

# Strengths and Weaknesses of Quantitative Climate Reconstructions Based on Late-Quaternary Biological Proxies

H. John B. Birks<sup>\*1</sup>, Oliver Heiri<sup>2</sup>, Heikki Seppä<sup>3</sup> and Anne E. Bjune<sup>4</sup>

<sup>1</sup>Department of Biology, University of Bergen, Thormøhlensgate 53A, NO-5006 Bergen, Norway; Bjerknes Centre for Climate Research, Allégaten 55, NO-5007 Bergen, Norway; Environmental Change Research Centre, University College London, London, WC1E 6BT, UK; and School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK

<sup>2</sup>Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland; Palaeoecology, Institute of Environmental Biology, Utrecht University, Budapestlaan 4, CD 3584 Utrecht, The Netherlands

<sup>3</sup>Department of Geosciences and Geography, University of Helsinki, Gustaf Hällströmin katu 2a, P.O. Box 64, FI-00014 Helsinki, Finland

<sup>4</sup>Bjerknes Centre for Climate Research, c/o Department of Biology, University of Bergen, Thormøhlensgate 53A, NO-5006 Bergen, Norway

**Abstract:** The importance of reconstructing past environments quantitatively in palaeoecology is reviewed by showing that many ecological questions asked of palaeoecological data commonly involve the reconstructions of past environment. Three basic approaches to reconstructing past climate from palaeoecological data are outlined and discussed in terms of their assumptions, strengths, and weaknesses. These approaches are the indicator-species approach involving bioclimate-envelope modelling; the assemblage approach involving modern analogue techniques and response surfaces; and the multivariate calibration-function approach. Topics common to all approaches are reviewed – presentation and interpretation, evaluation and validation, comparison, and general limitations of climate reconstructions. Challenges and possible future developments are presented and the potential future role of quantitative climate reconstructions in palaeoecology is summarised.

**Keywords:** Bayesian approaches, bioclimate-envelope modelling, climate, climate reconstruction, indicator species, modern analogue technique, mutual climate range methods, niche, palaeoecology, reconstruction evaluation, reconstruction presentation, reconstruction validation, response surfaces, spatial autocorrelation, weighted averaging, weighted averaging partial least squares.

## INTRODUCTION

Palaeoecology is the ecology of the past (Birks & Birks 1980) or as Rull (2010) proposes, it is “the branch of ecology that studies past ecological systems and their trends in time using fossils and other proxies”. Studies of ecological systems, past, present, or future, require consideration not only of their biotic components (e.g. species, populations, communities, biotic interactions, human activities) but also their abiotic components (e.g. climate, soil, topography, water chemistry, water temperature). If aspects of the abiotic component in a past ecosystem can be studied directly or reconstructed indirectly, these variables can be regarded as a set of predictors or forcing functions within the past ecological system under study. The biotic components of the system can thus be viewed as response variables. Important

ecological processes and properties such as ecosystem variability, rates and nature of biotic response to environmental change, base-line conditions, thresholds, resilience, and ecosystem novelty can then be deduced from the palaeoecological record (Willis *et al.* 2010a). These are all important topics in current biodiversity conservation research with its increasing focus on the need for the conservation of ecological and evolutionary processes in the face of global climate and related changes (Klein *et al.* 2009). Care must, of course, be taken to avoid circularity of argument where the abiotic aspects of a past ecological system are reconstructed from a set of biological variables, and the resulting reconstructions are then used to ‘explain’ the observed changes in the biological proxies used in the reconstructions. All reconstructions of the past environment must be independent of any response variables being used to study biotic responses to environmental change (Flessa & Jackson 2005a).

The development of quantitative techniques for inferring past environments from palaeoecological data (e.g. Birks 1995) and of multi-proxy studies in palaeoecology (Birks & Birks 2006) allows direct studies of biotic responses to

\*Address correspondence to this author at the Department of Biology, University of Bergen, PO Box 7803, NO-5020 Bergen, Norway; Tel: +44 5558 3350; E-mail: john.birks@bio.uib.no

Handling Editor: Dr. Valenti Rull

environmental change over a range of time scales in the past (Flessa & Jackson 2005a, 2005b). In multi-proxy studies, several palaeoecological variables are studied and one or a few are used as predictor variables to reconstruct major aspects of the past environment (e.g. climate, lake-water pH, catchment vegetation) and the other variables are analysed as response variables. This multi-proxy approach can allow the direct study of biotic responses to climate change in the past (e.g. Birks & Birks 2008) and the testing of hypotheses about, for example, the relative roles of climate and catchment variables in influencing within-lake biotic and abiotic records (e.g. Lotter & Birks 2003).

Palaeoecological records abound with examples of biotic responses to past environmental changes, but this wealth of information has rarely been exploited to answer ecological questions. Recent advances in palaeoecology (quantitative environmental reconstruction procedures; multi-proxy studies; improved fine-scale and fine-resolution sampling; improved understanding of the spatial scales of palaeoecological data; appropriate and powerful numerical techniques for data summarisation and synthesis; improved taxonomic resolution; improved chronological tools; increasing concern about site selection, problem formulation, and testable hypotheses; realisation that lake sediments are archives of a vast array of palaeoecological, biological, chemical, and physical variables; increasing realisation of the central importance of spatial and temporal scales in data interpretation – see, for example, Birks (1998), Seppä and Bennett (2003) Jackson and Williams (2004), Flessa and Jackson (2005a), Birks and Birks (2006), Smol (2008), and Jackson *et al.* (2009) for reviews of some of these advances) now allow palaeoecologists to explore many critical ecological questions, answers to which require the unique long-term temporal perspective provided by palaeoecology. What are the rates and nature of changes in ecological processes in response to environmental change? Which combination of abiotic and biotic processes results in a given ecological threshold being crossed? What ecological processes increase resilience to environmental change? Did novel ecosystems, so-called non-analogue systems, occur in the past? What combination of abiotic and biotic processes resulted in such novel ecosystems (Willis *et al.* 2010a)? Answers to these and related questions (Froyd & Willis 2008; MacDonald *et al.* 2008; Willis & Bhagwat 2010; Willis *et al.* 2010b) highlight the contributions that palaeoecology including environmental reconstructions can make not only in understanding ecological and evolutionary processes responsible for biodiversity patterns, but also in determining current and future management strategies necessary to ensure biodiversity conservation (Willis *et al.* 2010a).

Other ecological questions that palaeoecology can contribute insights to include the following. What are the major drivers of ecosystem dynamics at different time scales and do these drivers interact? How has climate change affected the distribution, magnitude, and frequency of disturbance regimes such as fires? What time lags have occurred between biotic response and climate change? Are there consistent patterns in time and/or space of population expansion, population extinction, and migration dynamics in relation to environmental change? Answers to all the questions raised here all require considerations of the past environment and

the rates, magnitude, and inherent variability of environmental change.

Reliable, robust, and unbiased high-resolution environmental reconstructions including climate are thus an integral part of applied palaeoecology (*Sensu* Birks 1996) and of ecological palaeoecology (MacDonald *et al.* 2008; Willis *et al.* 2010a) as many ecological and conservation questions asked of palaeoecological data require environmental reconstructions. Similarly the understanding of historical ‘ecological legacies’ (Flessa & Jackson 2005a, 2005b; Willis & Birks 2006) in present-day ecosystems requires records of past climate or land-use changes (Millar & Woolfenden 1999; Swetnam *et al.* 1999; Foster *et al.* 2003). For example, Heiri *et al.* (2006) elegantly illustrate the use of a chironomid-based quantitative climate reconstruction as a driver in a vegetation dynamics model to simulate tree-line fluctuations during the Holocene and to test hypotheses about the causes of tree-line changes.

This essay reviews the range of quantitative approaches available for inferring past climate from late-Quaternary palaeoecological records. Although we focus our discussion on climate reconstructions, many of the approaches considered are also used to reconstruct other environmental variables (e.g. lake-water pH and total phosphorus) from limnic organisms such as diatoms (Birks 1998, 2010). We make no attempt to provide a comprehensive review of the vast array of available palaeoclimatic reconstruction approaches. Instead we present a personal perspective on the strengths and weaknesses of the major approaches used to reconstruct past climate from biological proxies, with particular emphasis on pollen and chironomid assemblages, the two groups of fossil organisms that we have most practical experience of working with. We also discuss, where relevant, beetles, plant macrofossils, and other groups of fossils.

## BASIC PRINCIPLES AND ASSUMPTIONS

### Introduction

Since the pioneering days of Quaternary palaeoecology in the first half of the 20<sup>th</sup> century, scientists have used fossils preserved in lake and bog sediments and the composition and physical properties of these sediments as a basis for reconstructing aspects of the past local and regional environment. The early studies of Blytt (1881), Andersson (1909), Iversen (1944), von Post (1946), and others using plant macrofossils, animal remains, pollen assemblages, and peat stratigraphy laid the foundations for many of the current ideas about late-Quaternary climate changes (see Birks 2008; Birks & Seppä 2010 for historical reviews).

An important aim of many Quaternary palaeoecological studies today remains the reconstruction of features of the past environment, in particular climate, from fossil assemblages preserved in terrestrial or marine sediments to help address the types of question we raised in the Introduction. Although fossil assemblages are usually meticulously counted with individual fossil pollen, ostracods, chironomids, foraminifera, etc. being routinely identified and enumerated using considerable time and effort, the resulting environmental reconstructions are often qualitative and are presented as broad climate ranges which are interpreted as “arctic”,

“cool”, “temperate”, “continental”, “moist”, “dry”, etc (Porch 2010). There is, however, an ever-increasing need for quantification in Quaternary palaeoecology. This need is, in part, a response to demands for quantitative reconstructions of past climates as input to, or more commonly, for validation of hindcasts by general circulation or Earth system models of past climate or land-cover patterns (e.g. Braconnot *et al.* 2007a, 2007b; Renssen *et al.* 2009; Schmidt 2010), in mapping past climate patterns in time and space (e.g. Heiri *et al.* 2004; Morrill & Jacobsen 2005), in predicting the risk of future climate shifts (e.g. Enfield & Cid-Serrano 2006; Schmidt 2010), in integrating past climate reconstructions with historical societal changes (e.g. Caseldine & Turney 2010), and in providing insights into the rates of change that have characterised Earth’s past climate variability (e.g. Trenberth & Otto-Bliesner 2003; Caseldine *et al.* 2010). The need for quantification of environmental change is also, in part, related to demands for palaeoenvironmental reconstructions needed to answer ecological and palaeoecological questions concerning thresholds (e.g. Virah-Sawmy *et al.* 2009b), alternative stable states (e.g. Lytle 2005), resilience (e.g. Virah-Sawmy *et al.* 2009a), ecosystem variability (e.g. Hotchkiss *et al.* 2007; Tweiten *et al.* 2009), extinctions (e.g. Virah-Sawmy *et al.* 2009c, 2010; Willis & Bhagwat 2010; Willis *et al.* 2010b), lags in biotic responses (e.g. Birks & Birks 2008), baseline conditions (e.g. Willis *et al.* 2010a), and tree-population collapse (e.g. Calcote 2003).

In the last 30-40 years, quantitative methods for reconstructing climate (and other environmental variables) have been developed from a range of biological proxies such as pollen, plant macrofossils, insects (chironomids, coleopterans), molluscs, ostracods, diatoms, chrysophycean cysts, testate amoebae, and cladocerans preserved in lake sediments and peat profiles, or dinoflagellate cysts, diatoms, pollen, foraminifera, coccolithophores, and radiolarians preserved in marine sediment records. Records are available that quantitatively infer decadal- to centennial-scale climate variability from sedimentary deposits covering the instrumental record (e.g. Bigler *et al.* 2003; Larocque *et al.* 2009; Lamentowicz *et al.* 2010), the end of the last deglaciation (e.g. Coope *et al.* 1998; Birks & Ammann 2000, Brooks and Birks 2001; Heiri *et al.* 2007), the previous interglacial (e.g. Rioual *et al.* 2001; Kühl & Litt 2003; Klotz *et al.* 2003), and earlier parts of the Quaternary (e.g. Pross *et al.* 2000; Pross & Klotz 2002). In many cases these reconstructions have been validated by comparing them with independent climate records based on geochemical, isotope, or other biological data (e.g. Lotter *et al.* 2000; Heiri & Millet 2005; Peyron *et al.* 2005). For the pre-instrumental period (*ca.* 1800 AD) and especially for records from the mid- or early-Holocene, reconstruction methods based on biostratigraphical data are among the few approaches which can provide quantitative reconstructions that approach the temporal and spatial scales needed to allow the reconstructions to be compared with results from Earth system and general circulation models (e.g. Renssen & Isarin 2001; Renssen *et al.* 2009).

As with all methods for climate reconstruction, approaches for inferring past climate change based on biostratigraphical data may fail when important assumptions are violated that relate to the numerical approach used or to the ecological principles forming the basis of the reconstruction.

The quality of all quantitative reconstructions strongly depends on the quality of the data-sets used, on the temporal chronologies of the data-sets used, on the performance and robustness of the numerical approaches applied to reconstruct past climate, on the extent to which these approaches realistically quantify and model the relationship between the biostratigraphical proxies and their environment in the modern world and in the past, and on rigorous numerical, ecological, and palaeoenvironmental evaluations of the resulting reconstructions.

Here we provide a review of the three basic approaches most commonly used by palaeoecologists to infer past climate based on biostratigraphical data, namely the indicator-species approach; the assemblage or analogue approach; and the multivariate calibration-function (‘transfer function’) approach. Since the latter approach is the most widely used approach in the reconstruction of past climate based on late-Quaternary biostratigraphical data, we review and discuss it in more detail than the other two approaches.

We discuss the general principles, requirements, and assumptions of climate reconstructions based on these three approaches. We then discuss individual methods within these approaches and outline their strengths and weaknesses. We indicate how rigorous application of evaluation and validation techniques, estimation of realistic errors of prediction, and sound ecological reasoning can be used to detect and reduce some of the problems that can occur when using the general approach of reconstructing past climate from proxy biostratigraphical data. Our emphasis throughout is on the basic biological concepts, the assumptions, requirements, and data properties of the various approaches, and on their strengths and weaknesses. We also discuss how modern ecological and biogeographical studies on organism responses to future climate change involve the indicator-species approach, the assemblage approach, or the multivariate calibration-function approach. We outline how such predictive studies are facing many of the same problems that palaeoecologists are facing, namely the hidden assumptions and limitations of bioclimatic-envelope modelling, problems of spatial autocorrelation, and the robustness and applicability of multivariate calibration functions. We conclude by outlining new approaches to climate reconstruction involving forward modelling procedures and Bayesian approaches, and suggesting possible areas for future improvements in quantitative climate reconstructions based on biological proxy data.

