ulmaria, Betula nana*
			8000	*Abies alba, Alnus, Carpinus betulus, Corylus avellana, Picea abies* *Hedera helix, Pinus sylvestris, Tilia, Menyanthes trifoliata* *Plantago lanceolata, Potamogeton, Acer, Betula, Ulmus* *Fagus sylvatica, Juniperus communis, Quercus deciduous, Filipendula*
			6000	*Abies alba, Pinus sylvestris, Quercus deciduous, Tilia* *Calluna vulgaris, Ranunculus acris, Betula, Potamogeton* *Plantago lanceolata, Acer, Alnus, Scheuchzeria palustris* *Corylus avellana, Ulmus, Frangula alnus, Filipendula, Picea abies*
54	p	Fuschlsee	8000	*Acer, Alnus, Hedera helix, Quercus deciduous, Corylus avellana, Betula* *Pinus sylvestris, Tilia platyphyllos, Myriophyllum alterniflorum, Ulmus*
			6000	*Abies alba, Betula, Quercus deciduous, Tilia cordata* *Tilia platyphyllos, Plantago major, Ranunculus acris, Ulmus* *Acer, Alnus, Corylus avellana, Pinus sylvestris, Rumex acetosa*
56	p	Gola di Lago	13000	*Hippophae rhamnoides, Larix, Potamogeton natans, Betula pendula*
			12000	*Hippophae rhamnoides, Betula pendula, Juniperus communis*

Nr.	Dt.	Site	Time slice	Occuring taxa
53	m	Lanser See	13000	*Alnus, Betula pubescens, Pinus sylvestris* *Potamogeton gramineus, Scirpus lacustris*
			12000	*Alnus viridis, Betula nana, Larix* *Potamogeton, Pinus sylvestris*
55	p	Steregoiu	13000	*Alnus, Picea abies, Pinus sylvestris* *Betula, Ulmus, Filipendula*
			12000	*Alnus, Juniperus communis, Plantago lanceolata, Potamogeton* *Pinus sylvestris, Picea abies, Betula, Filipendula* *Plantago major, Rumex acetosa, Typha angustifolia*
			8000	*Quercus deciduous, Sambucus nigra, Tilia, Filipendula, Acer* *Rumex acetosa, Alnus, Betula, Corylus avellana, Ulmus*
			6000	*Corylus avellana, Fagus sylvatica, Hedera helix, Tilia* *Rumex acetosella, Alnus, Betula, Quercus deciduous* *Sambucus nigra, Ulmus, Filipendula*
58	p	Marais du Rosey Le Tronchet	13000	*Lycopus europaeus, Potamogeton praelongus, Filipendula* *Dryas octopetala, Sanguisorba officinalis, Betula pubescens* *Betula nana, Selaginella selaginoides, Juniperus communis* *Potamogeton gramineus, Potamogeton, Hippophae rhamnoides*
			12000	*Lycopus europaeus, Dryas octopetala, Pinus sylvestris* *Potamogeton praelongus, Juniperus communis* *Hippophae rhamnoides, Potamogeton filiformis* *Betula nana, Filipendula, Alnus, Plantago major, Populus tremula* *Corylus avellana, Selaginella selaginoides, Potamogeton natans*
			8000	*Abies alba, Alnus, Tilia, Filipendula* *Corylus avellana, Pinus sylvestris, Ulmus*
			6000	*Abies alba, Alnus, Tilia, Pinus sylvestris* *Corylus avellana, Filipendula, Ulmus, Picea abies*

Nr.	Dt.	Site	Time slice	Occuring taxa
57	m	Lago di Ganna	13000	*Pinus sylvestris, Ulmus, Rumex acetosella* *Filipendula, Juniperus communis, Betula pubescens*
			12000	*Alnus incana, Pinus cembra, Tilia cordata, Alnus viridis* *Corylus avellana, Populus tremula, Filipendula, Potamogeton* *Hippophae rhamnoides, Larex, Pinus sylvestris* *Acer, Juniperus communis, Tila*
59	p	Lac du Bouchet	12000	*Betula, Corylus avellana*
			8000	*Betula, Corylus avellana, Ulmus*
			6000	*Alnus, Betula, Corylus avellana, Hedera helix, Acer* *Quercus ilex, Tilia, Polygonum aviculare, Ulmus*
60	m,p	Lac de Saint	8000	*Abies alba, Alnus, Hedera helix, Quercus ilex* *Betula, Tilia, Plantago lanceolata, Potamogeton* *Nymphaea alba, Acer, Corylus avellana, Ulmus*
			6000	*Abies alba, Acer, Quercus ilex, Filipendula, Corylus avellana* *Alnus, Hedera helix, Nymphaea alba, Betula, Ulmus, Tilia*
63	p	Lake Vrana	13000	*Polygonum aviculare, Fagus sylvatica, Pinus sylvestris, Alnus* *Abies alba, Hedera helix, Quercus ilex, Buxus sempevirens, Betula* *Ulmus, Carpinus betulus, Corylus avellana, Quercus deciduous*
			12000	*Alnus viridis, Fraxinus excelsior, Juniperus communis* *Betula, Tilia, Corylus avellana, Ulmus*
			8000	*Alnus viridis, Betula, Fagus sylvatica, Fraxinus ornus* *Hedera helix, Hippophae rhamnoides, Tilia, Potamogeton* *Phillyrea, Quercus ilex, Acer, Corylus avellana* *Fraxinus excelsior, Ulmus, Filipendula*
			6000	*Abies alba, Alnus, Betula, Fraxinus excelsior, Acer* *Quercus ilex, Tilia, Corylus avellana, Ulmus*

Nr.	Dt.	Site	Time slice	Occuring taxa
61	p	Lago Padule	8000	*Abies alba, Betula, Tilia, Acer, Alnus, Ulmus*
			6000	*Abies alba, Acer, Alnus, Corylus avellana, Tilia, Ulmus*
62	p	Le Moura	13000	*Betula, Typha latifolia, Juniperus communis, Filipendula* *Pinus sylvestris, Nymphaea alba, Quercus deciduous*
64	p	Biscaye	13000	*Viburnum, Plantago lanceolata, Sanguisorba officinalis, Calluna vulgaris* *Filipendula, Hippophae rhamnoides, Quercus deciduous, Betula pendula*
			12000	*Betula, Hippophae rhamnoides, Juniperus communis* *Pinus sylvestris, Quercus deciduous, Filipendula*
			8000	*Corylus avellana, Hedera helix, Pinus sylvestris, Quercus ilex* *Tilia, Alnus, Betula, Populus tremula, Ulmus*
			6000	*Corylus avellana, Quercus ilex, Viburnum, Filipendula, Alnus* *Plantago lanceolata, Thalictrum aquilegifolium, Betula, Tilia, Ulmus*
65	m	Balma	13000	*Acer campestre, Corylus avellana, Pinus sylvestris* *Quercus pubescens, Taxus baccata, Prunus spinosa*
66	p	Biot	13000	*Betula, Quercus robur, Tilia, Ulmus, Corylus avellana*
			12000	*Alnus, Fraxinus excelsior, Tilia, Filipendula* *Acer, Betula, Corylus avellana, Quercus robur*
67	p	Malo Jezero	8000	*Acer, Carpinus betulus, Carpinus orientalis, Corylus avellana* *Fraxinus ornus, Quercus ilex, Quercus pubescens, Alnus* *Fagus sylvatica, Fraxinus excelsior, Ulmus, Plantago lanceolata*
			6000	*Acer, Carpinus betulus, Carpinus orientalis, Fraxinus excelsior* *Alnus, Fraxinus ornus, Pinus sylvestris, Quercus ilex, Ulmus* *Quercus pubescens, Chenopodium glaucum, Corylus avellana*
68	p	Las Pardillas	8000	*Alnus, Fraxinus excelsior, Quercus ilex* *Betula, Corylus avellana, Ulmus*
			6000	*Betula, Quercus ilex, Potamogeton natans* *Ulmus, Corylus avellana, Pinus sylvestris*