We draw on previous reviews by Birks (1981, 1995, 1998, 2010), ter Braak (1995), Seppä and Bennett (2003), Birks and Seppä (2004), Kumke *et al.* (2004a), Brewer *et al.* (2007), Simpson (2007, 2011), and Juggins and Birks (2011). Mathematical details of the relevant numerical procedures are not presented here – the interested reader should consult ter Braak and Looman (1986), ter Braak and Barendregt (1986), ter Braak and van Dam (1989), Birks *et al.* (1990), ter Braak and Juggins (1993), ter Braak *et al.* (1993), Birks (1995), ter Braak (1995), Kumke *et al.* (2004b), and Guiot and de Vernal (2007).

### Principles, Notation, and Terminology

There are three main approaches to the quantitative reconstruction of past climates from biostratigraphical data

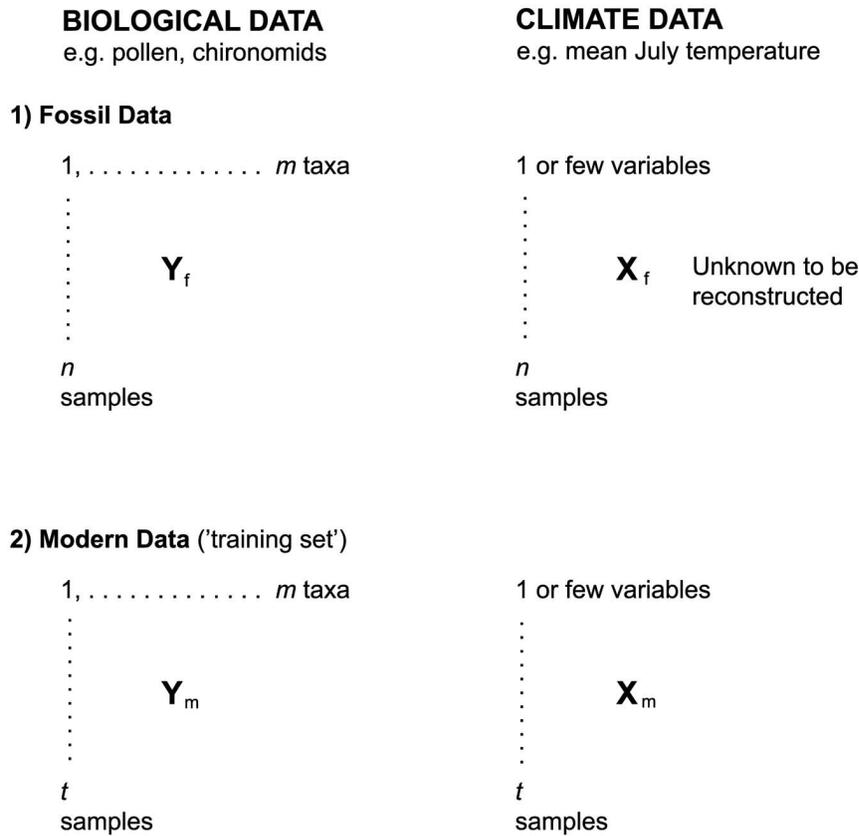


Fig. (1). Principles of quantitative palaeoclimatic reconstruction showing X<sub>f</sub>, the unknown past climate variable (e.g. mean July temperature) to be reconstructed from fossil assemblages Y<sub>f</sub> and the essential role of modern data consisting of modern biological data (presence/absence and/or abundances) (Y<sub>m</sub>) and associated climate data (X<sub>m</sub>).

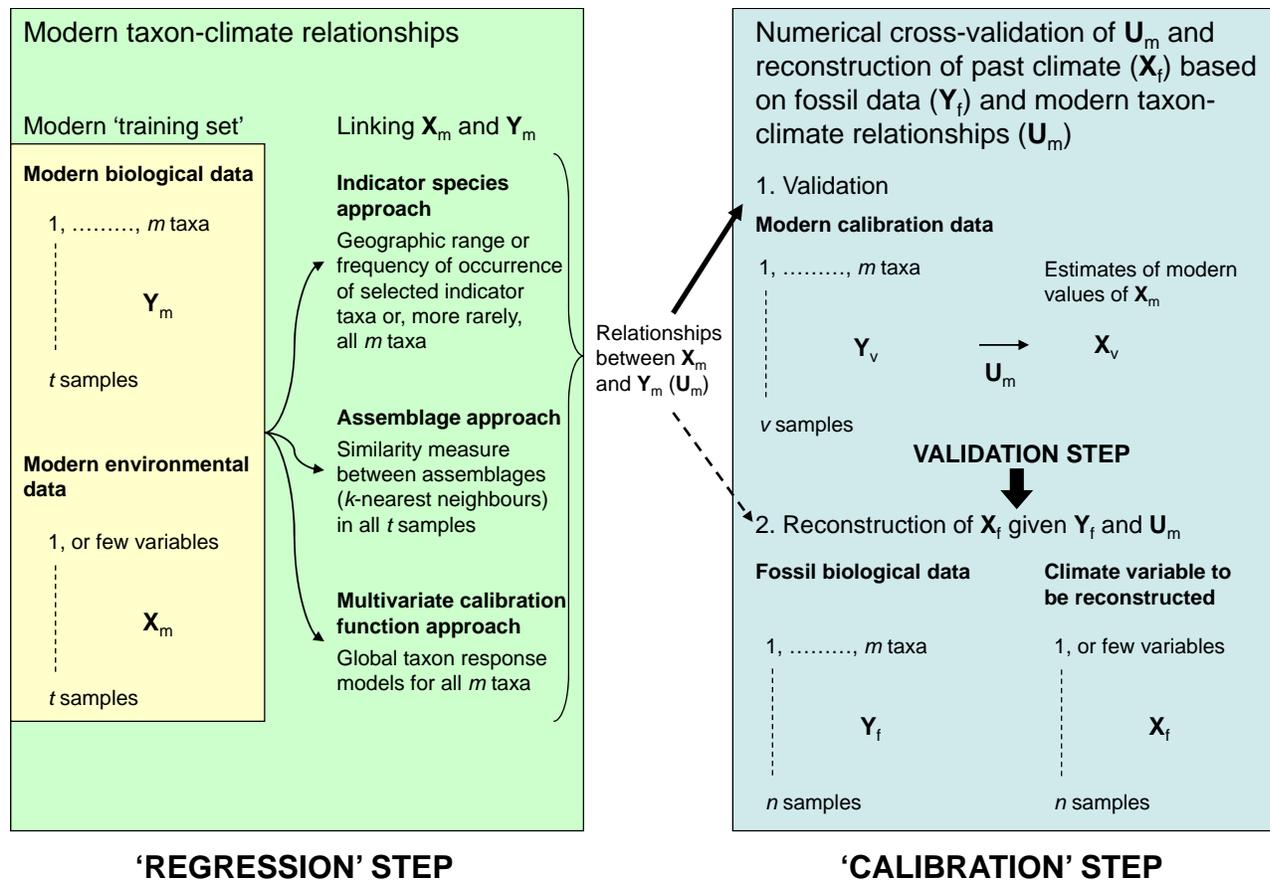
Table 1. Types of Modern Biological Data Used in the Three Major Approaches to Reconstructing Past Climates from Fossil Biological Data

Approach	Modern Biological Data Y <sub>m</sub>	Links with Modern Climate Data X <sub>m</sub>	Examples
Indicator-species approach	Presence/absence of one or few taxa (e.g. biogeographical distributions)	Visual Probability density functions	Iversen (1944) Conolly & Dahl (1970) Atkinson <i>et al.</i> (1987) Kühl <i>et al.</i> (2002)
Assemblage approach	Presence/absence of many taxa (e.g. biogeographical distributions) or modern quantitative assemblages at many sites	Modern analogue matching Response surfaces	Guiot (1990) Overpeck <i>et al.</i> (1985) Thompson <i>et al.</i> (2008) Bartlein <i>et al.</i> (1986) Huntley (1993) Prentice <i>et al.</i> (1991)
Multivariate calibration-function approach	Modern quantitative assemblages at many sites	Linear or non-linear regression and calibration	Imbrie & Kipp (1971) Birks <i>et al.</i> (1990) ter Braak & Juggins (1993)

(Birks & Birks 1980; Birks 1981, 1995; 1998; 2003). These are:

1. the indicator-species approach
2. the assemblage or analogue approach
3. the multivariate calibration-function approach.

All these approaches involve a space-for-time substitution and require information about the modern climatic preferences and tolerances of the taxa found as fossils. These data are usually obtained by exploring the distribution and abundance of organisms in relation to environmental variables in the modern world as an analogue to their expected



**Fig. (2).** The major stages in quantitative palaeoclimatic reconstruction of regression to derive from  $X_m$  and  $Y_m$  modern calibration functions  $U_m$  and calibration to estimate past climate  $X_f$  from fossil biological data. The additional stage of model validation involving test data or numerical cross-validation is also shown.

distribution and abundance in relation to the environmental variable of interest in the past (so-called 'space-for-time' substitution) (Jackson & Williams 2004). However, information on the response of organisms to environmental change can also be obtained from laboratory experiments (e.g. von Grafenstein *et al.* 1999; Brodersen *et al.* 2008) or by comparing variations in biotic assemblages during the instrumental period with time-series of measured changes in the environmental parameter (so-called 'calibration' in time or 'time-for-space' substitution) (e.g. Barnekow *et al.* 2007; Kamenik *et al.* 2009). The principle assumption of all climatic and all other environmental reconstructions in Quaternary palaeoecology is methodological uniformitarianism (Scott 1963; Gould 1965; Rymer 1978; Birks & Birks 1980; Huntley 2001), namely that modern-day observations and relationships can be used as a model for past conditions and, more specifically, that organism-environment relationships have not changed with time, at least in the late-Quaternary (Jackson & Overpeck 2000; Jackson & Williams 2004), so-called niche conservatism (Wiens & Graham 2005; Pearman *et al.* 2008a, 2008b; Bennett *et al.* 2010; Wiens *et al.* 2010).

The basic idea of all quantitative reconstructions and the notation we use throughout are summarised in Fig. (1). We have fossil biological data ( $Y_f$ ) (e.g. pollen or chironomid assemblages) at different depths or times in one or more

sediment sequences (so-called climate 'proxy' data) and we want to reconstruct the past environment ( $X_f$ ) (e.g. mean July temperature) for the times at which the fossil assemblages have been studied. Irrespective of the approach used to estimate or infer  $X_f$ , we need modern distributional and/or abundance data ( $Y_m$ ) about all or some of the species in the fossil assemblage and associated modern environmental data ( $X_m$ ) in order to relate the occurrence and/or abundance of  $Y_m$  to the modern environmental variable(s) ( $X_m$ ) that we are wanting to reconstruct ( $X_f$ ) from our fossil assemblages ( $Y_f$ ). The most commonly used forms of modern biological data ( $Y_m$ ) in the three main approaches are summarised in Table 1, along with how the contemporary relationships with the modern environment ( $X_m$ ) are established prior to reconstruction of  $X_f$  from  $Y_f$ .

Although  $X_f$  can be estimated or inferred in many different ways, in practice nearly all quantitative reconstructions involve two main stages (Fig. 2). First, the responses ( $U_m$ ) of the modern taxa ( $Y_m$ ) to contemporary climate ( $X_m$ ) are established. This is essentially a regression problem (ter Braak & Prentice 1988) and involves modern distributional data or a modern 'training set' of assemblages ('response' variables) with associated climate data ('predictor' variables). Second, the estimated modern responses ( $U_m$ ) of  $X_m$  in relation to  $Y_m$  are then used to infer past climate variables  $X_f$ .

from the composition of the fossil assemblages. This is a calibration problem (ter Braak & Prentice 1988). These two stages are summarised in Fig. (2). Before application to a fossil data-set, the robustness and reliability of the modern responses must be assessed by some form of cross-validation. This cross-validation normally provides the basis for producing robust estimates of the prediction error associated with a given approach to environmental reconstruction. This validation step is also shown in Fig. (2).

The term ‘transfer function’ is commonly used to refer to all of these reconstruction approaches (e.g. Guiot & de Vernal 2007). In particular, the modern analogue technique (see below under The Assemblage Approach) is frequently regarded as a ‘transfer function’ method. We avoid the use of the term ‘transfer function’ altogether because it has a clear and very different usage in statistics (Everitt 2003). In its place we use the term ‘calibration function’ approach to refer to reconstruction methods such as two-way weighted averaging (WA) and weighted averaging partial least squares regression and calibration (WA-PLS) where there are an explicit underlying taxon-climate response model and global estimation of taxon parameters (see below The Multivariate Calibration-Function Approach). Other methods that fall into this approach but that are not discussed in this essay include principal components regression (Imbrie & Kipp 1971), partial least squares regression and calibration (Martens & Næs 1989), Gaussian logit regression and calibration (ter Braak & van Dam 1989; Birks *et al.* 1990), multinomial logit regression and calibration (ter Braak & van Dam 1989; ter Braak *et al.* 1993; ter Braak 1995), and correspondence analysis regression (Roux 1979) – see Birks (1995) for a review of these ‘calibration function’ methods and related model-free environmental reconstruction techniques. ter Braak and Prentice (1988) and ter Braak (1996) provide a clear account of the distinction between regression and calibration. ter Braak (1995, 1996) notes that calibration is a type of multiple-species bioassay where species are used to infer environmental values from species data.

Modern training sets ( $\mathbf{Y}_m$ ) (e.g. pollen, chironomids) contain many taxa (e.g. 100-300) whereas there may be 50-200 samples. Data are usually quantitative and commonly expressed as percentages of the total sample count. They are thus closed, multivariate compositional data with a constant-sum constraint (ter Braak 1995). They often contain many zero values (up to 75% of all elements in the data matrix can be zero) for sites where taxa are absent or unrecorded. The data are complex, showing noise, redundancy, and internal correlations, and often contain what may, in a particular model, be regarded as outliers. Taxon abundance is usually a unimodal function of the environmental variables (ter Braak 1996).

Modern climate data ( $\mathbf{X}_m$ ) usually contain many fewer variables (c. 1-15) than the matching biological data. Climate data rarely contain zero or absent values. Quantitative climate variables often follow a normal distribution and commonly show linear relationships and high correlations between variables (e.g. mean July temperature and growing-day degrees). There is thus much data redundancy. The reliability and quality of the modern climate data can be problematical in some instances.

## Basic Requirements

There are at least seven major requirements in a quantitative climate reconstruction using any of the three main approaches (Birks 1995). These requirements are a result of the complex numerical and ecological properties of modern biological and climate data such as non-linear species-environment responses. These requirements are as follows (from Birks 1995).

1. A biological system is required that produces abundant identifiable fossils and that is responsive and sensitive to the climate variable(s) of interest today at the spatial and temporal scales of study.
2. A large high-quality data-set of modern biological data ( $\mathbf{Y}_m$ ) (species occurrences and absences, surface-samples from lakes containing modern assemblages of sub-fossil pollen, chironomids, etc.) and associated climate data ( $\mathbf{X}_m$ ) is available. This data-set should be representative of the likely range of climate variables in the past ( $\mathbf{X}_f$ ), be of consistent and detailed taxonomy and nomenclature (Finkelstein *et al.* 2006), be of comparable quality (data-extraction procedures, spatial scale, count size, sampling methodology, preparation procedures, counting techniques) and, in the case of surface samples, be from the same type of sedimentary environment as the fossil samples (e.g. lakes) (Bjune *et al.* 2010) and hence have comparable taphonomies (cf. Minckley & Whitlock 2000; Whitmore *et al.* 2005; Gonzales *et al.* 2009a; Tonello *et al.* 2009; Fletcher *et al.* 2010; Goring *et al.* 2010).
3. The fossil data-sets ( $\mathbf{Y}_f$ ) used for reconstruction purposes should be of comparable taxonomy and nomenclature, quality, and sedimentary environment as the modern data ( $\mathbf{Y}_m$ ).
4. Good independent chronological control is required for the fossil data-sets to permit correlations and comparisons and, if required, to allow an assessment of the rates of biotic response in relation to climate changes (Birks & Ammann 2000; Birks *et al.* 2000).
5. Robust numerical methods for the regression and calibration steps are required that can adequately model the complex, non-linear, and often unimodal relationships between modern taxa and climate. The numerical methods used are theoretically, statistically, and ecologically sound, are easy to understand, perform well with large and small data-sets and taxon-poor or taxon-rich assemblages, are not too demanding in terms of computing resources, and the relevant computer programs and data-sets are available to the research community.
6. Reliable and realistic numerical estimation of standard errors of prediction for the modern data-set as a whole ( $\mathbf{X}_m, \mathbf{Y}_m$ ) and for each reconstructed climate value ( $\mathbf{X}_f$ ) is required. As the reliability of the reconstructed climate values may vary from one fossil sample to another depending on, for example, composition or preservation, sample-specific standard errors of prediction are needed.

- Critical ecological, numerical, and palaeoclimatic evaluations of all reconstructions ( $\mathbf{X}_f$ ) are needed, as all numerical regression and calibration procedures are designed to produce a result. It is, however, essential to evaluate if the reconstruction is ecologically meaningful, and numerically and palaeoclimatologically reliable, and to validate it by comparison with other independent records.

### Basic Assumptions

There are at least six major assumptions in all climate reconstructions from palaeoecological data (Imbrie & Kipp 1971; Birks *et al.* 1990; Birks 1995, 2003). Originally it was thought that there were only five key assumptions (Birks 1995) but the work of Telford and Birks (2005; 2009) has highlighted a sixth major assumption. These assumptions apply, in varying degrees, to all the three major approaches to climate reconstruction.

- The  $m$  taxa in the training set ( $\mathbf{Y}_m$ ) are systematically related to the climate ( $\mathbf{X}_m$ ) in which they live.
- The climate variable(s) ( $\mathbf{X}_f$ ) to be reconstructed is, or is linearly, or at least monotonically, related to an ecologically important determinant in the system of interest and that it is on a continuous quantitative scale. Presence/absence (1/0) of a particular climate mode cannot be reconstructed using the methods discussed here.
- The  $m$  taxa in the training set ( $\mathbf{Y}_m$ ) are the same biological entities as in the fossil data ( $\mathbf{Y}_f$ ) and their ecological responses to  $\mathbf{X}_m$  and  $\mathbf{X}_f$  have not changed over the time represented by the fossil assemblages. Thus the contemporary spatial patterns of taxon abundance ( $\mathbf{Y}_m$ ) in relation to  $\mathbf{X}_m$  can be used to reconstruct changes in  $\mathbf{X}_f$  through time from  $\mathbf{Y}_f$  ('space-for-time' substitution).
- The mathematical methods used adequately model the species responses in  $\mathbf{Y}_m$  to  $\mathbf{X}_m$  and produce calibration functions ( $\mathbf{U}_m$ ) with sufficient predictive power to allow accurate and unbiased reconstructions of  $\mathbf{X}_f$ .
- Other environmental variables than the one(s) of interest ( $\mathbf{X}_f$ ) have had negligible influence on  $\mathbf{Y}_f$  during the time window of interest, the joint distribution of these variables of interest in the past was the same as today, or their effect on  $\mathbf{Y}_f$  did not lead to past changes in assemblage states resembling shifts indicative in the modern environment of changes in the variable of interest.
- In model validation and in the estimation of prediction errors by some form of cross-validation, the test data are statistically independent of the training set (Telford & Birks 2005; 2009).

We now consider the three major approaches to climate reconstruction, starting with the oldest and most-used approach, the indicator-species approach.

## THE INDICATOR-SPECIES APPROACH

### Basic Principles and One or a Few Indicator Species

Fossil occurrences of a species with known modern environmental preferences and tolerances provide a basis for inferring the past environment (Birks 1981). Assuming methodological uniformitarianism (Rymer 1978) and niche conservatism (Wiens & Graham 2005; Pearman *et al.* 2008a; Wiens *et al.* 2010), the past environment is reconstructed to have been the modern environmental, Grinnellian (Soberón 2007) or gamma-niche (Morin & Lechowicz 2008) of the species found as a fossil, namely "the range of values of environmental factors that are necessary and sufficient to allow a species to carry out its life history" (James *et al.* 1984). In reality we can only delimit from observational data the realised Grinnellian (cf. Soberón & Peterson 2005) and within it the realised regeneration niche (Grubb 1977) if data about reproduction and regeneration are available. Colwell and Rangel (2009) clarify the relationship between the Grinnellian (an environmental feature) and the Hutchinsonian niche (a species property) and its biotope (physical setting). The indicator-species approach requires information (Table 1) about what environmental factors may influence the distribution and regeneration today of the species concerned. (Species abundances are almost always not considered in the indicator-species approach in palaeoecology (but see Fauquette *et al.* 1998, 1999)).

The commonest means of obtaining such environmental information is to compare present-day distributions of species with selected climate variables of potential eco-physiological significance (Dahl 1998), such as the mean temperature of the coldest month, maximum summer temperature, or some measure of overall growing-season warmth such as respiration sums. If the geographical trend of an ecoclimate variable covaries with the present geographical distribution of the species of interest, a cause-and-effect relationship is often assumed (Birks 1981). For example, Conolly and Dahl (1970) related the modern distribution of *Salix herbacea* in the British Isles to the 23°C maximum summer temperature isotherm for the highest points where *S. herbacea* grows today. Macrofossil records show that this dwarf willow occurred widely in lowland Britain and Ireland during the late-glacial Younger Dryas stadial where today maximum summer temperatures are 26-27°C, suggesting there was a depression in maximum summer temperature of 3-4°C during the Younger Dryas (see Dahl 1998).

This correlative approach of comparing modern distributions with contemporary climate variables, pioneered by Andersson (1902, 1903, 1909), has proliferated in the last 10-15 years (Soberón 2007; Franklin 2010). Now called bioclimate-envelope modelling and involving a wide range of modelling methods (e.g. Guisan & Zimmermann 2000; Anderson *et al.* 2003; Thuiller *et al.* 2003; Araújo *et al.* 2005a, 2005b; Leathwick *et al.* 2005; Araújo & Guisan 2006; Elith *et al.* 2006; Phillips & Dudik 2008; Baselga & Araújo 2009; Elith & Leathwick 2009; Franklin 2010), it is widely used to predict impacts of future climate change on species distributions (e.g. Currie 2001; Shafer *et al.* 2001; Thuiller *et al.* 2005, 2006; Hamann & Wang 2006;

Heikkinen *et al.* 2006; Rehfeldt *et al.* 2006; Levinsky *et al.* 2007; Huntley *et al.* 2008; Iverson *et al.* 2008; Jeschke & Strayer 2008; Thuiller *et al.* 2008; Lawler *et al.* 2009; Soberón & Nakamura 2009; Franklin 2010; Yates *et al.* 2010). There is currently much debate about the basic assumptions underlying this approach (e.g. Pearson & Dawson 2003; Hampe 2004; Pearson & Dawson 2004; Heikkinen *et al.* 2006; Dormann 2007a; Beale *et al.* 2008; Jiménez-Valverde *et al.* 2008; Thuiller *et al.* 2008; Araújo *et al.* 2009; Aspinall *et al.* 2009; Beale *et al.* 2009; Duncan *et al.* 2009; Morin & Thuiller 2009; Peterson *et al.* 2009; Willis & Bhagwat 2009; Chapman 2010; Mouton *et al.* 2010; Real *et al.* 2010; Willis *et al.* 2010b), namely that present-day distributions are controlled by climate (Araújo & Luoto 2007; Beale *et al.* 2009; Blach-Overgaard *et al.* 2010; Chapman 2010), that the distributions are in equilibrium with climate today (Svenning & Skov 2004; Araújo & Pearson 2005; Svenning & Skov 2005, 2007; de Marco *et al.* 2008; Svenning *et al.* 2008; Normand *et al.* 2009), that the distributional data and the climate data are reliable (Rolland 2003; Kitricos & Leriche 2010) and, in the case of montane and alpine biota, that the modern climate data are from the same altitudes as where the species being modelled actually grow (Dahl 1951, 1998; Randin *et al.* 2006; Lundquist & Cayan 2007; Pape *et al.* 2009; Randin *et al.* 2009a; Scherrer & Körner 2010, 2011), and that the observed realised distributional area is close to the potential and fundamental distributional areas (Colwell & Rangel 2009; Soberón & Nakamura 2009). Despite all the effort in recent years that has gone into identifying and understanding the major sources of uncertainty associated with the use of bioclimate envelopes to model and predict current and future species distributions and developing performance criteria for different models (e.g. Kadmon *et al.* 2003; Araújo *et al.* 2005b; Guisan & Thuiller 2005; Luoto *et al.* 2005; Thuiller *et al.* 2005; Araújo & Guisan 2006; Araújo & Rahbek 2006; Elith *et al.* 2006; Heikkinen *et al.* 2006; Lawler *et al.* 2006; Araújo & Luoto 2007; Botkin *et al.* 2007; Heikkinen *et al.* 2007; Luoto *et al.* 2007; Beale *et al.* 2008; Green *et al.* 2008; Luoto & Heikkinen 2008; Araújo *et al.* 2009; Diniz-Filho *et al.* 2009; Elith & Leathwick 2009; Engler *et al.* 2009; Titeux *et al.* 2009; Randin *et al.* 2009b; Franklin 2010; Hoffman *et al.* 2010; Mouton *et al.* 2010; Smulders *et al.* 2010), the key assumption remains, namely that climate is assumed to limit the observed distribution in bioclimate envelope models. The greatest uncertainty is whether this assumption holds for the species modelled today (Beale *et al.* 2008; Duncan *et al.* 2009; Chapman 2010), and hence in predictions for the future (Dormann 2007a) and in the use of such bioclimatic-envelope models as a basis for inferring past climate from fossil remains.

An additional critical question in all bioclimate-envelope models is what climate variables should be included in such models (Gavin & Hu 2005, 2006; Norris *et al.* 2006; Jackson *et al.* 2009). Temperature comprises a large set of eco-physiologically relevant variables (Dahl 1998). The simple variable, annual mean temperature, covaries spatially with many broad-scale biotic patterns at regional and global scales (Leith & Whittaker 1975). Seasonal or monthly means may provide additional explanatory power (in a statistical modelling sense) (Jackson *et al.* 2009). These climate means and integrated eco-physiological measures (e.g. respiration

sum, growing-degree days) are widely used in bioclimatic-envelope modelling and perform well in predicting modern distributions. Of course, organisms have maximum and minimum temperature thresholds and their sensitivity may vary between individuals and populations (e.g. Pigott 1970, 1975; Carter & Prince 1981; Pigott 1982; Larcher 2003; Norris *et al.* 2006; Ibáñez *et al.* 2007, 2008; Cote *et al.* 2008; Jackson *et al.* 2009; Proctor 2009). As Jackson *et al.* (2009) discuss, a proximal temperature predictor of a tree-range limit might, in reality, consist of the probability that temperature falls below a threshold within a certain number of days after bud burst, the frequency of years in which summer temperatures persist below a certain threshold, or freezing resistance (e.g. Sakai 1971; Sakai & Weiser 1973; Bigler *et al.* 2007; Morin *et al.* 2007; Morin & Thuiller 2009). The potential number and variety of such proximal variables and hence niche dimensions may be enormous and be beyond the current limits of empirical, correlative procedures (Jackson *et al.* 2009). Such high dimensionality is usually ignored by assuming a strong covariance between a simple and easy-to-obtain climate variable (e.g. mean annual temperature) and ecologically critical variables (e.g. respiration sum) (Dahl 1998). The apparent success of bioclimate-envelope modelling comes from this strong spatial covariance between easily measured variables and the poorly understood and largely unknown ecologically critical variables (Jackson *et al.* 2009). From a palaeoecological viewpoint, a critical question is whether this spatial covariance remains constant through late-Quaternary time (assumption 5 above). Jackson *et al.* (2009) argue that although the strength and direction (negative or positive) of this climatic covariance may be conserved through time, the covariance structure (slope, intercept, dispersion) may change as atmospheric circulation patterns respond to different forcing factors, feedbacks, and changing atmospheric composition (Jackson & Overpeck 2000; Williams & Jackson 2007; Porch 2010). Such inconstancy will clearly reduce the predictive power and reliability of bioclimate-envelope models when applied to past and future climate states with different covariance structures than today's climate (Jackson *et al.* 2009; Porch 2010).