Nr.	Dt.	Site	Time slice	Occuring taxa
69	m	Troubat	13000	*Hippophae rhamnoides, Viburnum lantana, Viburnum opulus*
70	p	Lake Dalgoto Lake Ribno	8000	*Abies alba, Betula, Carpinus betulus, Carpinus orientalis* *Tilia, Plantago lanceolata, Pinus sylvestris, Alnus, Acer* *Corylus avellana, Ulmus, Polygonum aviculare, Typha angustifolia*
			6000	*Carpinus betulus, Carpinus orientalis, Corylus avellana* *Tilia, Filipendula, Plantago lanceolata, Betula* *Abies alba, Alnus glutinosa, Ulmus*
71	m,p	Laghi di Monticchio	13000	*Hedera helix, Tilia, Ulmus, Fagus sylvatica* *Filipendula, Acer, Betula*
			12000	*Alnus, Corylus avellana, Fagus sylvatica, Quercus evergreen* *Tilia, Filipendula, Hippophae rhamnoides, Nymphaea alba* *Ulmus, Betula*
			8000	*Carpinus orientalis, Hedera helix, Tilia, Najas minor* *Ulmus, Najas marina, Acer, Corylus avellana*
			6000	*Abies alba, Carpinus betulus, Carpinus orientalis, Tilia, Acer* *Fraxinus ornus, Hedera helix, Najas marina, Najas minor, Ulmus* *Nymphaea alba, Alnus glutinosa, Corylus avellana, Nuphar lutea*
72	p	Canada Cruz	8000	*Quercus ilex*
			6000	*Acer, Betula, Corylus avellana* *Quercus evergreen, Potamogeton, Ulmus* *Pinus sylvestris, Typha angustifolia*
73	m	Antas	8000	*Alnus, Corylus avellana, Phillyrea, Quercus evergreen*
			6000	*Betula, Quercus evergreen*
74	m	San Rafael	8000	*Alnus, Corylus avellana, Phillyrea, Quercus evergreen*
			6000	*Alnus, Corylus avellana, Phillyrea, Quercus evergreen*

List of abbreviations

M_{opt} ..... optimal number of Clusters
P_{ann} ..... Annual precipitation
T_{Jan} ..... January temperature
T_{Jul} ..... July temperature
Abal_m ..... *Abies alba*
Acer_m ..... *Acer*
Aln_hm ..... *Alnus*
Alpl_m ..... *Alisma plantago-aquatica*
Alvi_w ..... *Alnus viridis*
Aral_m ..... *Arctostaphylos alpinus*
Behu_m ..... *Betula humilis*
Bena_f ..... *Betula nana*
Bepu_m ..... *Betula pubescens*
Betu_m ..... *Betula*
BIC ..... Bayesian Information Criterion
BP ..... Before Present
Cabe_f ..... *Carpinus betulus*
Cael_m ..... *Carex elongata*
Cafl_m ..... *Carex flava*
Cani_m ..... *Carex nigra*
Caor_m ..... *Carpinus orientalis*
Caps_m ..... *Carex pseudocyperus*
Caro_m ..... *Carex rostrata*
Cavu_m ..... *Calluna vulgaris*
CE ..... Central-European area
Coav_m ..... *Corylus avellana*
Copa_m ..... *Comarum palustre*
EM ..... Expectation-Maximisation
Emni_m ..... *Empetrum nigrum*
EOF ..... Empirical Orthogonal Function
EPD ..... European Pollen Database

Fasy_m ..... *Fagus sylvatica*
Fili_m ..... *Filipendula*
Frex_m ..... *Fraxinus excelsior*
Fror_m ..... *Fraxinus ornus*
GCM ..... Global Circulation Model
GRIP ..... Greenland Icecore Project
Hehe_m ..... *Hedera helix*
Hirh_m ..... *Hippophae rhamnoides*
Hivu_m ..... *Hippuris vulgaris*
Juco_m ..... *Juniperus communis*
ka ..... kilo years
LGM ..... Last Glacial Maximum
Lipe_m ..... *Linum perenne*
LST ..... Laacher See Tephra
Lyeu_m ..... *Lycopus europaeus*
MAT ..... Modern Analogue Technique
MCMC ..... Markoc Chain Monte Carlo Modeling
MD ..... Mahalanobis Distance
Myal_m ..... *Myriophyllum alterniflorum*
Nama_h ..... *Najas marina*
NW ..... Northwest-European area
PDF,pdf ..... probability density function
pdf2d ..... Two-dimensional mixture model
pdf3d ..... Three-dimensional mixture model
Phil_m ..... *Phillyrea*
Piab_m ..... *Picea abies*
Pisy_m ..... *Pinus sylvestris*
Plla_m ..... *Plantago lanceolata*
Plme_m ..... *Plantago media*
Pona_m ..... *Potamogeton natans*
Pot2_m ..... *Populus tremula*
Pota_m ..... *Potamogeton*
Qude_m ..... *Quercus decidous*
Quev_m ..... *Quercus evergreen*
Quil_m ..... *Quercus ilex*
Qupu_m ..... *Quercus pubescens*
Raac_m ..... *Ranunculus acris*
Rasc_h ..... *Ranunculus sceleratus*
Ruac_h ..... *Rumex acetosa*
Sahe_m ..... *Salix herbacea*
Scpa_m ..... *Scheuchzeria palustris*

Sese_m ..... *Selaginella selaginoides*
SVD ..... Singular Value Decomposition
SW ..... Southwest-European area
THC ..... Thermohaline Circulation
Tico_m ..... *Tilia cordata*
Tili_m ..... *Tilia*
Tyan_h ..... *Typha angustifolia*
Tyla_h ..... *Typha latifolia*
Vibu_m ..... *Viburnum*
YD ..... Younger Dryas

Bibliography

Ammann, B., 1989: Late Quaternary Palynology at Lobsigensee. Regional Vegetation History and Local Lake Development. *Dissertationes Botanicae*, **137**, 1–157.

Ammann, B., M. Andree, L. Chaix, U. Eicher, S. A. Elias, W. Hofmann, H. Oescher, U. Siegenthaler, K. Tobolski, B. Wilkinson, and H. Züllig, 1985: *Swiss lake and mire environments during the last 15000 years*, Dissertationes Botanicae, chapter Lobsigensee - Late-glacial and Holocene environments of a lake on the Central Swiss Plateau. Studies in the Late Quaternary of Lobsigensee 11. Number 87, 127–170.

Ammann, B. and K. Tobolski, 1983: Vegetational development during the Late-Würm at Lobsigensee (Swiss Plateaux).- Studies in the Late-Quaternary of Lobsigensee 1. *Revue de Paléobiologie*, **2**, 163–180.

Andersson, E., J. Haseler, J. Unden, P. Courtier, P. Kelly, D. Vasiljevic, C. Brankovic, C. Cardinali, C. Gaffard, A. Hollingsworth, C. Jakob, P. Janssen, E. Klinker, A. Lanzinger, M. Miller, F. Rabier, A. Simmons, B. Strauss, J.-N. Thépaut, and P. Viterbo, 1998: The ECMWF implementation of three-dimensional variational analysis (3D-var). III: Experimental results. *Quaterly Journal of the Royal Meteorological Society*, **124**, 1831–1859.

Arakawa, A., 1972: Numerical simulation of weather and climate: design of the UCLA general circulation model. Technical Report 7, University of California, Department of Meteorology, Los Angeles, California.

Atkinson, T. C., K. R. Briffa, and G. R. Coope, 1987: Seasonal temperatures in Britain during the past 22000 years, reconstructed using beetle remains. *Nature*, **325**, 587–592.