As not all possible climate states are realised today (Jackson & Overpeck 2000; Jackson & Williams 2004; Williams & Jackson 2007), this implies that in contemporary data an 'apparent' threshold value for some climate parameter (e.g. summer temperature) may not actually be ecologically relevant even though it is consistently correlated with 'real' threshold values of a climate variable of critical physiological importance today. For example, the apparent threshold in summer temperature may, in reality, be related to a real threshold in winter or spring temperature. As a result, the indicator-species approach to reconstruction may fail under climate conditions in the past (or predicted conditions in the future) simply because these climate states with different climate parameters and covariance structures are not realised today. A palaeoecological example is from the middle Weichselian (marine isotope stage 3a) (Helmens *et al.* 2007) in northern Finland where a disagreement is found between air temperatures inferred from aquatic organisms (chironomids using the calibration-function approach and aquatic plants and other invertebrates using the indicator-species approach) and temperatures inferred by

calibration functions from fossil pollen assemblages. One hypothesis to explain these discrepancies is that cold winter temperatures influenced the terrestrial vegetation but not the aquatic organisms, which in lakes are effectively isolated from sub-zero temperatures by winter ice-cover. This hypothesis implies that whereas both aquatic organisms and pollen assemblages co-vary strongly with summer, winter, and mean annual temperature in the modern environment, they were affected to a different extent by summer and winter temperature in the past in a 'summer-warm' tundra or forest-tundra landscape representing a non-analogue climate state. Although the problem of non-analogue climates arises in all climate reconstruction procedures, we suspect that the indicator-species approach and the assemblage approach (see below) are the most susceptible to biases resulting from non-analogue climates.

The robustness of bioclimate-envelope modelling results can be greatly enhanced by explicitly incorporating mechanistic links between the functional traits of organisms and their environments into species distribution models (e.g. Kearney & Porter 2004, 2009; Kearney 2006; Colwell & Rangel 2009; Buckley *et al.* 2010). Physiological and biophysical principles can be used to link distributional data to the physiological responses and constraints of species to provide a mechanistic view of the fundamental niche which can then be mapped and compared with the realised niche today. Such mechanistic niche modelling (Kearney 2006) has considerable potential in predicting or reconstructing responses in novel or non-equilibrium conditions (Colwell & Rangel 2009).

In some instances it may be more realistic to consider species distributions in relation to two or more climate variables (e.g. Hintikka 1963; Gavin & Hu 2005, 2006). This bivariate or even multivariate (Gavin & Hu 2005, 2006) approach was pioneered by Samuelsson (1916) for the northern limit of *Corylus avellana* in Scandinavia and developed by Iversen (1944) in his classic work on *Viscum album*, *Ilex aquifolium*, and *Hedera helix* in Denmark (see Fægri (1950) for a discussion of the importance of this bivariate approach). Samuelsson (1916) found that at the northern limit of *Corylus*, lower summer temperature could be compensated for by a longer growing season and thus that the presence of fossil *Corylus* nuts did not give a direct, single palaeotemperature value (Andersson 1902, 1903, 1909) but a series of such values, each corresponding to a given growing-season duration (Fægri 1950). On the basis of detailed field observations, Iversen (1944) delimited the 'thermal limits' or bioclimatic-envelopes within which *Viscum*, *Ilex*, and *Hedera* flowered and produced seed (the Grinnellian and regeneration niches within Denmark). Iversen showed that *Ilex aquifolium* is intolerant of cold winters but tolerant of cold summers. *Hedera helix* is intolerant of winters with mean temperatures colder than -1.5°C but requires warmer summers than *Ilex*. *Viscum album* is tolerant of cold winters but requires warmer summers than either *Ilex* or *Hedera*. These shrubs are ideal 'indicator species' in palaeoecology because their pollen is readily identifiable to species level, it is not blown great distances so interpretation problems arising from far-distance transport do not arise, and their fruits (nuts and berries) are rapidly dispersed by birds. Their distributions, at least at the scale studied by Iversen (1944), are likely to be in equilibrium

with climate. The validity of this assumption has been elegantly shown by Walther *et al.* (2005) who showed that the northern limit of *Ilex* has shifted northwards in the last 50 years, presumably in response to warmer winters in recent decades. From fossil pollen occurrences, Iversen (1944) applied his modern climate envelopes to infer that mid-Holocene summers were 2-3°C warmer and winters were 1-2°C warmer than today in Denmark.

### Many Indicator Species

This general bioclimate-envelope approach has been extended to several taxa simultaneously to identify areas of climate overlap based on pollen (e.g. Grichuk 1969; Grichuk *et al.* 1984; Markgraf *et al.* 1986; Kershaw & Nix 1988; McKenzie & Busby 1992), plant macrofossils (e.g. Sinka & Atkinson 1989; Mosbrugger & Utescher 1997; Sharpe 2002), chironomids (e.g. Dimitriadis & Cranston 2001), molluscs (e.g. Moine *et al.* 2002), ostracods (e.g. Horne 2007; Horne & Mezquita 2008), and beetles (e.g. Atkinson *et al.* 1987; Elias 1994, 1997; Elias *et al.* 1999; Elias 2001; Porch 2010), the so-called mutual climatic range (MCR) approach. The assumptions are identical to other bioclimate-envelope methods discussed above. Although the mutual climate range method is simple in principle, problems arise when, as a result of non-analogue climate states, taxa co-occur in fossil assemblages that do not have overlapping climate ranges today. In such circumstances, a possible solution is to find the least dissimilar combination(s) of taxa (Guiot & de Vernal 2007). The reconstructed values are ranges for each of the climate variables considered. In general, these ranges are narrower for species-rich assemblages. However, in very species-poor assemblages, the presence of a taxon with an extremely narrow range of occurrence today may strongly bias the reconstructed climate ranges. Various attempts have been made to narrow the climate ranges reconstructed by MCR. One attempt uses a sine model and maximum likelihood envelopes to obtain more reliable modern climate data for beetles in a topographically and climatically diverse landscape such as New Zealand (Marra *et al.* 2004; 2006). A second attempt at narrowing MCR ranges involves an ingenious probabilistic approach (Klotz & Pross 1999; Pross *et al.* 2000; Klotz *et al.* 2003; 2004). Instead of simply calculating environmental envelopes with an equal reconstructed probability of all climatic states contained within the envelope, this method estimates mutual climate probability spheres for fossil samples. A large number of MCRs are calculated based on the entire set of possible pollen assemblages that can be expected to occur in the modern environment based on present-day plant distributional data and modern climate data on a 0.5° x 0.5° grid (New *et al.* 1999). MCRs for the variable of interest calculated for the gridded synthetic pollen assemblages are compared to the observed values of the variable of interest. Probability distribution patterns within the different MCRs calculated based on the synthetic assemblages can be established and analysed statistically. The synthetic floras show that only a part of the estimated MCR is actually represented by measured values of the variable of interest in the modern environment. This restricted envelope is then used to correct MCRs calculated for fossil assemblages, assuming that MCRs inferred for fossil assemblages can be corrected based on the probability distribution pattern observed for similar

envelopes in the modern environment. These narrower ranges are considered to represent the probability interval for the climate variable of interest. For graphical display, only the mean values of the upper and lower limits of the probability intervals are shown (e.g. Feurdean *et al.* 2008a, 2008b). This approach illustrates how the MCR technique that relies entirely on total geographical ranges exaggerates the climate ranges, whereas the synthetic assemblage approach produces narrower, more likely climate ranges based on the probability distribution of the climate states within the MCRs calculated for the synthetic assemblages.

Kühl (2003) and Kühl *et al.* (2002) have combined the Iversen (1944), Grichuk (1969), MCR, and co-existence approaches (Atkinson *et al.* 1987; Mosbrugger & Utescher 1997) to provide a potentially more rigorous approach to climate reconstruction based on presence/absence data (e.g. plant macrofossils). Kühl's approach involves probability density functions (the so-called pdf-method) (Table 1). It combines modern plant distributional data for individual taxa with gridded climate data and summarises the modern climate-taxon relationships as conditional probability density functions (climate|taxon present) rather than as ranges or thresholds. A normal distribution is used to estimate the pdfs robustly. As the joint distribution of climate variables within Europe is not uniform, and as this distribution may have changed with time, a weighting of the modern climate variables is included to account for this non-uniform distribution and to make the estimated pdfs independent of the climate distributions today (Kühl *et al.* 2002). By assuming statistical independence, a joint pdf is calculated for a fossil assemblage as a product of the pdfs of the individual taxa in the assemblage. Each taxon is weighted by the dispersion of the modern climate range, so 'narrow' indicators receive 'high' weight. The product of the pdf's is the most likely past climate and its mean and standard deviation or confidence intervals are interpreted as the most probable climate and its uncertainty. The reconstructed climate is based on the complete distribution and climate data of all the taxa included and because it is a multiplicative technique, no single taxon is decisive (Kühl & Litt 2003). The pdf approach has been used to reconstruct Holocene (Litt *et al.* 2009; Kühl *et al.* 2010) and last interglacial and early Weichselian climate from pollen assemblages (expressed as presences and absences) (Kühl 2003; Kühl & Litt 2003; Kaspar *et al.* 2005). The pdf climate reconstructions have been compared, in many cases, with climate reconstructions based on quantitative pollen assemblage data using a modern analogue technique (see The Assemblage Approach below). These comparisons suggest that the pdf method generally underestimates past climate variability but may provide robust estimates of long-term climate trends rather than of climate variability of small amplitude (Litt *et al.* 2009; Kühl *et al.* 2010). This is not surprising as the bioclimatic envelopes of many central European trees and shrubs are very broad so that presence/absence data can only provide palaeoclimate information when the site in question is near the range margins of some of the 'key' indicator taxa (e.g. *Hedera*, *Ilex*, *Viscum*, *Taxus*). The pdf method has been extended by Neumann *et al.* (2006) as a Bayesian indicator taxa model and used by Kühl *et al.* (2007, 2010) and Kühl and Gobet (2010). This extension combines the conditional pdfs with the prior probability density of climate to estimate

Bayesian posterior probability density functions (Robertson *et al.* 1999). Two-dimensional normal distributions are used as parametric functions to describe the conditional pdfs of July and January temperature, thereby allowing the influence of these two variables to be considered together and hence for both variables to be reconstructed simultaneously (Kühl & Gobet 2010).

### Strengths and Weaknesses

There are many assumptions behind the various indicator-species-based approaches to reconstruct climate from fossil assemblages, just as there are in modern bioclimate-envelope modelling (Dormann 2007a). Perhaps the greatest limitation is that the indicator-species approaches are based on presence/absence data of taxa even where relative abundances of the taxa are often available, especially for pollen, chironomids, and ostracods. The indicator-species approach, relying on presence/absence data only, is very sensitive to count size, as the probability of finding a single pollen grain of a critical indicator taxon (e.g. *Hedera helix*) will increase as more and more pollen are counted. In some situations, when there are assemblages that no longer exist together today, abundances cannot be used in the assemblage or calibration-function approaches. For example, Pliocene fossil pollen assemblages from southern Europe contain not only Mediterranean taxa but also taxa that occur in China and North America today. In such instances, an indicator-species approach is the only choice. Fauquette *et al.* (1998, 1999) improved the basic MCR approach by using taxon abundances (as abundance classes of absent, present, strongly present, abundant) to refine the climate range identified by the presence of a given pollen taxon. For example, *Quercus* deciduous-type pollen is considered to be present at 5% with an annual temperature range between -20 and 30°C. Once its percentage passes a defined threshold, the taxon is considered to be abundant and the climate range is reduced. *Quercus* deciduous-type is considered abundant over 20% and the annual temperature range changes to between 5 and 23°C. This approach warrants further use as it is a considerable improvement over the use of presence/absence pollen data only in MCR and related techniques such as the probability density function methods of Kühl *et al.* (2002).

MCR approaches assume that a taxon has an equal probability of occurrence anywhere within its climate range (Huppert & Solow 2004; Horne & Mezquita 2008). This assumption is unlikely to be valid for ecological or biogeographical reasons (Christiansen & Fenchel 1977; ter Braak & Prentice 1988; Hengeveld 1990, but see Sagarin *et al.* (2006) for a contrasting analysis) and has been shown not to be true in many empirical studies (e.g. Horne & Mezquita 2008), in particular where response surfaces have been constructed to illustrate the probability of occurrence of a taxon at particular points in climate space (e.g. Huntley *et al.* 1995; Thompson *et al.* 1999a, 1999b, 2000; Shafer *et al.* 2001; Thompson *et al.* 2001; Bray *et al.* 2006). Given such response surfaces and/or the frequency of species occurrences in relation to selected climate variables, robust numerical methods of non-linear regression and calibration using maximum likelihood estimation exist for estimating not climatic ranges but the most likely value of the climate

variable of interest and its statistical uncertainty (e.g. ter Braak & Barendregt 1986; ter Braak & Looman 1986; ter Braak & Prentice 1988; Huppert & Solow 2004).