Baales, M., O. Jöris, M. Street, F. Bittmann, B. Weninger, and J. Wiethold, 2002: Impact of the Late Glacial Eruption of the Laacher See Volcano, Central Rhineland, Germany. *Quaternary Research*, **58**, 273–288.

Bakke, J., O. Lie, E. Heegaard, T. M. Dokken, G. Haug, H. H. Birks, P. Dulski, and T. Nilsen, 2009: Rapid oceanic and atmospheric changes during the Younger Dryas cold period. *Nature Geoscience*, **2**, 202–205.

Balaga, K., 1990: The development of Lake Lukcze and changes in the plant cover of the south-western part of the Leczna-Wlodawa lake district in the last 13000 years. *Acta Palaeobotanica*, **30**, 77–146.

Bar-Matthews, M., A. Ayalon, M. Gilmour, A. Matthews, and C. Hawkesworth, 2003: Sea-land oxygen isotopic relationships from planktonic foraminifera and speleothems in the Eastern Mediterranean region and their implication for paleorainfall during interglacial intervals. *Geochimica et Cosmochimica Acta*, **67**, 3181–3199.

Bartley, D. D. and A. V. Morgan, 1990: The palynological record of the King´s Pool, Stafford, England. *New Phytologist*, **116**, 177–194.

Bayes, T., 1763: An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, **53**, 370–418.

Bennett, K. D., 1983: Devensian late-glacial and flandrian vegetational history at Hockham mere, Norfolk, England. *New Phytologist*, **95**, 489–504.

Berger, A. L., 1978: Long-term Variations of Daily Insolation and Quaternary Climate Changes. *Journal of the Atmospheric Sciences*, **35**, 2362–2367.

Berger, J. O., 1985: *Statistical Decision Theory and Bayesian Analysis*. Springer Series in Statistics, Springer Verlag, second edition edition.

Binka, K., A. Ciesla, B. Lacka, T. Madeyska, B. Marciniak, K. Szeroczynska, and K. Wickowski, 1991: The development of Bledowo Lake (Central Poland) - A palaeoecological study. *Studia Geologica Polonica*, **100**, 1–83.

Birks, H. H., 1970: Studies in the vegetational history of Scotland. I. A Pollen diagram from Abernethy Forest, Inverness-Shire. *Journal of Ecology*, **58**, 827–846.

— 1972: Studies in the vegetational history of Scotland. III. A radiocarbon-dated pollen diagram from Loch Maree, Ross and Cromarty. *New Phytologist*, **71**, 731–754.

Birks, H. H. and R. W. Matthews, 1978: Studies in the vegetational history of Scotland. V. Late Devensian and early Flandrian Pollen and Macrofossil Stratigraphy at Abernethy Forest, Inverness Shire. *New Phytologist*, **80**, 455–484.

Birks, H. J. B., 1993: The importance of Plant Macrofossils in Late-Glacial Climatic reconstructions: An example from Western Norway. *Quaternary Science Reviews*, **12**, 719–726.

— 1994: Late-glacial climatic patterns and ecotones in western Norway. *Vegetation History Archaeobotany*, **3**, 107–119.

— 1995: *Statistical modelling of quaternary science data*, Quaternary Research Association, volume 5, chapter Quantitative paleoenvironmental reconstructions. 161–266.

Björkman, L., A. Feurdean, K. Cinthio, B. Wohlfarth, and G. Possnert, 2002: Lateglacial and early Holocene vegetation development in the Gutaiului Mountains, northwestern Romania. *Quarternary Science Reviews*, **21**, 1039–1059.

Bjune, A. E., 2005: Holocene vegetation history and tree-line changes on a north-south transect crossing major climate gradients in southern Norway - evidence from pollen and plant macrofossils in lake sediments. *Review of Palaeobotany and Palynology*, **133**, 249–275.

Bjune, A. E., J. Bakke, A. Nesje, and H. J. B. Birks, 2005: Holocene mean July temperature and winter precipitation in western Norway inferred from palynological and glaciological lake-sediment proxies. *The Holocene*, **15**, 177–189.

Bjune, A. E., H. J. B. Birks, and H. Seppä, 2004: Holocene vegetation and climate history on a continental-oceanic transect in northern Fennoscandia based on pollen and plant macrofossils. *Boreas*, **33**, 211–223.

Boinska, U., 1987: Analysis of macrofossils in bottom deposits of Lake Strazym (Bordnica Lake district).- PAN. *Acta Palaeobotanica*, **27**, 305–310.

Borrows, C. J., 1974: Plant macrofossils from Late-Devensian deposits at Nant Ffrancon, Caernarvonshire. *New Phytologist*, **73**, 1003–1033.

— 1975: Radiocarbon dates from Late-Devensian deposits, Nant Ffrancon, Caernarvonshire. *New Phytologist*, **75**, 167–171.

Bortenschlager, S., 1984: Die Vegetationsentwicklung im Spätglazial: Das Moor beim Lanser See III, ein Typprofil für die Ostalpen. *Dissertationes Botanicae*, **72**, 71–79.

Bos, J. A. A., 1998: *Aspects of the Lateglacial-Early Holocene Vegetation Development in Western Europe - palynological and palaeobotanical investigations in Brabant (The Netherlands) and Hessen (Germany)*. Number 10 in LPP Contributions Series, Utrecht.

Bos, J. A. A., S. J. P. Bohncke, and C. R. Janssen, 2006: Lake- level fluctuations and small-scale vegetation patterns during the late glacial in The Netherlands. *Journal of Paleolimnology*, **35**, 211–238.

Braconnot, P., B. Otto-Bliesner, S. Harrison, S. Joussaume, J. Y. Peterchmitt, A. Abe-Ouchi, M. Crucifix, E. Driesschaert, T. Fichefet, T. D. Hewitt, M. Kageyama, A. Kitoh, A. Laîné, M.-F. Loutre, O. Marti, U. Merkel, G. Ramstein, P. Valdes, S.-L. Weber, Y. Yu, and Y. Zhao, 2007: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum - Part 1: experiments and large-scale features. *Climate of the Past*, **3**, 261–277.

Brauer, A., C. Endres, C. Günter, T. Litt, M. Stebich, and J. W. F. Negendank, 1999: High resolution sediment and vegetation responses to Younger Dryas climate change in varved lake sediments from Meerfelder Maar, Germany. *Quaternary Science Reviews*, **18**, 321–329.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research*, **111**, doi:doi:10.1029/2005JD006548.

Carrión, J. S., M. Munuera, M. Dupré, and A. Andrade, 2002: Abrupt vegetation changes in the Segura Mountains of southern Spain throughout the Holocene. *Journal of Ecology*, **89**, 783–797.

Cheddadi, R., G. Yu, J. Guiot, S. P. Harrison, and I. C. Prentice, 1997: The climate of Europe 6000 years ago. *Climate Dynamics*, **13**, 1–9.

Clark, P. U., S. J. Marshall, G. K. C. Clarke, S. W. Hostetler, J. M. Licciardi, and J. T. Teller, 2001: Freshwater Forcing of Abrupt Climate Change During the Last Glaciation. *Science*, **293**, 283–287.

Collins, M., 2007: Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society A*, **365**, 1957–1970.

Cook, E. R., R. D. D'Arrigo, and M. E. Mann, 2001: A Well-Verified, Multiproxy Reconstruction of the Winter North Atlantic Oscillation Index since A.D. 1400. *Journal of Climate*, **15**, 1754–1764.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational analysis (3D-VAR). I: Formulation. *Quarterly Journal of the Royal Meteorological Society*, **124**, 1783–1807.

Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. 11: Numerical results. *Quarterly Journal of the Royal Meteorological Society*, **113**, 1329–1347.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-VAR, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, **120**, 1367–1387.

Dahl-Jensen, D., K. Mosegaard, N. Gundestrup, G. D. Clow, S. J. Johnsen, A. W. Hansen, and N. Balling, 1998: Past Temperatures Directly from the Greenland Ice Sheet. *Science*, **282**, 268–271.

Davis, B. A. S., S. Brewer, A. C. Stevenson, and J. Guiot, 2003: The temperature of Europe during the Holocene reconstructed from pollen data. *Quaternary Science Reviews*, **22**, 1701–1716.

Debret, M., V. Bout-Roumazeilles, F. Grousset, M. Desmet, J. F. McManus, N. Massei, D. Sebag, J.-R. Petit, Y. Copard, and A. Trentesaux, 2007: The origin of the 1500-year climate cycles in Holocene North-Atlantic records. *Climate of the Past*, **3**, 569–575.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, **B39**, 1–38.

Digerfeldt, G., 1977: The Flandrian development of Lake Flarken. Regional vegetation history and palaeolimnology. Technical Report 13, University of Lund, Department of Quaternary Geology.

— 1982: The Holocene development of Lake Sambosjon. 1. The regional vegetation history. *Lundqua Report*, **23**, 1–24.

Digerfeldt, G., J.-L. de Beaulieu, J. Guiot, and J. Mouthon, 1997: Reconstruction and palaeoclimatic interpretation of the Holocene lake-level changes in Lac de Saint-Léger, Haute-Provence, southeast France. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **136**, 231–258.

Ditlevsen, P., M. Kristensen, and K. Andersen, 2005: The Recurrence Time of Dansgaard-Oeschger Events and Limits on the Possible Periodic Component. *Journal of Climate*, **18**, 2594–2603.

Ditlevsen, P. D., K. K. Andersen, and A. Svensson, 2007: The DO-climate events are probably noise induced: statistical investigation of the claimed 1470 years cycle. *Climate of the Past*, **3**, 129–134.

Easterling, D. R. and M. F. Wehner, 2009: Is the climate warming or cooling? *Geophysical Research Letters*, **36**, doi:10.1029/2009GL037810.

Egger, J., 1977: On the linear theory of the linear response to sea surface temperature anomalies. *Journal of the Atmospheric Sciences*, **34**, 603–614.

Elbern, H. and H. Schmidt, 2001: Ozone episode analysis by four-dimensional variational chemistry data assimiltaion. *Journal of Geophysical Research D*, **106**, 3569–3590.

EPICA Community Members, 2004: Eight glacial cycles from an Antarctic ice core. *Nature*, **429**, 623–628.

Eskes, H. J., A. J. M. Piters, P. F. Levelt, M. A. F. Allaart, and H. N. Kelder, 1999: Variational assimilation of GOME total-column ozone satellite data in a 2D latitude-longitude tracer-transport model. *Journal of the Atmospheric Sciences*, **56**, 3560–3572.

Fairbanks, R. G., 1990: The Age and Origin of the "Younger Dryas Climate Event" in Greenland Ice Cores. *Paleoceanography*, **5**, 937–948.

Farrera, I., S. P. Harrison, I. C. Prentice, G. Ramstein, J. Guiot, P. J. Bartlein, R. Bonnefille, M. Bush, W. Cramer, U. von Grafenstein, K. Holmgren, H. Hooghiemstra, G. Hope, D. Jolly, S.-E. Lauritzen, Y. Ono, S. Pinot, M. Stute, and G. Yu, 1999: Tropical climates at the Last Glacial Maximum: a new synthesis of terrestrial palaeoclimate data. I. Vegetation, lake-levels and geochemistry. *Climate Dynamics*, **15**, 823–856.

Fraley, C. and A. Raftery, 2002: Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, **97**, 611–631.

— 2007: Bayesian Regularization for Normal Mixture Estimation and Model-Based Clustering. *Journal of Classification*, **24**, 155–181.

Früchtl, M. K., 1998: *Beiträge zur spätglazialen Florenentwicklung Zentralmitteleuropas aufgrund pflanzlicher Großreste (unter besonderer Berücksichtigung von Großrestfloren aus der Hitsche (Eifel) und aus Krumpa (Geiseltal))*. Ph.D. thesis, Universität Bonn.

Gaillard, M.-J., 1984: Evolution de la Végétation du Moyen-Pays Romand. *Dissertationes Botanicae*, **77**.

— 1985: Late-glacial and Holocene environments of some ancient lakes in the western Swiss Plateau. *Dissertationes Botanicae*, **87**, 273–336.

Ganopolski, A. and S. Rahmstorf, 2001: Rapid changes of glacial climate simulated in a coupled climate model. *Nature*, **409**, 153–158.

Gauthier, P., 1992: Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus*, **44A**, 2–17.

Gauthier, P., C. Charette, C. Fillion, P. Koclas, and S. Laroche, 1999: Implementation of a 3D variational data assimilation system at the Canadian Centre. Part I: The global analysis. *Atmosphere-Ocean*, **37**, 103–156.

Gebhardt, C., 2003: *Variational Reconstruction of Quaternary Temperature Fields using Mixture Models as Botanical-Climatological Transfer Fuctions*. Ph.D. thesis, Meteorologisches Institut der Universität Bonn.

Gebhardt, C. and A. Hense, 2001: Statistical-dynamical analysis of wind observations using a variational approach. *Theoretical and Applied Climatology*, **68**, 137–154.

Gebhardt, C., N. Kühl, A. Hense, and Litt. T., 2008: Reconstruction of Quaternary temperature fields by dynamically consistent smoothing. *Climate Dynamics*, **30**, 421–437.

Gentle, J. E., 2003: *Random Number Generation and Monte Carlo Methods. Statistics and Computing.*. Springer-Verlag, New York.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, 1996: *Markov chain Monte Carlo in practice.*. Chapman & Hall, London, UK.

Gisecke, T., 2000: Pollenanalytische und sedimentchemische Untersuchungen zur natürlichen und anthropogenen Geschichte im Schlaubetal. *Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin*, **39**, 89–112.

— 2005a: Holocene dynamics of the southern boreal forest in Sweden. *The Holocene*, **15**, 858–872.

— 2005b: Holocene forest development in the central Scandes Mountains, Sweden. *Vegetation History and Archaeobotany*, **14**, 133–147.

Godwin, H. and P. A. Tallantire, 1951: Studies of the Post-glacial history of British vegetation XII. Hockam Mere, Norfolk. *Journal of Ecology*, **39**, 185–307.

Golub, G. and W. Kahan, 1965: Calculating the Singular Values and Pseudo-Inverse of a Matrix. *Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis*, **2**, 205–224.

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*, M. Neishtadt, ed., 41–57.

Guiot, J. and M. Couteaux, 1992: Quantitative climate reconstruction from pollen data in the Grand Duchy of Luxembourg since 15 000 yr BP. *Journal of Quaternary Science*, **7**, 303–309.

Harmata, K., 1987: Late-Glacial and Holocene history of vegetation at Roztoki and Tarnowiec near Jaslo (Jaslo-Sanok Depression). *Acta Palaeobotanica*, **27**, 43–65.

Heegard, E., H. Birks, and R. Telford, 2005: Relationships between calibrated ages and depth in stratigraphical sequences: an estimation procedure by mixed-effect regression. *The Holocene*, **15**, 612–618.

Heikkilä, M. and H. Seppä, 2003: A 11,000 yr palaeotemperature reconstruction from the southern boreal zone in Finland. *Quaternary Science Reviews*, **22**, 541–554.

Heinz, C. and M. Barbaza, 1998: Environmental changes during the Late Glacial and Post-Glacial in the central Pyrenees (France): new charcoal analysis and archaeological data. *Review of Palaeobotany and Palynology*, **104**, 1–17.

Hense, A., R. Glowienka-Hense, H. von Storch, and U. Stähler, 1990: Northern hemisphere atmospheric response to changes of atlantic ocean sst on decadal timescales: a gcm experiment. *Climate Dynamics*, **4**, 157–174.

Holton, J. R., 1992: *An Introduction to Dynamic Meteorology*. Academic Press, third edition edition.