A major problem in the indicator-species approach is how to test the reliability of these methods (Mouton *et al.* 2010). Model validation ideally requires an independent test-set but when the two sets of variables (Figs. 1 and 2) being considered (geographical distributions of taxa and broad-scale climatic variables) show strong spatial auto-correlation (Beale *et al.* 2010; de Knecht *et al.* 2010), it is impossible to test a MCR model with an independent test-set (Telford & Birks 2005; 2009) because the geographical position of any test-set used will fall within the geographical range of the primary biological data used in the model (Araújo *et al.* 2005a). To address this issue in an MCR approach applied to beetle data, Bray *et al.* (2006) proposed “to avoid circular reasoning in these experiments (to evaluate the accuracy and sensitivity of the MCR method), care was taken to use only samples of modern beetles that have not previously been used in the construction of the primary MCR database”. However, this seems to be impossible to implement as the beetle data source they quote for the MCR database are broad-scale distribution maps and atlases containing the areas with the sites used to evaluate the MCR method. Porch (2010) uses leave-one-out cross-validation (jack-knifing) to evaluate beetle-based climate-envelope models in Australia where the modern beetle data are based on collections and museum specimens from 5500 separate locations. However, problems of spatial autocorrelations (Telford & Birks 2005; 2009) will arise nevertheless with data of these types, as discussed in the next section on The Assemblage Approach. Comparison of reconstructions with palaeoclimate records developed from other, independent climate proxies therefore seems the only way of evaluating MCR-based inferences whereas methods for producing reliable and robust estimates of prediction errors are presently not available for the standard MCR approach.

Despite its antiquity in palaeoecology, the indicator-species approach remains a potentially useful tool in reconstructing past climates from fossil assemblages of, for example, plant macrofossils, molluscs, and beetles where taxon abundances in the sediments are not necessarily a direct reflection of the former abundance of the taxon in the past. It is potentially valuable when the fossil assemblages appear to have no modern analogues, for example in the early or middle Pleistocene or the Pliocene, when no reliable modern assemblages are available because of intense human impact in, for example, central Europe or the Mediterranean basin, or when there are no data-sets of modern assemblages available at all (e.g. Porch *et al.* 2009). In the latter case, modern distribution data from a range of sources can be relatively easily incorporated into species-climate envelopes even when the biological sampling techniques are very different to standard palaeoecological approaches (e.g. Dimitriadis & Cranston 2001). It is important to remember, however, that data quality as well as data quantity determine the reliability of any bioclimate-envelope model (Mateo *et al.* 2010), as they do in all ecological or biogeographical studies.

Given the increasing availability of modern distributional data of taxa in the form of atlases and on-line data-bases

(Graham *et al.* 2004), of fossil pollen records in data-bases (Gajewski 2008), of gridded climate data (New *et al.* 1999), of robust non-linear numerical methods for regression modelling and calibration reconstructions (Araújo & New 2007; Thuiller *et al.* 2009), and of suitable software (e.g. Guo & Liu 2010; Tarroso & Rebelo 2010; Warren *et al.* 2010), what are the limiting factors in a detailed and rigorous application of the indicator-species approach in palaeoecology?

The basic assumptions of bioclimate-envelope modelling (Dormann 2007a) are as relevant in predicting modern and future distributions as they are in reconstructing past climates. The modern distributional data used in such modelling should include not only presences but also absences (Phillips & Dudík 2008; Phillips *et al.* 2009; Soberón & Nakamura 2009; Ward *et al.* 2009; Lobo *et al.* 2010, but see Phillips & Elith (2010) for using presence data only) if probabilities of occurrence are to be estimated (cf. Argáez *et al.* 2005). Transformation of abundance data (e.g. pollen percentages) into ordinal classes reflecting absence, presence, and abundance (Fauquette *et al.* 1998, 1999) should be used more frequently. The appropriateness of the modern climate data used is very critical (Daly 2006; Ashcroft *et al.* 2008, 2009; Fridley 2009; Randin *et al.* 2009a; Kitricos & Leriche 2010; Roubicek *et al.* 2010; Tabor & Williams 2010). In species-climate modelling it is assumed that modern climate data used are the actual climate values where the species occurs in its known localities. Climate stations tend to be at low altitudes whereas cold-tolerant species tend to occur at high altitudes, especially in the southern parts of their range (Lundquist & Cayan 2007; Pape *et al.* 2009). Unless the climate data are adjusted for the altitude of the localities where the species occurs, there will be a bias towards higher temperatures in climate reconstructions based on, for example, macrofossil or coleopteran assemblages of alpine and montane species. Although the interpolation algorithms and statistical models for modern climate estimation are sophisticated and high-resolution digital elevation models can be used to incorporate adiabatic lapse rates and regional climate (Randin *et al.* 2006), they still predict meteorological (2 m above ground surface) air conditions rather than actual life conditions for alpine plants and coleopterans (Scherrer & Körner 2010, 2011). It is questionable whether such gridded or interpolated climate data reflect the actual alpine environment where the majority of organisms live in micro-habitats strongly decoupled from atmospheric conditions and interacting with micro-topography at a very fine spatial scales (Lookingbill & Urban 2003; Scherrer & Körner 2010, 2011). Spatial heterogeneity in climate is a very important environmental feature and is of critical importance in the reliability and usefulness of bioclimate-envelope models that use mesoscale climate data (Ashcroft *et al.* 2009; Ackerly *et al.* 2010; Bennie *et al.* 2010). When the biological distribution data are derived from distribution maps and atlases, the altitudes at which the species grow are often unknown so the climate data from lowland climate stations used in MCR are likely to be ecologically unrealistic. Another critical point is that as new data-analytical techniques are rapidly being developed, existing taxon-climate data-sets should be made available to researchers developing new modelling procedures. As Huppert and Solow (2004) report in their work on developing new numerical techniques

to reconstruct climate from fossil beetle assemblages, “despite repeated attempts, we have been unable to acquire data to permit a direct comparison of the results of our method with the results of the MCR method. We recognize the value of such a comparison and we regret the unavailability of such data”. As many indicator-species approaches, particularly the MCR method, place major emphasis on the range limits of species rather than the probability of occurrence within their range, understanding the dynamics of range limits in relation to climate is critical (Gaston 2009; Pigot *et al.* 2010). Recently, Keith *et al.* (2008), Anderson *et al.* (2009), Buckley *et al.* (2010), and Huntley *et al.* (2010) have shown the importance and feasibility of moving from simple bioclimatic-envelope models to models that incorporate dynamic climate change and metapopulation dynamics (see also Jackson *et al.* 2009 and Midgley *et al.* 2010). Palaeoecologists using bioclimate-envelope approaches to reconstruct past climate from fossil assemblages could profitably test the robustness of their correlative methods by applying appropriate statistical tests to establish if the distribution of the taxa used in climate reconstructions show statistical relationships with climate, along the lines of Beale *et al.* (2008, 2009) and Chapman (2010). A final and currently insoluble problem concerns reconstructions of past climatic variables under conditions with novel or non-analogue climates (Jackson & Williams 2004; Williams & Jackson 2007). Here the covariance structure between the climate variable being reconstructed such as mean July temperature and an eco-physiologically relevant variable such as respiration sum may not be the same as it is today (Jackson *et al.* 2009). Climate reconstructions under such conditions remain a major challenge (Morin & Lechowicz 2008; Gonzales *et al.* 2009b; Porch 2010). Predicting biotic responses in a future world with non-analogue climate (MacDonald 2010) is equally challenging.

## THE ASSEMBLAGE APPROACH

### Basic Principles

In contrast to the indicator-species approach, the assemblage approach considers the fossil assemblage as a whole and the relative abundances of all the different fossil taxa (Birks & Birks 1980). It has been widely used in an intuitive, non-quantitative way for many decades. For example, fossil pollen assemblages can be interpreted, using a range of criteria, as reflecting tundra, pine forest, or deciduous forest. Past environmental inferences for the fossil assemblages are then based on the present-day environment in which these vegetation types occur today. In the 1960s the interpretation of fossil pollen assemblages in terms of past vegetation types was put on a firmer basis through systematically studying modern pollen assemblages from known vegetation types (Wright 1967) and then by comparing modern and fossil pollen assemblages in a qualitative modern-analogue matching procedure (e.g. McAndrews 1966; Wright *et al.* 1967; Birks 1973a, 1973b). More recently, the assemblage approach has developed a more quantitative basis, the so-called modern analogue technique (MAT) and the related response-surface approach (Table 1).

The basic idea of MAT is to compare numerically, using a dissimilarity measure (e.g. squared chord distance)

(Overpeck *et al.* 1985; Prell 1985; Simpson 2007, 2011) a fossil assemblage with modern assemblages. Having found the modern sample(s) that is (are) most similar to the fossil assemblage, the past climate for the fossil sample is inferred to be equivalent to the state of the climate variable(s) of interest for the analogous modern sample(s). The procedure is repeated for all fossil samples and a simultaneous reconstruction of one or several climate variables can be made on the basis of the modern analogues. The climate reconstruction(s) can be based on the modern sample that most closely resembles the fossil assemblage or, more reliably, it can be based on the mean or the weighted mean of, say, the 10 or 25 most similar modern samples, with weights being the inverse of the dissimilarities so that modern samples with the lowest dissimilarity (= highest similarity) have the greatest weight in the climate reconstruction. Williams and Shuman (2008) found for a large North American data-set of modern pollen assemblages that strong weighting by the inverse of the squared chord distance dissimilarity reduced the predictive ability of their MAT model, whereas there were little differences if an unweighted average or an average weighted by the inverse of the chord distance was used.

A critical question is to discover how many analogues to include in a reconstruction. An estimate of how many analogues to include in the reconstruction must be derived by numerical cross-validation. Procedures involving receiver operating characteristic (ROC) curves (Gavin *et al.* 2003; Wahl 2004) are not appropriate here as the analogues being sought are not 1/0 vegetation types (as in analogue matching (Simpson 2011) with the ROC approach) but the closest multivariate assemblage. Leave-one-out estimates of the climate variable(s) of interest based on the modern assemblage data-set are obtained for 1, 2, ...  $k$  modern samples when the inferred or predicted value for the one sample left out is determined from the modern training data-set (excluding the single test sample). The procedure is repeated, leaving out another sample to form a test sample, until all samples have been excluded once. The predicted and observed values for the excluded samples based on 1, 2, ...  $k$  analogues are differenced to calculate the prediction residuals. The root mean square error of prediction (RMSEP) and associated model statistics such as the mean bias, the maximum bias, and the coefficient of determination ( $r^2$ ) between predicted and observed values (Birks 1995; Barrows *et al.* 2000; Barrows & Juggins 2005) are then calculated (Simpson 2007, 2011). The value of  $k$  that gives the lowest RMSEP is then used in climate reconstructions from fossil assemblages and this RMSEP is sometimes assumed to be the prediction error for all the palaeoenvironmental inferences. With improving computing resources, it is now possible to derive the value of  $k$  for models with the lowest RMSEP as assessed by the more computer-intensive procedure of bootstrapping (Birks *et al.* 1990) the modern samples and to generate bootstrap-derived sample-specific errors for each fossil sample using many (usually 1000) bootstrap cycles (Simpson 2011). An alternative measure of reconstruction uncertainty suggested by ter Braak (1995) is the standard error of the estimates calculated as the weighted sample variance among the  $k$  analogues where the weights are the inverse of the dissimilarities (Simpson 2011). This measure has several attractive features (ter Braak 1995):

1. it is independent of the magnitude of dissimilarity to the  $k$  most similar analogues
2. if the magnitude of the standard error is large relative to the values computed for the modern samples, this indicates that no close modern analogues exist for the fossil assemblage and the climate estimate cannot be trusted
3. the fossil assemblage may be similar to several modern analogues that are very dissimilar in terms of their modern climate, thereby resulting in a high standard error. As a consequence of this multiple analogue situation, the assemblages are not responsive to the climate variable of interest, at least in the region of the climate gradient represented by the fossil sample.

As ter Braak (1995) notes, it can be considered an advantage of MAT that it highlights such problems. MAT is more commonly known in the data-analytical literature as  $k$ -nearest neighbours (ter Braak 1995). Analogue matching (Simpson 2011) is a method in classification (assignment) or discrimination. MAT is also one of several approaches to inverse regression involving non-parametric smoothers (ter Braak 1995). However, it does not fit a global model. Instead it fits a local model to small subsets of the training data for each fossil sample. To be effective MAT therefore requires a good coverage of samples in local space which implies a well-populated network of samples across the climate gradient(s) of interest (Juggins & Birks 2011).

As in all other reconstruction procedures, it is essential that the fossil data-sets used for climate reconstruction are from the same sedimentary environment as the modern training set. The most striking violation of this basic requirement is provided by the recent study of Fletcher *et al.* (2010) where they use a 3350-sample data-set of terrestrial modern pollen assemblages from a variety of continental sedimentary environments in a MAT-based climate reconstruction for a *marine* core from the Alboran Sea, despite the well-known differences in taphonomy of terrestrial and marine pollen assemblages. The authors deleted *Pinus* from the modern and fossil samples because of its over-representation in marine assemblages. Clearly the resulting climatic reconstructions may be heavily biased by taphonomic differences between marine and terrestrial pollen assemblages and by deletion of the most abundant pollen taxon (see also Dormoy *et al.* 2009).

In terms of the underlying mathematical properties of MAT and the revised MAT approach of Waelbroeck *et al.* (1998) that combines MAT with the response-surface method of Bartlein *et al.* (1986) (the so-called revised analogue method (RAM)), they are inverse reconstruction procedures (ter Braak 1995). They have no assumed underlying linear or unimodal species-environment model, they use full dimensionality (all taxa), and estimate smooth non-parametric local functions rather than parametric global functions, resulting in their inability to extrapolate beyond the range of the modern training set. Although extremely different in their numerical computations and algorithms, artificial neural networks (ANN) have identical mathematical properties when applied to palaeoenvironmental reconstructions. Malmgren and Nordlund (1997), Malmgren *et al.*

(2001), Burrows *et al.* (2005), and Kucera *et al.* (2005) describe the use of ANN in marine sea-surface temperature reconstructions; Peyron *et al.* (1998, 2000, 2005), Tarasov *et al.* (1999a, 1999b), and Guiot *et al.* (1996) provide examples of their use with terrestrial pollen data; and Racca *et al.* (2001; 2003; 2004) discuss ANN applications in palaeolimnology. Telford and Birks (2005) show that in cross-validation tests with a truly independent test-set, MAT and ANN perform poorly as both involve estimating smooth local functions. In practice, ANN appears, in some ways, to be a computationally demanding way of implementing simple MAT! There are several disadvantages to ANN. First, it is difficult to interpret network coefficients in any ecologically meaningful way as ANN is essentially a 'black box' tool. Second, artificial neural networks are very flexible functions and with large numbers of taxa in training sets, they are very prone to overfitting. Telford *et al.* (2004) and Telford and Birks (2005) show the importance of using, in addition to the training set and an independent test-set, a separate optimisation data-set to optimise the networks based on the training set or to select the optimal number of analogues in MAT and RAM.