Hu, X. and L. Xu, 2004: Investigation on Several Model Selection Criteria for Determining the Number of Cluster. *Neural Information Processing - Letters and Reviews*, **4**, 1–10.

Hultén, E., 1964: *The circumpolar plants, Vol. I. Vascular cryptogams, conifers, monocotyledons.*. Almkvist & Wiksell, Stockholm.

— 1971: *The circumpolar plants, Vol. II. Dicotyledons*.

I., S. and B. Ammann, 2003: Lateglacial and Holocene vegetation belts in the Pirin Mountains (southwestern Bulgaria). *The Holocene*, **13**, 97–107.

Isarin, R. F. B., H. Renssen, and J. Vandenberghe, 1998: The impact of the north atlantic ocean on the younger dryas climate in northwestern and central europe. *Journal of Quaternary Science*, **13**, 447–453.

Iversen, J., 1944: Viscum, Hedera and Ilex as climate indicators. *Geologiska Foereningens Stockholm Foerhandlingar*, **66**, 463–483.

Jahns, S. and C. van den Bogaard, 1998: New palynological and tephrostratigraphical investigations of two salt lagoons on the island of Mljet, south Dalmatia, Croatia. *Vegetation History and Archaeobotany*, **7**, 219–234.

Jalas, J. and J. Suominen, 1973: *Atlas Florae Europaeae. Distribution of vascular Plants in Europe, Vol. 2: Gymnospermae*.

— 1976: *Atlas Florae Europaeae. Distribution of vascular Plants in Europe, Vol. 3: Salicaceae to Balanophoraceae.*.

— 1989: Atlas Florae Europaeae. Distribution of vascular Plants in Europe, Vol. 8: Nymphaceae to Ranunculaceae.

Jankovska, V., 1980: *Paläogeobotanische Rekonstruktion der Vegetationsentwicklung im Becken Trebonska panev während des Spätglazials und Holozäns*, volume A 11 of *Vegetace CSSR*. Praha.

Jansen, E., J. Overpeck, K. R. Briffa, J.-C. Duplessy, F. Joos, V. Masson-Delmotte, D. Olago, B. Otto-Bliesner, W. R. Peltier, S. Rahmstorf, R. Ramesh, D. Raynaud, D. Rind, O. Solomina, R. Villalba, and D. Zhang, 2007: *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., chapter Paleoclimate.

Janssen, C. R., M. J. J. Cup-Uiterwijk, H. J. Edelman, J. Mekel-te Riele, and J. P. Pals, 1975: Ecologic and paleoecologic studies in the Feigne d'Artimont (Vosges, France). *Plant Ecology*, **30**, 165–178.

Juvigné, E., 1977: La zone de dispersion des poussières émises par une des dernières eruption du volcan du Laachersee (Eifel). *Zeitschrift für Geomorphologie*, **21**, 323–342.

Klaßen, J., A. Hense, and U. Römer, 1994: Climate anomalies north of 55° N associated with tropical climate extremes. *International Journal of Climatology*, **14**, 829–842.

Kleeman, R., 2008: Stochastic theories for the irregularity of ENSO. *Philosophical Transactions of the Royal Society A*, **366**, 2511–2526.

Koperowa, W., 1970: Late-Glacial and Holocene history of the vegetation of the eastern part of the "Jaslo-Sanok Doly" (Flysch Carpathians). *Acta Palaeobotanica*, **11**, 1–43.

Korhola, A., K. Vasko, H. Toivonen, and H. Olander, 2002: Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modelling. *Quaternary Science Reviews*, **21**, 1841–1860.

Kubitz, B., 2000: *Die holozäne Vegetations- und Siedlungsgeschichte in der Westeifel am Beispiel eines hochauflösenden Pollendiagrammes aus dem Meerfelder Maar*. Number 339 in Dissertationes Botanicae, Cramer, Stuttgart.

Kühl, N., 2002: *Die Bestimmung botanisch-klimatologischer Transferfunktionen und die Rekonstruktion des bodennahen Klimazustandesn in der Eem-Warmzeit*. Ph.D. thesis, University of Bonn.

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

Kühl, N., T. Litt, C. Schölzel, and A. Hense, 2007: Eemian and Early Weichselian temperature and precipitation variability in northern Germany. *Quaternary Science Reviews*, **26**, 3311–3317.

Kuhlemann, J., E. J. Rohling, I. Krumrei, P. Kubik, S. Ivy-Ochs, and M. Kucera, 2008: Regional Synthesis of Mediterranean Atmospheric Circulation During the Last Glacial Maximum. *Science*, **321**, 1338–1340.

Kuprijanowicz, M., 2000: Przemiany siedlisk I Roslinnosci torfowisk uroczyska Stare Biele w Puszczy Knyszynskiej. Poznoglacjalne i holocenskie zmiany roslinnosci okolic uroczyska. Technical Report 70, Politechnika Bialostocka.

Latalowa, M., 1982: Postglacial vegetational changes in the eastern baltic coastal zone of Poland. *Acta Palaeobotanica*, **XXII**, 179–249.

Latalowa, M. and R. K. Borowka, 2006: The Alleröd/Younger Dryas transition in Wolin Island, northwest Poland, as reflected by pollen, macrofossils, and chemical content of an organic layer separating two aeolian series. *Vegetation History and Archaeobotany*, **15**, 321–331.

Le Dimet, F. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**, 97–110.

Lee, Y., K. Y. Lee, and J. Lee, 2006: The Estimating Optimal Number of Gaussian Mixtures Based on Incremental k-means for Speaker Identification. *International Journal of Information Technology*, **12**, 13–21.

Leroyer, C. and C. Heinz, 1992: Complémentarité des études palynologiques et anthracologiques: les exemples pyrénéens de La Balma Margineda (Andorre) et de Belesta (Pyrénées-Orientales, France). *Bulletin de la Societe botanique de France*, **139**, 281–295.

Lewis, J. M., 2005: Roots of ensemble forecasting. *Monthly Weather Review*, **133**, 1865–1885.

Li, Z., I. M. Navon, and Y. Zhu, 2000: Performance of 4D-Var with Different Strategies for the Use of Adjoint Physics with the FSU Global Spectral Model. *Monthly Weather Review*, **128**, 668–688.

Liedberg, J. B., 1988: The Late Weichselian Macrofossil Flora in Western Skåne, Southern Sweden. *Lundqua Thesis*, **24**, 1–42.

Likas, A., N. Vlassis, and J. J. Verbeek, 2003: The global k-means clustering algorithm. *Pattern Recognition*, **36**, 451–461.

Liou, K. N., 2002: *An introduction to atmospheric radiation*, Academic Press, volume 84 of *International Geophysics Series*. 583pp.

Lisiecki, L. E. and M. E. Raymo, 2005: A Pliocene-Pleistocene stack of 57 globally distributed benthic D18O records. *Paleoceanography*, **20**, doi:10.1029/2004PA001071.

Litt, T., A. Brauer, T. Goslar, J. Merkt, K. Balaga, H. Müller, M. Ralska-Jasiewiczowa, M. Stebich, and J. Negendank, 2001: Correlation and synchronisation of Lateglacial continental sequences in northern central Europe based on annually laminated lacustrine sediments. *Quaternary Science Reviews*, **20**, 1233–1249.

Litt, T., C. Schölzel, N. Kühl, and A. Brauer, 2009: Vegetation and climate history in the Westeifel Volcanic Field (Germany) during the past 11000 years based on annually laminated lacustrine maar sediments. *Boreas*, doi:10.1111/j.1502-3885.2009.00096.x.

Lorenz, E. N., 1963: Deterministic non-periodic flow. *Journal of the Atmospheric Sciences*, **20**, 130–141.

Lotter, A. F., 1988: Paläoökologische und paläolimnologische Studie des Rotsees bei Luzern. Pollen-, großrest-, diatomeen- und sedimentanalytische Untersuchungen. *Dissertationes Botanicae*, **124**, 1–187.