There are many examples of the use of MAT for reconstructing Holocene climates from pollen data (e.g. Bartlein & Whitlock 1993; Cheddadi *et al.* 1998a; Kerwin *et al.* 2004; Nakagawa *et al.* 2002; Sawada *et al.* 2004; Peryon *et al.* 2005; Ortu *et al.* 2006, 2008, 2010; Tarasov *et al.* 2007; Feurdean *et al.* 2008a, 2008b; Bordon *et al.* 2009; Köhl *et al.* 2010). MAT has been extended by Guiot (1990), Guiot *et al.* (1989; 1992), and Cheddadi *et al.* (1998b) to reconstruct several climatic variables from pollen assemblages for the last glacial-interglacial cycle using so-called 'palaeobio-climatic operators' as weights to emphasise preferentially the climatic signal within fossil pollen data (see below).

A realistic assessment of the ability of MAT, RAM, and ANN (and other calibration-function based reconstruction procedures) to reconstruct modern climate can be made not by simple leave-one-out cross-validation but by the more powerful  $n$ -fold leave-out cross-validation or split-sampling of the modern data-sets ( $\mathbf{X}_m$ ,  $\mathbf{Y}_m$ ) (Barrows & Juggins 2005). This involves splitting the data into  $n$  (usually 5 or 10) random subsets containing 20% or 10% of the modern data points, respectively. Each subset is removed from the data-set in turn and the remaining modern samples are used to generate reconstruction models that are then applied to the omitted samples. The process is repeated for each subset to give a leave-out climate estimate for each modern sample. This data-splitting procedure is repeated 5 or 10 times, each with random partitions of the modern data to provide a more reliable estimate of the model performance than is provided by simple leave-one-out cross-validation (Barrows & Juggins 2005).

## Response Surfaces

Before discussing the limitations and inherent problems of MAT (and ANN), we will outline the so-called response-surface approach (Table 1) because, in reality, it is a form of MAT using smoothed modern pollen data rather than the individual pollen assemblages. Response surfaces are three-dimensional graphical representations of the occurrence and/or abundance of individual taxa in modern climate space

(Huntley 1993). The  $x$  and  $y$  axes represent climate variables (e.g. mean July temperature, mean January temperature) and the  $z$  axis represents the occurrence or relative abundance of the taxon of interest. Such modern taxon-climate response surfaces have been constructed to summarise patterns of modern taxon abundances along major climate gradients. The surfaces have been fitted by multiple regression (e.g. Bartlein *et al.* 1986), locally weighted regression (e.g. Huntley *et al.* 1989; Prentice *et al.* 1991; Bartlein & Whitlock 1993), or generalised additive models (e.g. Šmilauer & Birks 1995). Palaeoenvironmental reconstructions (e.g. Webb *et al.* 1987; 1993a; 1993b; Allen *et al.* 1995; Watts *et al.* 1996a; 1996b; Bartlein *et al.* 1998; Webb *et al.* 1998; Allen *et al.* 2000, 2002; Kerwin *et al.* 2004) are made by 'stacking' modern surfaces to produce synthetic modern assemblages for a series of grid nodes, usually 20 x 20 nodes in modern climate space. These synthetic assemblages are then compared to fossil assemblages using a dissimilarity measure, usually the squared chord distance as in MAT. Climate values for the 10 grid nodes with synthetic pollen spectra most similar to the fossil assemblage are then used to infer the past climate. The final inferred value is a mean of the climate variable weighted by the inverse of the squared chord distance (Prentice *et al.* 1991). Climate reconstructions are thus done by MAT but the modern data consist of fitted or smoothed pollen values in relation to modern climate and not the original pollen values. ter Braak (1995) calls this a 'smooth response-surface approach via local weighted averaging'. The fitted response-surface values naturally smooth the data to varying degrees depending on the extent of smoothing and procedure used (ter Braak 1995). Much inherent local site-scale variability that is assumed to be unrelated to broad-scale climate patterns is thus removed (Bartlein & Whitlock 1993).

MAT and response surfaces including RAM are widely used with quantitative terrestrial pollen data and with marine fossil assemblages (e.g. Barrows & Juggins 2005; Kucera *et al.* 2005; Guiot & de Vernal 2007). Thompson *et al.* (2008) present an ingenious and rigorous use and validation of MAT with qualitative presence/absence data of plant macrofossils. They applied MAT with four different similarity coefficients to a continental-scale set of modern plant distribution presence/absence data and associated bioclimate data for North America. For each of 32,211 grid points, they calculated the similarity between the floristic data in the grid point of interest and all other points on the grid (excluding nearby points within 50 km of the grid point of interest to minimise the effects of spatial autocorrelation). The climate of the points with most similar floras was used to estimate the climate at the grid point of interest. They showed the accuracy of the climate estimates is strongly linked to the number of taxa in the target assemblage and like all MAT (and ANN) procedures, it is constrained to produce estimates within the range of observed values. They then applied their MAT procedures to a Last Glacial Maximum plant macrofossil assemblage from a packrat midden in southern Nevada (Thompson *et al.* 1999c) to derive estimates of moisture conditions and growing season warmth, and compared the MAT estimates with results from a mutual climate range method (Thompson *et al.* 1999c). They found that the MAT methods produced more plausible reconstructions of temperature-related variables than moisture-related variables.

## Weaknesses

MAT, ANN, and response surfaces suffer from a number of problems (Brewer *et al.* 2007) including assemblages containing taxa that respond to different environmental variables, the estimation of realistic errors of prediction, fossil assemblages that have no modern analogues or have multiple modern analogues, and assessing model performance realistically because of inherent spatial autocorrelation that all methods using local-estimation procedures such as MAT and ANN are particularly susceptible to. They, like the multivariate calibration-function approach discussed in the next section, are data-demanding techniques in terms of data quality and data quantity. It is particularly important that all the modern and fossil assemblages are from the same sedimentary environment to minimise variation due to taphonomy, basin size, relevant source areas, etc (cf. Fletcher *et al.* 2010). A further limitation with MAT and related methods is that because they do not assume any response model or have any underlying statistical model or basis, it is not possible to develop model evaluation or diagnostic procedures analogous to regression diagnostics in statistical modelling, nor to have any clear principles of model selection and testing. Unlike methods with a statistical basis and assumed response model, MAT, ANN, etc, have no clear testable assumptions and are thus, in many ways, a 'black box', particularly ANN and some variants of MAT. Without an underlying model or statistical basis, it is difficult to develop 'artificial' simulated data with realistic assumptions for numerical evaluation experiments (e.g. ter Braak 1995).

In order to deal with the problems of taxa responding to different environmental variables, Guiot (1990) derives 'palaeo-bioclimatic operators' to weight preferentially the dissimilarity calculated between the modern and fossil pollen assemblages. An ordination (usually a principal components analysis) is made of all the modern samples, and the distribution of taxa are examined to identify the component axis that is most strongly related to the climatic variable of interest. The taxon scores on this axis are then used as weights. Other variants of these operators or weighting procedures are discussed by Birks (1995).

The error involved in reconstructing climatic values is usually derived from the dispersion of the climatic values of the selected analogues around the weighted average of the climatic value. This approach has the advantage of considering the higher and lower estimates separately, resulting in non-symmetric and possibly more realistic error estimates. However, the total error is underestimated (Brewer *et al.* 2007). Nakagawa *et al.* (2002) use a linear relationship calculated between observed and estimated climate values at modern pollen sites. The confidence intervals for this linear relation are then applied to derive probability distributions and/or confidence intervals for the estimates of past climate. This procedure needs a dense network of modern climate stations. Potentially the most reliable procedure for error estimation in MAT is bootstrapping (Simpson 2011) to derive sample-specific error estimates for past climate. However, the reliability of such bootstrap estimates in MAT in the presence of spatial autocorrelation has not yet been investigated.

A very significant problem for MAT and related methods including ANN, is situations where no good modern analogues can be found for a fossil assemblage (e.g. Ortu *et al.* 2006, 2010; Minckley *et al.* 2008). Such non-analogue situations can result from, for example, unique biotic interactions, vegetation dynamics, novel climates (low CO<sub>2</sub> concentrations, unique combinations of summer and winter irradiance), anthropogenic influence, and lack of modern samples (Jackson & Williams 2004). A critical question therefore is what is a 'good analogue' and what is a non-analogue. Definitions are difficult and involve total taxonomic composition, the relative abundance of the numerical dominants, the identification of specialists and generalists, and life-forms and functional types. These biological attributes are not easy to measure and even more difficult to summarise in a single dissimilarity or analogue measure (Juggins & Birks 2011).

Davis *et al.* (2003) (see also Davis & Brewer 2009) modified the MAT approach to help circumvent the non-analogue problem. They grouped the modern and fossil pollen taxa into plant functional types (PFTs) that are broad classes of plants defined by stature, life-form, phenology, physiology, leaf-form, and climatic tolerances (Duckworth *et al.* 2000). Analogue matching was carried out on the basis of the PFT abundances. As each PFT represents a set of ecologically similar taxa (Prentice *et al.* 1996), the assumption is made that a similar climate will always support a similar composition of PFTs, even if the actual species composition within a PFT may vary for non-climatic reasons (Brewer *et al.* 2007). This approach of grouping pollen taxa into PFTs has also been used by Peyron *et al.* (1998) and Tarasov *et al.* (1999a, 1999b) in their reconstructions of Europe and Eurasia during the Last Glacial Maximum (see also Peyron *et al.* 2000) and at 6000 BP, and, rather surprisingly, in reconstructions of late-Holocene climate by Köhl *et al.* (2010). Because of the large climatic amplitudes of many PFTs, this approach can lead to very large reconstruction errors, for example of 5–10°C for annual mean temperature.

An ingenious approach to trying to cope with the non-analogue problem is to construct so-called expanded response-surfaces (Gonzales *et al.* 2009b). The method assumes that pollen taxon abundances follow a symmetrical unimodal distribution along climatic gradients, that taxa with truncated distributions in modern climate space occupy a subset of their fundamental or potential niche, and thus the truncated distributions can simply be expanded by mirroring around the distributional mode to recover the portion of the potential niche not realised today in modern climate space (Gonzales *et al.* 2009b). This approach should usefully be compared with conventional and more rigorous statistical modelling of taxon responses in relation to climate by generalised linear (ter Braak & Looman 1987) or generalised additive modelling (Yee & Mitchell 1991), both of which allow some degree of extrapolation.

A further problem in the general MAT approach is the existence of multiple analogues (Huntley 1996, 2001). These arise when the fossil assemblage is similar to several modern samples that differ widely in their modern climate. For example, modern assemblages dominated by *Pinus* pollen can be derived from northern, central, or Mediterranean Europe, all of which have very different climates today (ter

Braak *et al.* 1996; Birks 1998). Clearly in these circumstances, the MAT approach will provide unreliable and ambiguous reconstructions. To help circumvent this problem, information from a second climate proxy has been used to constrain the climatic reconstruction. Cheddadi *et al.* (1997) reconstructed past climate from pollen data using MAT constrained by palaeolake-level data. The constraint consists of restricting the set of modern pollen samples considered as analogues for the fossil samples to those locations where the implied change in annual precipitation minus evapotranspiration (P-E) is consistent with the regional change in moisture balance as indicated by the palaeolake-level data. This constraint led to an improvement in the spatial coherency of the reconstructed palaeoclimate anomalies, particularly (and not surprisingly) for P-E (Brewer *et al.* 2007). Other proxies used as constraints have included biome reconstructions derived from assemblages of plant macrofossils or fossil pollen (e.g. Magny *et al.* 2001, 2003; Ortu *et al.*, 2006, 2010), beetle-based climate reconstructions, and sedimentary parameters (see Birks 1995 for a review). The use of certain proxies as a constraint has the disadvantage that these proxies can no longer be used to verify independently the reconstruction results (Brewer *et al.* 2007). However, if the climate response of the proxy is well understood, the use of a constraint may provide a more consistent set of results than those from an unconstrained MAT (e.g. Magny *et al.* 2001, 2003; Muller *et al.* 2003; Bordon *et al.* 2009). The need of using such constraints raises questions about the suitability and robustness of MAT as a climate-reconstruction procedure when applied to pollen data, especially when the constraints applied are themselves derived from the pollen data (e.g. biomes).

### Spatial Autocorrelation

The problem of spatial autocorrelation (Legendre 1993) in modern assemblage and associated environmental data-sets and its influence on the performance of environmental reconstruction techniques, both the assemblage approach and the multivariate calibration-function approach (see next section) has, until recently (Telford & Birks 2005; 2009), been ignored. Cross-validation (leave-one-out, split-sampling into a training or calibration data-set and a test-set, and bootstrapping) assumes statistical independence of samples, an assumption that is violated in the presence of spatial autocorrelation (Legendre 1993), leading to over-optimistic estimates of RMSEP and other model performance statistics.

Telford and Birks (2005), using an extensive modern foraminiferal assemblage data-set from the Atlantic showed that MAT (and ANN – see also Telford *et al.* (2004)) utilises spatial structure in the assemblage data that is uncorrelated with the environmental variable being modelled and thus improves the apparent predictive-model abilities of MAT and ANN. This residual structure is likely to be related to spatial structures in environmental variables other than the one being reconstructed (see Telford 2006). When a spatially independent test-set (e.g. South Atlantic) is used to assess the MAT and ANN model performances when the calibration data-set is from the North Atlantic, MAT and ANN perform no better than standard calibration-function methods such as two-way weighted averaging (see The Multivariate Calibration-Function Approach), suggesting that previously

reported MAT and ANN model performance statistics in the palaeoceanographic literature are over-optimistic (Telford *et al.* 2004). These findings highlight that spatial autocorrelation is a problem in the assemblage approach to environmental reconstruction when the environmental variable(s) of interest is spatially smooth and where reconstruction methods such as MAT and ANN involve local non-parametric estimation of functions rather than global parametric estimation as in two-way weighted averaging.

The extent to which spatial autocorrelation affects other environmental variables and other proxies has been addressed by Telford and Birks (2009). In modern data-sets with strong spatial autocorrelation (e.g. pollen and July sunshine, planktonic foraminifera and summer sea-surface temperature, benthic foraminifera and salinity), MAT generally outperforms other reconstruction techniques, whereas MAT underperforms with data-sets with weak spatial autocorrelation (e.g. diatoms and lake-water pH). The most likely explanation for this behaviour is that MAT finds a fit between the species assemblages and the environment that is local rather than global, and as a result MAT is not robust to the presence of spatial autocorrelation. In light of these results, Telford and Birks (2009) recommend that MAT not be used with modern data-sets where strong spatial autocorrelation is present despite its widespread use in marine studies (Guiot & de Vernal 2007) and terrestrial pollen-analytical studies (Williams & Shuman 2008).

Telford and Birks (2009) present two simple means for evaluating the influence of spatial autocorrelation on the performance of environmental-inference models. The first is simply to delete samples at random and derive modern performance statistics (RMSEP, etc) for different sized modern data-sets. Samples are also deleted that are geographically close to the test sample and performance statistics for both types of deletion are calculated and compared. If strong spatial autocorrelation is present, deleting geographically close samples will preferentially delete the most similar samples. Because of spatial autocorrelation these close samples will bias inferences and lead to seemingly 'good' performance of the reconstruction model. Thus their deletion should drastically decrease model performance. In contrast, random deletion should have much less impact on model performance. The second method involves *h*-block cross-validation and provides a means of evaluating a reconstruction model in the presence of spatial autocorrelation. In time-series analysis, *h*-block cross-validation involves deleting an observation from a model set along with *h* observations on either side to minimise the effects of temporal autocorrelation (Burman *et al.* 1994). The idea has been extended to spatial autocorrelation (Telford & Birks 2009) by defining radius *h* based on the range of a geostatistical variogram of the model residuals from the basic reconstruction model. RMSEP and related model statistics are calculated for *h*-block cross-validation and compared with the same statistics based on simple leave-one-out cross-validation. Telford and Birks (2009) showed a consistent increase (up to 230%) in RMSEP in *h*-block cross-validation using MAT with modern assemblage data-sets with high spatial autocorrelation.

Despite this work, some investigators claim that spatial autocorrelation does not bias MAT models (e.g. Guiot & de

Vernal 2007; Fréchet *et al.* 2008; Bonnet *et al.* 2010). The basis for this claim is generally to select at random a test-set (e.g. 10% or 20% of the modern data) and to estimate RMSEP for this test-set. In this procedure, the test-set samples are not spatially and hence not statistically independent and so cannot be used to test correctly for the influence of spatial autocorrelation. To do this appropriately, the test-set must be *spatially independent* from the calibration data-set (e.g. Telford & Birks 2005). However, using spatially independent test-sets is rarely possible, so the two procedures of Telford and Birks (2009) outlined above should be used to test for the influence of spatial autocorrelation on the performance of MAT and ANN when applied to spatially structured assemblage data-sets.

Biogeographers and ecologists are also beginning to wrestle with the very real problems of spatial autocorrelation in the numerical analysis of modern species distribution data (e.g. Lichstein *et al.* 2002; Araújo *et al.* 2005; Betts *et al.* 2006; Rangel *et al.* 2006; Dormann *et al.* 2007; Dormann 2007a, 2007b, 2007c; Beale *et al.* 2010). As in palaeoecology, there is currently a lively debate about the need or otherwise to account for the effects of spatial autocorrelation (e.g. Lennon 2000; Betts *et al.* 2006, 2009; Beale *et al.* 2007, 2010; Hawkins *et al.* 2007; Bini *et al.* 2009; Dormann 2009; de Knecht *et al.* 2010) in bioclimate-envelope modelling and other forms of species distribution modelling.

Despite the long tradition in biogeography of species-climate modelling and analysis, this has almost exclusively involved individual species responses (e.g. Elith & Leathwick 2009; Franklin 2010). Recently, biogeographers (Baselga & Araújo 2009, 2010) are considering species assemblages in relation to climate variables to predict assemblage dynamics under future climates using the recently developed technique of canonical quadratic ordination (Yee 2004, 2006). Other predictions of future assemblage change include Hamann and Wang (2006) and Rehfeldt *et al.* (2006).

As in the indicator-species approach, there are increasingly close research parallels between the palaeoecological community on the one hand, and the ecological and biogeographical communities using the assemblage approach on the other. Both research communities should be aware of current trends within both research communities as there is likely to be much of relevance to each research community.

## THE MULTIVARIATE CALIBRATION-FUNCTION APPROACH

The third approach to palaeoclimate reconstruction is the multivariate approach involving calibration functions. This approach involves underlying statistical models with global estimation of parametric functions for all the taxa present. This allows, to some degree, some extrapolation (ter Braak *et al.* 1993; ter Braak 1995) and helps to overcome some of the problems inherent in the assemblage approach.

The basic idea of quantitative climate reconstructions involving multivariate calibration functions (Table 1) is summarised in Figs. (1 and 2). As with the indicator-species and assemblage approaches, it assumes that there is one or more climate variable(s) ( $X_i$ ) to be reconstructed from fossil

biotic assemblages ( $\mathbf{Y}_f$ ) consisting of  $m$  taxa in  $n$  samples. To estimate or reconstruct  $\mathbf{X}_f$ , we need to model numerically the responses of the same  $m$  taxa today in relation to modern climate variable(s) ( $\mathbf{X}_m$ ). This requires a modern ‘training set’ or ‘calibration set’ of  $m$  taxa at  $t$  sites or modern samples ( $\mathbf{Y}_m$ ) studied as assemblages preserved in modern surface sediments (e.g. surface (0-1 cm) lake muds) with associated modern climate variables ( $\mathbf{X}_m$ ) for the same  $t$  modern sites or samples. The modern relationships between  $\mathbf{Y}_m$  and  $\mathbf{X}_m$  are modelled numerically by some form of regression analysis and the resulting calibration function is then used to transform the fossil data ( $\mathbf{X}_f$ ) into quantitative estimates of the past climate variable(s) ( $\mathbf{X}_f$ ).

Since Imbrie and Kipp (1971) revolutionised Quaternary palaeoclimatology by presenting, for the first time, a numerical procedure (principal components regression) for quantitatively reconstructing past environments (in their case sea-surface summer and winter temperatures and salinity) from fossil foraminiferal assemblages, several numerical techniques have been developed for deriving multivariate calibration functions (see Birks 1995; ter Braak 1995; Birks 1998; Guiot & de Vernal 2007; Juggins & Birks 2011 for reviews).

### Basic Considerations

There are four basic considerations when deriving multivariate calibration functions (ter Braak 1995). First, there is the choice between classical and inverse regression approaches. The classical approach is of the general form:

$$\mathbf{Y}_m = f(\mathbf{X}_m) + \text{error}$$

The function  $f(\ )$  is estimated by linear or non-linear multivariate regression from the modern training set ( $\mathbf{X}_m$ ,  $\mathbf{Y}_m$ ). The estimate of  $f(\ )$  is then ‘inverted’ to infer the past climate  $\mathbf{X}_f$  from  $\mathbf{Y}_f$ :

$$\mathbf{X}_f \approx f^{-1}(\mathbf{Y}_f)$$

‘Inversion’ involves finding the past climate variable that maximises the likelihood of observing the fossil assemblages in that climate. If the function  $f(\ )$  is non-linear, which it almost always is, non-linear optimisation procedures are required and these can raise programming problems (Birks 2001).

Alternatively there is the simpler but ecologically slightly curious inverse approach of:

$$\mathbf{X}_m = q(\mathbf{Y}_m) + \text{error}$$

where the difficult inversion step is avoided by estimating directly the function  $q(\ )$  from the training set by inverse regression of  $\mathbf{X}_m$  on  $\mathbf{Y}_m$ , in contrast to the classical approach where  $\mathbf{Y}_m$  is regressed on  $\mathbf{X}_m$ . The inferred past climate ( $\mathbf{X}_f$ ), given a fossil assemblage of  $\mathbf{Y}_f$  is simply:

$$\mathbf{X}_f = q(\mathbf{Y}_f)$$

Statisticians have debated the relative merits of the classical and inverse approaches (ter Braak 1995). In practice, inverse models nearly always perform as well as classical methods. Inverse models appear to perform best if the fossil assemblages are similar in composition to samples in the central part of the modern training data, whereas classical methods may be better at the extremes and under

some extrapolation as in non-analogue situations (Birks 1998).

Second, there is the question of an assumed species-response model. Some climate reconstruction procedures (e.g. MAT, ANN – see above) have no underlying assumed species response model. Multivariate calibration-function approaches assume either linear or unimodal species-environment responses (Birks 1995). It is a general law of nature that organism-environment relationships are non-linear and taxon abundance is usually a unimodal function of the environmental variable (ter Braak 1996). Each taxon grows best or is most competitive at a particular optimal value of that variable and cannot survive or is less competitive where the value of that variable is too low or too high (ter Braak 1996). Thus all taxa tend to occur over a characteristic but limited environmental range and within this range to be most abundant at or near their environmental optima (see Whittaker (1956, 1967) and Odum (1971) for further discussions of the unimodal species response model in ecology). Although species responses may be more complex than implied by the unimodal model, this model is nevertheless useful in developing numerical descriptive techniques for data showing mostly unimodal responses, just as linear models are useful in statistical analysis of data that are only approximately linear (ter Braak & Prentice 1988). Modern pollen data collected over a continental or sub-continental scale covering broad climate gradients (e.g. Williams *et al.* 2006) may show complex multimodal responses (Birks & Telford 2006). An additional problem associated with such data-sets includes the existence of multiple analogues, as discussed under The Assemblage Approach.

Third, there is the question of model dimensionality. Should all taxa be considered individually (full dimensionality) or should some axes or numerical combinations of taxa (e.g. principal components, correspondence analysis axes, partial least squares (PLS) components, weighted averaging PLS components) (reduced dimensionality) be used?

Fourth, there is the estimation procedure to consider. Should a global estimation procedure be used that estimates parametric functions such as taxa optima across the complete modern training set and thus allows some extrapolation, or should a local estimation procedure be adopted that estimates non-parametric local functions that do not allow any extrapolation?

The two multivariate calibration-function techniques that we will discuss here (two-way weighted averaging (WA) regression and calibration and weighted averaging partial least squares (WA-PLS) regression and calibration) both assume a unimodal organism-environmental response model, involve global parametric estimation, and are inverse regression procedures. They differ in using full dimensionality (WA) and reduced dimensionality (WA-PLS).

Our experience with a wide range of data-sets from many different environments suggest that there are good reasons for favouring methods with an assumed taxon response model and global parametric estimation. These include:

1. It is possible to test if a taxon has a statistically significant relationship to a particular climate variable by some form of linear or non-linear regression ana-

lysis (e.g. Lotter *et al.* 1997, 1998; Seppä *et al.* 2004; Tonello *et al.* 2009; Herzschuh & Birks 2010).

2. It is possible to generate ‘artificial’ simulated data under realistic ecological assumptions and with known numerical properties for comparative experiments on model performance (e.g. ter Braak & Juggins 1993; ter Braak *et al.* 1993; ter Braak 1995).
3. Models with assumed response models and involving global estimation procedures have clear and testable assumptions and are less of a ‘black box’ than, for example, artificial neural networks (Telford *et al.* 2004; Telford & Birks 2005).
4. It is possible to develop model evaluation or diagnostic procedures analogous to regression diagnostics in statistical modelling (e.g. Birks *et al.* 1990; Birks 1995; 1998).
5. As such methods have an underlying statistical basis, it is possible to adopt the well-established principles of statistical model selection and testing (e.g. minimal adequate model, maximum numbers of degrees of freedom, principle of parsimony) in calibration-function methodology, thereby minimising ‘ad hoc’ aspects of selecting, for example, taxon weighting functions in modern analogue procedures (Birks 1995; 1998).
6. It appears from recent work that methods involving the estimation of global parametric functions are more robust to problems of spatial autocorrelation (Telford & Birks 2005; 2009) than methods that involve local non-parametric estimation.

It is more difficult to present reasons for favouring classical or inverse approaches, as discussed above. The main reason for favouring methods that use full dimensionality is, as we will discuss later, that they are easier to fit and there is less danger of ‘overfitting’, which can easily occur in methods that can involve many PLS components unless the user is aware of the dangers of overfitting in statistical modelling and of the need to develop a ‘minimal adequate model’ (Birks 1998).

### Weighted Averaging and Weighted Averaging Partial Least Squares – Strengths and Weaknesses

The basic idea behind two-way weighted averaging (WA) (ter Braak & van Dam 1989; Birks *et al.* 1990; ter Braak 1996) is that if a taxon shows a unimodal relationship with a particular climate variable  $x$ , its abundance will tend to be highest at sites with values of  $x$  close to the taxon’s environmental optimum or niche peak. A simple and ecologically realistic estimate of the taxon’s optimum for  $x$  is thus the average of all the  $x$  values for sites at which the taxon occurs, weighted by the taxon’s relative abundance. The estimated optimum is the abundance-weighted average of  $x$  (WA regression). Taxon absences have no weight and are thus disregarded. An estimate of a site’s value for climate variable  $x$  is the weighted average of the optima for  $x$  for all the taxa present (WA calibration). The underlying theory of WA and the conditions under which it approximates the theoretically more rigorous Gaussian logit regression and

calibration are discussed by ter Braak and Barendregt (1986), ter Braak and Looman (1986), and ter Braak (1996).

WA has gained considerable popularity in many branches of Quaternary palaeoecology in the last two decades for several reasons (Birks 1998).