— 1991: bsolute dating of the late-glacial period in Switzerland using annually laminated sediments. *Quaternary Research*, **35**, 321–330.

— 1997: Exkursionführer zur 21. Moorexkursion: Waldgrenzprobleme und hochauflösende Stratigraphien.

— 1999: Late-glacial and Holocene vegetation history and dynamics as shown by pollen and macrofossil analyses in anually laminated sediments from Soppensee, central Switzerland. *Vegetation History and Archaeobotany*, **8**, 165–184.

Lotter, A. F., H. J. B. Birks, U. Eichner, W. Hofmann, J. Schwander, and L. Wick, 2000: Younger Dryas and Alleroed summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **159**, 349–361.

Lotter, A. F., U. Eicher, U. Siegenthaler, and H. J. B. Birks, 1992: Late-glacial climatic oscillations as recorded in Swiss lake sediments. *Journal of Quaternary Science*, **7**, 187–204.

Loutre, M. F. and A. Berger, 2000: Future climate changes: Are we entering an exceptionally long interglacial? *Climatic Change*, **46**, 61–90.

Magny, M., J. Guiot, and P. Schoellammer, 2001: Quantitative Reconstruction of Younger Dryas to Mid-Holocene Paleoclimates at Le Locle, Swiss Jura, Using Pollen and Lake-Level Data. *Quaternary Research*, **56**, 170–180.

Mahalanobis, P. C., 1936: On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2**, 49–55.

Mangerud, J., 1970: Late Weichselian Vegetation and Ice-Front Oscillations in the Bergen District, Western Norway. *Norsk geografisk Tidsskrift*, **24**, 121–148.

Marek, S., 2000: Przemiany siedlisk I Roslinnosci torfowisk uroczyska Stare Biele w Puszczy Knyszynskiej. Analiza biostratygraficzna profili torfowych. *Politechnika Bialostocka*, 70–77.

Merkt, J. and H. Müller, 1999: Varve chronology and palynology of the Lateglacial in Northwest Germany from lacustrine sediments of Hämelsee in Lower Saxony. *Quaternary International*, **61**, 41–59.

Meusel, H. and E. Jäger, 1992: *Vergleichende Chorologie der zentraleuropäischen Flora*, volume Bd. III, Karten.

Meusel, H., E. Jäger, and E. Weinert, 1964: *Vergleichende Chorologie der zentraleuropäischen Flora*, volume Bd. I, Karten.

— 1974: *Vergleichende Chorologie der zentraleuropäischen Flora*, volume Bd. II, Karten.

Miotk-Szpiganowicz, G., 1992: The history of vegetation of Bory Tucholskie and the role of man in the lighht of palynological investigations. *Acta Palaeobotanica*, **32**, 39–122.

Muscheler, R., J. Beer, and P. Kubik, 2004: *Solar Variability and Its Effects on Climate*, American Geophysical Union, chapter Long-Term Solar Variability and Climate Change based on Radionuclide Data from Ice Cores. 221–235.

Neumann, F., C. Schölzel, T. Litt, A. Hense, and M. Stein, 2007: Holocene vegetation and climate history of the northern Golan heights (Near East). *Vegetation History and Archaeobotany*, **16**, 329–346.

New, M., M. Hulme, and P. Jones, 1999: Representing Twentieth-Century Space-Time Climate Variability. Part I: Development of a 1961-90 Mean Monthly Terrestrial Climatology. *Journal of Climate*, **12**, 829–856.

NGRIP-Members, 2005: High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature*, **431**, 147–151.

Nicol-Pichard, S. and M. Dubar, 1998: Reconstructions of late-glacial and holocene environments in south-east France based on the study of a 66-m long core from Biot, Alpes Maritimes. *Vegetation History and Archaeobotany*, **7**, 11–15.

Obidowicz, A., 1988: *Lateglacial and Holocene environment changes: Vistula Basin 1988. Excursion Guide Book*, Symposium, AGH, Cracow, Poland, chapter The Puscizna Rekowianska raised bog. 87–90.

— 1989: Type Region P-a: Inner West Carpathians- Nowy Targ Basin. *Acta Palaeobotanica*, **29**, 11–15.

— 1990: Eine pollenanalytische und moorkundliche Studie zur Vegetationsgeschichte des Podhale-Gebietes (West-Karpaten). *Acta Palaeobotanica*, **30**, 147–219.

— 1993: Wahania gornej granicy lasu w poznym plejstocenie i holocenie w Tatrach [Fluctuations of the forest timberline in the Tatra Mountains during the last 12 000 years]. *Dokumentacja Geograficzna*, **4-5**, 31–42.

Oeggl, K., 1992: Sediment- und Makrofossilienanalysen aus dem Lanser See in Tirol (Austria): Ein Beitrag zur spätglazialen Bio- und Chronostratigraphie der Ostalpen. *Flora*, **186**, 317–339.

Ogg, J. G., G. Ogg, and F. M. Gradstein, 2008: *The Concise Geologic Time Scale*. Cambridge University Press.

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

Pantaléon-Cano, J., E.-I. Yll, R. Pérez-Obiol, and J. M. Roure, 2003: Palynological evidence for vegetational history in semi-arid areas of the western Mediterranean (Almería, Spain). *The Holocene*, **13**, 109–119.

Paus, A., 1989: Late Weichselian vegetation, climate, and floral migration at Eigebakken, South Rogaland, southwestern Norway. *Review of Palaeobotany and Palynology*, **61**, 177–203.

— 1990: Late Weichselian and early Holocene vegetation, climate, and floral migration at Utsira, North Rogaland, southwestern Norway. *Norsk Geologisk Tidsskrift*, **70**, 135–152.

Peixoto, J. and A. H. Oort, 1992: *Physics of the Climate*, American Institute of Physics. 520pp.

Pernkopf, F. and D. Bouchaffra, 2005: Genetic-Based EM Algorithm for Learning Gaussian Mixture Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 1344–1348.

Pirrus, R., A.-M. Rõuk, and A. Liiva, 1987: *Palaeohydrology of the temperate zone II. Lakes*, Valgus, Tallin, Estonia, chapter Geology and stratigraphy of the reference site of Lake Raigastvere in Saadjärv drumlin field. 101–122.

Pokorny, P., 2002: A high-resolution record of Late-Glacial and Early-Holocene climatic and environmental change in the Czech Republic. *Quaternary International*, **91**, 101–122.

Pokorny, P. and V. Jankovská, 2000: Long-term vegetation dynamics and the infilling process of a former lake (Svarcenberk, Czech Republic). *Folia Geobotanica*, **35**, 433–457.

Prentice, I. C., W. Cramer, S. P. Harrison, R. Leemans, R. A. Monserud, and A. M. Solomon, 1992: A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, **19**, 117–134.

Prentice, I. C., J. Guiot, B. Huntley, D. Jolly, and R. Cheddadi, 1996: Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. *Climate Dynamics*, **12**, 185–194.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical recipes in FORTRAN*, Cambridge University Press. 963pp.

Rabier, F. and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Quarterly Journal of the Royal Meteorological Society*, **118**, 649–672.

Ralska-Jasiewiczowa, M., D. Demske, and B. van Geel, 1998: *Lake Gosciaz, Central Poland. A Monographic Study*, Krakow, chapter Late-Glacial vegetation history recorded in the Lake Gosciaz sediments.

Ramstein, G., M. Kageyama, J. Guiot, H. Wu, C. Hély, G. Krinner, and S. Brewer, 2007: How cold was Europe at the Last Glacial Maximum? A synthesis of the progress achieved since the first PMIP model-data comparison. *Climate of the Past*, **3**, 331–339.

Rankama, T. and I. Vuorela, 1988: Between inland and coast in Metal Age Finland - human impact on the premeval forests of Southern Hame during the Iron Age. *Memoranda Societatis Pro Fauna et Flora Fennica*, **64**, 25–34.

Reille, M., 1993: L'interface Tardiglaciaire-Holocène dans un site du littoral atlantique sud-européen: le Moura (Pyrénées Atlantiques, France). *Comptes rendus de l'Académie des sciences*, **316**, 463–468.

Reille, M. and V. Andrieu, 1995: The late Pleistocene and Holocene in the Lourdes Basin, Western Pyrenees, France: new pollen analytical and chronological data. *Vegetation History and Archaeobotany*, **4**, 1–21.

Reille, M. and J. L. de Beaulieu, 1988: History of the Würm and Holocene vegetation in western Velay (Massif Central, France): a comparison of pollen analysis from three corings at Lac du Bouchet. *Review of Palaeobotany and Palynology*, **54**, 233–248.

Renssen, H., R. F. B. Isarin, D. Jacob, R. Podzun, and R. J. Vandenberghe, 2001: Simulation of the Younger Dryas climate in Europe using a regional climate model nested in an AGCM: preliminary results. *Global and Planetary Change*, **30**, 41–57.

Robertson, I., D. Lucy, L. Baxter, A. M. Pollard, R. G. Aykroyd, A. C. Barker, A. H. C. Carter, V. R. Switsur, and J. S. Waterhouse, 1999: A kernel-based Bayesian approach to climatic reconstruction. *The Holocene*, **9**, 495–500.

Robinson, D. E. and J. H. Dickson, 1988: Vegetational history and land use: a radiocarbon-dated pollen diagram from Machrie Moor, Arran, Scotland. *New Phytologist*, **109**, 223–251.

Rösch, M., 1983: Geschichte der Nußbaumer Seen (Kanton Thurgau) und ihrer Umgebung seit dem Ausgang der letzten Eiszeit aufgrund quartärbotanischer, stratigraphischer und sedimentologischer Untersuchungen. *Mitteilungen der Thurgauischen Naturforscher Gesellschaft*, **45**, 1–110.

— 1985: *Swiss lake and mire environments during the last 15000 years*, Dissertationes Botanicae, chapter Nussbaumer Seen - Spät- und postglaziale Umweltveränderungen einer Seengruppe im östlichen Schweizer Mittelland. Number 87, 337–339.

— 1989: Pollenprofil Breitnau-Neuhof: Zum zeitlichen Verlauf der holozänen Vegetationsentwicklung im südlichen Schwarzwald. *Carolinea*, **47**, 15–24.

— 2000: Long-term human impact as registered in an upland pollen profile from the southern Black Forest, southwestern Germany. *Vegetation History and Archaeobotany*, **9**, 205–218.

Rybnickova, E. and K. Rybnicek, 1985: Palaeogeobotanical Evaluation of the Holocene Profile from the Rezabinec Fish-pond. *Folia Geobotanica et Phytotaxonomica*, **20**, 419–437.

Sánchez, G. M. F. and G. E. Hannon, 1999: High-altitude vegetational pattern on the Iberian Mountain chain (north-central Spain) during the Holocene. *The Holocene*, **9**, 39–57.

Sasaki, Y., 1958: An Objective Analysis Based on the Variational Method. *Journal of the Meteorological Society of Japan*, **36**, 77–88.

— 1970: Some basic formalisms in numerical variational analysis. *Monthly Weather Review*, **98**, 875–883.

Schloss, S., 1979: *Pollenanalytische und stratigraphische Untersuchungen im Sewensee. Ein Beitrag zur spät- und postglazialen Vegetationsgeschichte der Südvogesen*. Number 52 in Dissertationes Botanicae.

Schmidt, R., J. Müller, R. Drescher-Schneider, R. Krisai, K. Szerocynska, and A. Baric, 2000: hanges in lake level and trophy at Lake Vrana, a large karstic lake on the Island of Cres (Croatia), with respect to palaeoclimate and anthropogenic impacts during the last approx. 16,000 years. *Journal of Limnology*, **59**, 113–130.

Schneider, R. and K. Tobolski, 1983: alynologische und stratigraphische Untersuchungen im Lago di Ganna (Varese, Italien). *Botanica Helvetica*, **93**, 115–122.

Schölzel, C., 2005: *Palaeoenvironmental Transfer Functions in a Bayesian Framework with Application to Holocene Climate Variability in the Near East*. Ph.D. thesis, Meteorologisches Institut der Universität Bonn.

Schölzel, C. and P. Friederichs, 2008: Multivariate non-normally distributed random variables in climate research - introduction to the copula approach. *Nonlinear Processes in Geophysics*, **15**, 761–772.

Schölzel, C., A. Hense, P. Hübl, N. Kühl, and T. Litt, 2002: Digitization and geo-referencing of botanical distribution maps. *Journal of Biogeography*, **29**, 851–856.

Seppä, H. and H. J. B. Birks, 2001: July mean temperature and annual precipitation trends during the Holocene in the Fennoscandia tree-line area: pollen based climate reconstruction. *The Holocene*, **11**, 527–539.

— 2002: Holocene Climate Reconstructions from the Fennoscandian Tree-Line Area Based on Pollen Data from Toskaljavri. *Quaternary Research*, **57**, 191–199.

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

Seppä, H., D. Hammarlund, and K. Antonsson, 2005: Low and high-frequency changes in temperature and effective humidity during the Holocene in South central Sweden: Implications for atmospheric and oceanic forcings of climate. *Climate Dynamics*, **25**, 285–297.

Smith, A. G. and I. C. Goddard, 1991: A 12500 years record of vegetational history at Sluggan Bog, County Antrim, Northern Ireland (incorporating a pollen zone scheme for the non-specalist). *New Phytologist*, **118**, 167–187.

Smyth, P., 2000: Model selection for probabilistic clustering using cross-validated likelihood. *Statistics and Computing*, **10**, 63–72.

Smyth, P., K. Ide, and M. Ghil, 1999: Multiple Regimes in Northern Hemisphere Height Fields via Mixture Model Clustering. *Journal of the Atmospheric Sciences*, **56**, 3704–3723.

Srodon, A., 1983: The history of fir in Poland (Jodla pospolita w historii naszych Lasow).- Nasze drezewa Lesne. Monografie Popularnonaukowe 4, Polska Akademia Nauk, Instytut Dendrologii.

— 1990: Postglacial history of the common spruce (Picea excelsa (Lam.) Lk.) in the low Beskids separating the east from the west Carpathians. *Acta Palaeobotanica*, **30**, 221–226.

Stauffer, D. R. and N. L. Seaman, 1994: Multiscale Four-Dimensional Data Assimilation. *Journal of Applied Meteorology*, **33**, 416–434.

Stebich, M., 1999: *Palynologische Untersuchungen zur Vegetationsgeschichte des Weichsel-Spätglazial und Frühholozän an jährlich geschichteten Sedimenten des Meerfelder Maares (Eifel)*. Number 320, Cramer, Stuttgart.

Stewart, D., A. Walker, and J. H. Dickson, 1984: Pollen diagrams from Dubh Lochan, near Loch Lomond. *New Phytologist*, **98**, 531–549.

Stocker, T. F., D. G. Wright, and W. S. Broecker, 1992: The Influence of High-Latitude Surface Forcing on the Global Thermohaline Circulation. *Paleoceanography*, **7**, 529–541.

Swanson, K. L. and A. A. Tsonis, 2009: Has the climate recently shifted? *Geophysical Research Letters*, **36**, doi:10.1029/2008GL037022.

Switsur, V. R. and R. G. West, 1973: University of Cambridge natural Radiocarbon measurements XI. *Radiocarbon*, **15**, 534–544.

Szczepanek, K., 1971: The Staszow karst in the light of palaeobotanical studies (South Poland). *Acta Palaeobotanica*, **12**, 64–140.

Tarasov, P. E., J. Guiot, R. Cheddadi, A. A. Andreev, L. G. Bezusko, T. A. Blyakharchuk, N. I. Dorofeyuk, L. V. Filimonova, V. S. Volkova, and V. P. Zernitskaya, 1999: Climate in northern Eurasia 6000 years ago reconstructed from pollen data. *Earth and Planetary Science Letters*, **171**, 635–645.

Telford, R., E. Heegard, and H. Birks, 2004a: All age-depth models are wrong: but how badly? *Quaternary Science Reviews*, **23**, 1–5.

— 2004b: The intercept is a poor estimate of a calibrated radiocarbon age. *The Holocene*, **14**, 296–298.

Thacker, W. C. and R. B. Long, 1988: Fitting Dynamics to Data. *Journal of Geophysical Research C*, **93**, 1227–1240.

Thépaut, J.-N. and P. Courtier, 1991: Four-dimensional data assimilation using the adjoint of a multi-level primitive-equation model. *Quaterly Journal of the Royal Meteorological Society*, **117**, 1225–1254.

Titterington, D. M., A. F. M. Smith, and U. E. Makov, 1985: *Statistical Analysis of finite mixture distributions*. John Wiley & Sons.

Tobolski, K., 1990: Paläoökologische Untersuchungen des Siedlungsgebietes im Lednica Landschaftspark (Nordwestpolen). *Offa*, **47**, 109–131.

Toivonen, H., H. Mannila, A. Korhola, and H. Olander, 2001: Applying Bayesian Statistics to Organism-Based Environmental Reconstruction. *Ecological Applications*, **11**, 618–630.

Tonkov, S., H. Panovska, G. Possnert, and E. Bozilova, 2002: The Holocene vegetation history of Northern Pirin Mountain, southwestern Bulgaria: pollen analysis and radiocarbon dating of a core from Lake Ribno Ban derishko. *The Holocene*, **12**, 201–210.

Trouet, V., J. Esper, N. E. Graham, A. Baker, J. D. Scourse, and D. C. Frank, 2009: Persistent Positive North Atlantic Oscillation Mode Dominated the Medieval Climate Anomaly. *Science*, **324**, 78–80.

Vasko, K., H. T. T. Toivonen, and A. Korhola, 2000: A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. *Journal of Paleolimnology*, **24**, 243–250.

Voigt, R., 1996: Paläolimnologische und vegetationsgeschichtliche Untersuchungen an Sedimenten aus Fuschlsee und Chiemsee. *Dissertationes Botanicae*, **270**.

von Storch, H., E. Zorita, J. M. Jones, Y. Dimitriev, F. González-Rouco, and S. F. B. Tett, 2004: Reconstructing Past Climate from Noisy Data. *Science*, **308**, 679–682.

von Storch, H., E. Zorita, J. M. Jones, F. Gonzalez-Rouco, and S. F. B. Tett, 2006: Response to Comment on "Reconstructing Past Climate from Noisy Data". *Science*, **312**, 529c.

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press.

Vrac, M., A. Chedin, and E. Diday, 2005: Clustering a Global Field of Atmospheric Profiles by Mixture Decomposition of Copulas. *Journal of Atmospheric and Oceanic Technology*, **22**, 1445–1459.

Watson, C. S., 1996: The vegetational history of the northern Apennines, Italy: information from three new sequences and a review of Regional vegetational change. *Journal of Biogeography*, **23**, 805–841.

Watts, W. A., J. R. M. Allen, and B. Huntley, 1996a: Vegetation History and Palaeoclimate of the last Glacial Period at Lago Grande di Monticchio, Southern Italy. *Quaternary Science Reviews*, **15**, 133–153.

Watts, W. A., J. R. M. Allen, B. Huntley, and S. C. Fritz, 1996b: Vegetation History and Climate of the last 15000 years at Lago di Monticchio, Southern Italy. *Quaternary Science Reviews*, **15**, 113–132.

Webster, P. J., 1981: Mechanism determing the atmospheric response to sea surface temperature anomalies. *Journal of the Atmospheric Sciences*, **38**, 554–571.

Wegmüiller, H. P., 1976: Vegetationsgeschichtliche Untersuchungen in den Thuralpen und im Faningebiet (Kanton Appenzell,St. Gallen, Graubtinden/Schweiz). *Botanische Jahrbücher für Systematik*, **97**, 226–307.

Woillard, G., 1975: Recherches palynologique sur le pleistocène dans l'est de la Belgique et dans les Vosges lorraines. *Acta Geogr. Lovaniensia*, **14**, 1–117.

— 1978: The last interglacial-glacial cycle at Grand Pile in the Northeastern France. manuscript.

Wu, H., J. Guiot, and Brewer. S., 2007: Climatic changes in Eurasia and Africa at the last glacial maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling. *Climate Dynamics*, **29**, 211–229.

Zagwijn, W. H., 1996: An Analysis of Eemian Climate in Western and Central Europe. *Quaternary Science Rewievs*, **15**, 451–469.

Zhu, J. and M. Kamachi, 2000: An adaptive variational method for data assimilation with imperfect models. *Tellus*, **52A**, 265–279.

Zoller, H. and H. Kleiber, 1971: Vegetationsgeschichtliche Untersuchungen in der montanen und subalpinen Stufe der Tessintäler. *Verhandlungen der Naturforschenden Gesellschaft in Basel*, **81**, 1–156.

Zwahlen, R., 1985: *Swiss lake and environments during the last 15000 years*, Dissertationes Botanicae, chapter Lörmoos - Late-glacial and Holocene environments of an ancient lake on the Central Swiss Plateau. Number 87, 171–184.

Danksagung

Zunächst möchte ich natürlich meinen Betreuern Thomas Litt und Andreas Hense herzlich danken. Sie akzeptierten mich als Doktoranden, trauten mir zu, dieses Thema zu bearbeiten und unterstützten mich bestmöglich durch unzählige Ratschläge und Hilfestellungen. Weiterhin danke ich Christoph Gebhardt, Christian Schölzel und Norbert Kühl für etliche Hinweise und Gespräche, die zum Gelingen dieser Arbeit beitrugen. Ebenfalls einen großen Beitrag an dieser Arbeit haben Sabine Hettler und Nils Riedel, die für Beschaffung der Vegetationsdaten verantwortlich waren. Für das Korrekturlesen meiner Arbeit danke ich Isabel, Annika, René, Henning und Matthieu.
Viele Menschen am Meteorologischen und Paläontologischen Institut waren nette Kollegen oder wurden im Laufe der Zeit zu guten Freunden. Hier sind natürlich meine langjährigen Bürokollegen René, Henning, Marco, Sabine und Nils zu nennen, aber auch Steffi, Susanne, Insa, Isabel, Thomas, Elham, Michael und nicht zuletzt natürlich Kerstin, Annika, Christoph, Robin, Timo und Matthieu. Ihr habt mich mit Süßigkeiten und Kaffee versorgt, mit mir am Kicker gestanden und natürlich auch das ein- oder andere Bier vernichtet. Dank Euch werde ich immer gerne an meine Uni-Zeit zurückdenken.
Natürlich danke ich auch meinen Freunden aus der "Grillgruppe" die vom ersten bis zum letzten Semester an meiner Seite standen und ohne die alles nur halb so schön gewesen wäre.
Mein gesamter Weg von der Schule über das Studium bis zur Promotion wäre natürlich nie möglich gewesen, ohne die Unterstützung, den Rückhalt und den Ansporn meiner Eltern Rosa-Maria und Wolfgang sowie meiner Geschwister Janina und Felix.
Nun bleibt noch ein Mensch dem ich garnicht genug danken kann: Verena Dir bin ich so unendlich dankbar für die vielen aufmunternden Worte und auch manchmal den nötigen Tritt in den Hintern, für die Geduld mit der Du meine 60-70 Stunden Wochen der letzten Monate ertragen hast und einfach dafür, dass Du der liebevollste Mensch bist, den ich mir an meiner Seite vorstellen kann.