1. It combines ecological realism (unimodal taxon responses and the species-packing model) with mathematical and computational simplicity, sound underlying ecological and statistical theory, and good empirical power.
2. It does not assume linear taxon-environment responses, it is relatively insensitive to outliers, and, perhaps surprisingly, it is not hindered by the high inherent correlations between taxa whose abundances are expressed as ‘closed’ percentages or by the large numbers of taxa in training sets ( $\mathbf{Y}_m$ ) (Birks 1994).
3. Because of its computational simplicity, it is possible to use computer-intensive bootstrapping to derive sample-specific root mean square errors of prediction (RMSEP) for all estimates of the past climate ( $\mathbf{X}_f$ ) (Birks *et al.* 1990).
4. WA performs well in ‘non-analogue’ situations (Hutson 1977; ter Braak *et al.* 1993). In such situations, environmental reconstructions are based on the WA of the optima of the taxa in common between the modern ( $\mathbf{Y}_m$ ) and the fossil ( $\mathbf{Y}_f$ ) assemblages. As long as we have reliable estimates of the optima for the fossil taxa with high numerical importance, WA inferences are often relatively robust. This is why WA resembles more a multivariate indicator-species approach in which the indicative values of individual taxa are weighted by their abundances than a modern analogue procedure.
5. Because WA estimates global parameters for all taxa across the full sampled climate gradient ( $\mathbf{X}_m$ ), it only models the variance in  $\mathbf{Y}_m$  that is correlated with  $\mathbf{X}_m$ . It is therefore relatively robust to any spatial structure in the modern data, unlike the modern analogue technique and artificial neural networks (see above).
6. Many studies have shown that WA consistently performs well with noisy, taxon-rich compositional data with many taxa absent from many samples and extending over a relatively long climate gradient.
7. WA is relatively robust to violations of the underlying assumptions of a unimodal species-climate response (ter Braak & Verdonschot 1995; ter Braak 1996)

WA does, however, have three important potential weaknesses (ter Braak & Juggins 1993).

1. WA is sensitive to the distribution of samples within the modern training set along the climate gradient of interest (ter Braak & Looman 1986). With large datasets (>400 samples), WA has been shown to be surprisingly robust to this distributional requirement (Ginn *et al.* 2007).
2. WA ignores residual correlations in the modern biological data, namely correlations that remain in the

biological data after fitting the climate variable of interest. These residual correlations result from environmental variables not considered directly in WA.

3. WA suffers from ‘edge effects’ that lead to non-linear distortions at the gradient lengths. This problem is particularly acute for training sets with a long single dominant gradient and high compositional turnover. It leads to an overestimation of optima at the low end of the gradient and an underestimation at the high end. This in turn leads to biases in the inferred values (ter Braak & Juggins 1993).

The incorporation of partial least squares (PLS) regression (Martens & Næs 1989) into WA to create WA-PLS (ter Braak & Juggins 1993) helps to overcome, in part at least, the second weakness of WA by utilising residual correlations to improve the estimates of taxon climate parameters and the third weakness, namely the ‘edge-effect’ problem.

ter Braak and Juggins (1993), ter Braak *et al.* (1993), ter Braak (1995), and Birks (1995; 1998) describe the mathematical basis of WA-PLS and this is not repeated here. The main advantage of WA-PLS is that it often, but not always, produces calibration-function models with lower RMSEP and bias than simple two-way WA. There are two possible reasons for this improved performance.

1. WA-PLS reduces the ‘edge-effect’ problem. Due to the ‘edge effect’ in WA, there is not only a bias in the inferred values of the climate variable but also in the model residuals at the ends of the gradients. WA-PLS implicitly involves a weighted inverse deshrinking regression that pulls the inferred values towards the training-set mean. WA-PLS utilises both patterns in the residuals to update the calibration function, thereby reducing errors, and patterns in the systematic bias to reduce, to some extent, the ‘edge-effect’ problem.
2. In real life, there are often additional environmental variables that may influence the biological assemblages. WA ignores any structure resulting from these variables and assumes that environmental variables other than the one of interest have negligible influence, or that their joint distribution with the climate variable of interest in the past is the same as today (assumption 5 above). WA-PLS uses this additional structure to try to improve estimates of the taxon parameters in the final calibration function. For optimal performance, the joint distribution of these environmental variables in the past should be the same as in the modern data (ter Braak & Juggins 1993).

The main disadvantage of WA-PLS compared to WA is that great care is needed in model selection. As more components are added, the WA-PLS model appears to fit the modern data better and better as the model error decreases and becomes 0 when the number of components fitted equals the number of samples. Such a model, the so-called full model, has zero predictive power as it has no degrees of freedom. An independent test-set is needed to evaluate different WA-PLS models based on 1, 2, ...,  $n$  components

where  $n$  is usually 4-10, and thus to find the optimal model with the lowest prediction error and bias for the test-set. In practice, there are usually no independent test-sets and WA-PLS model evaluation has to be based on cross-validation to derive approximate estimates of model performance. The final WA-PLS model to use in climate reconstruction is selected on the basis of low RMSEP, low bias, and a small number of ‘useful’ components (‘minimal adequate model’) (Birks 1998). In some cases, the number of ‘useful’ components (defined by Birks (1998) as giving >5% reduction in the RMSEP of the one-component WA-PLS model) does not guard against selecting an overfitted model. A simple solution to avoid overfitting is to use van der Voet’s (1994) randomisation  $t$ -test to test the equality of predictions from two transfer-function models (see Racca *et al.* (2001) and Velle *et al.* (2011) for applications of this test). Sample-specific prediction errors for past climate variables can be estimated by Monte Carlo simulation or by bootstrapping (Birks *et al.* 1990; Birks 1995, 1998).

A further problem with WA-PLS is that as it exploits residual structure in the data, which is often related to the effects of environmental variables other than the one of interest. WA-PLS can, in theory, be more susceptible to biases related to different covariations of environmental variables in the past than occur in the modern training set. It is not uncommon to find discrepancies between reconstructions based on WA-PLS when two or three components are used, even though the ‘optimal’ model (based on model performance statistics) suggest a two- or three-component model is appropriate. It is also important that the residual structure used by WA-PLS is ecologically useful. We have worked with several pollen and chironomid data-sets for which additional WA-PLS components improved the performance statistics of the *modern* training set but when these additional components are applied to fossil stratigraphic data, the components simply increase the noise in the structure without showing more climatically-relevant structure or strengthening the signal relative to the noise in the reconstruction.

Examples of situations in which WA-PLS may lead to misleading results are training sets in which the relationship between climate and modern proxy-data is affected by non-climatic processes acting to different extents along the observed environmental gradients. For example, it is very likely that the strong relationship observed between chironomid assemblages and summer temperature is mediated and reinforced to a significant extent by climate effects on water chemistry parameters (e.g. nutrient and oxygen concentrations) (Brodersen & Quinlan 2006; Velle *et al.* 2010; Verbruggen *et al.* 2010). If, in modern training sets, these water chemistry parameters are affected in parts of the gradient by processes unrelated to climate, such as human activity or acidification, it is possible that WA-PLS will correct for any resulting offset in inferred values within a developed chironomid-based calibration function for temperature. However, this would result in a correction of an offset that is unrelated to climatic effects on chironomid assemblages and could potentially lead to a bias in the reconstructions obtained from fossil assemblages unaffected by confounding factors important in the modern training set. Other examples of potentially problematic situations for

WA-PLS might include training sets in which the relationship between an environmental parameter and assemblage data is driven by different seasonal components of the parameter of interest in different parts of the observed environmental gradient (e.g. winter temperature versus summer temperature as a dominant parameter affecting plant assemblages in different climatic conditions) or training sets in which the environmental gradient of interest covers major transitions in bedrock type which might influence species distributions (e.g. calcareous versus non-calcareous bedrock).

In a series of numerical experiments using a 191- and 310-sample modern pollen-climate data-set from Norway, northern Sweden, and Svalbard, Bjune *et al.* (2010) showed that WA-PLS model performance statistics were only slightly improved by increasing the size of the training set. However, an important result was that down-core reconstructions of mean July temperature showed much less sample-to-sample variation ('noise') when the larger training set was used. This emphasises the importance of considering not only modern model statistics but also the resulting climate reconstructions.

Our general experience is that all model performance statistics are only a *rough* guide to model selection. The problem is that all these performance statistics are based on predicting temperature from modern assemblages with known climate. The statistics do not indicate how reliable the models are when they are applied to fossil assemblages where the past climate is not known. Some careful and critical two-way interaction is needed between assessing modern model performance statistics and the resulting down-core reconstruction (see Velle *et al.* 2010).

Like WA, WA-PLS performs surprisingly well and considerably better than modern analogue procedures when none of the fossil assemblages are similar to the modern assemblages (ter Braak *et al.* 1993; ter Braak 1995). For very strong extrapolation beyond the modern training set, WA often performs better than WA-PLS. Like WA, WA-PLS is a weighted multivariate indicator-species approach but where all taxa are used in the reconstruction and estimates of the relevant taxon parameters are derived from the modern training data rather than from modern autecological observations or biogeographical data. Telford and Birks (2005) show that, unlike WA, WA-PLS can produce biased results and overly optimistic prediction errors in the presence of strong spatial autocorrelation. This occurs because, besides reducing the predictive error by reducing the 'edge-effects' inherent in WA, WA-PLS may acquire residual variation in certain parts of the environmental gradient that can be spatially autocorrelated to the climate variable of interest.

One approach to eliminate the problem of model selection and model overfitting in WA-PLS and to reduce the 'edge-effects' of WA is to use a monotonic smoothing spline (Wold 1992; ter Braak & Juggins 1993; Telford *et al.* in prep) rather than the conventional linear inverse or classical deshrinking regression in WA (Birks *et al.* 1990; Birks 1995). Comparisons using several diatom-environment data-sets (Telford *et al.* in prep) indicate that a simple WA model with a monotonic spline often has the best predictive power,

and produces the simplest model ('minimal adequate model') in terms of number of parameters to estimate and to fit.

### Modern Ecological Applications of Multivariate Calibration Functions

In the previous sections, we have seen how palaeoecologists, ecologists, and biogeographers are independently considering the same or related problems in the indicator-species approach and the assemblage approach. Is there a similar parallelism in the multivariate calibration-function approach? The answer is yes.

Plant ecologists have long been using modern plant assemblages as a basis for assessing modern environmental conditions, so-called bio-indication, by means of ecological indicator values (e.g. Ellenberg 1948; Ellenberg *et al.* 1992). Ordinal environmental indicator values ('optima') for the central European flora (Ellenberg *et al.* 1992) and the western European flora (Hill *et al.* 2000; 2004; 2007) have been estimated on the basis of field, laboratory, and phytosociological studies for light, moisture, soil reaction, nitrogen, temperature, and salt tolerance. Given a modern assemblage, the environmental value of the sites is inferred to be the average of the species indicator values (if presence/absence data only are available) or the abundance-weighted average of the indicator values if quantitative abundance data are available. Similar index-based approaches have been developed and are widely used for biomonitoring of freshwater ecosystems. Examples include the Benthic Quality Index (Wiederholm 1980) based on abundance-weighted averages of indicator values for selected chironomid species or abundance-weighted indices reflecting past and present water quality based on diatom assemblages (e.g. Renberg & Hellberg 1982). Diekmann (2003) provides a thorough review of species indicator values as a tool in applied ecology, and additional excellent examples of their use include Dzwonko (2001), Hawkes *et al.* (1997), Hill and Carey (1997), and Holtland *et al.* (2010). The underlying mathematical basis for the use of indicator values is outlined by ter Braak and Barendregt (1986) and ter Braak and Gremmen (1987).

Recently, Brady *et al.* (2010) have used exactly the same multivariate calibration-function approach in palaeoclimatology of WA-PLS to develop modern climate-plant calibration functions based on 107 forest plots in Oregon and Washington. Their aim is to revisit and record the plots in the future and to use their modern calibration function to infer what climatic changes there have been. Such an approach could be of great value in remote arctic or alpine areas where there are very few meteorological stations. However, developing a modern calibration function in such areas is not possible without reliable climate data. It will be interesting to see, under conditions of future climate change, how robust and how useful the approach of Brady *et al.* (2010) will be.

We now turn to topics common to all approaches – presentation and interpretation, evaluation and validation, comparisons, and general limitations of climate reconstructions. In addition we discuss future developments and challenges.

## PRESENTATION AND INTERPRETATION OF CLIMATE RECONSTRUCTIONS

In palaeoclimatology, it is common to present reconstructions or model hindcasts as climate ‘anomalies’ or ‘deviations’ from present-day values (e.g.  $\Delta = -1.5^{\circ}\text{C}$ ,  $\Delta = 2.1^{\circ}\text{C}$ ). There are several options in the presentation of climate reconstructions based on biological assemblages using the three main approaches described above:

1. as the inferred or reconstructed value in  $^{\circ}\text{C}$ , mm of precipitation, etc (e.g. Seppä & Birks 2001)
2. as deviations ( $\Delta$ ) from the observed present-day value at the study site
3. as deviations ( $\Delta$ ) from the inferred present-day value at the study site (Heikkilä & Seppä 2010)
4. as deviations ( $\Delta$ ) from the mean of all reconstructed values for the last 100, 200, or 250 years (e.g. Heikkilä & Seppä 2003; Seppä & Poska 2004), or
5. as deviations ( $\Delta$ ) (possibly standardised to unit variance, so-called ‘z-scores’) from the mean of all reconstructed values for the sequence under study (Osborn & Briffa 2006).

Despite the considerable advances in techniques for reconstructing past climate from fossil assemblages, much work is needed on how to present optimally reconstructed climate temporal-series and how to detect consistent trends and patterns within and between such temporal-series (Birks & Seppä 2004). The simple threshold-count approach of Osborn and Briffa (2006) is a promising and robust procedure for comparing different climatic temporal-series after standardisation.

Expression of reconstructions as deviations from the modern inferred values may have the advantage that the deviations accurately reflect past changes in temperature even when absolute inferences are consistently too high or too low. This may be the case if the local thermal regime in a region or ecosystem is not adequately reflected by or in equilibrium with the thermal variable that is reconstructed. Examples are reconstructions of past air temperature based on lacustrine proxy-indicators (e.g. chironomids, diatoms, cladocerans) from lakes which receive considerable amounts of cool water via tributaries, pollen-based reconstructions from landscapes affected by higher or lower than expected insolation (e.g. south- or north-facing slopes) or, for WA and WA-based methods, reconstructions which infer values close to the edges of the examined temperature gradient and are therefore potentially affected by ‘edge effects’.

Despite considerable advances in multivariate calibration-function methodology and in developing organism-climate training sets, our abilities to interpret and compare time-series of palaeoclimate reconstructions have hardly developed beyond visual comparison of time-series. Willis *et al.* (2007) illustrate a statistically rigorous way of comparing two or more palaeoecological biological and environmental time-series using auto-regressive procedures.

Although it is possible, thanks to improved computing power and efficient programming, to derive sample-specific errors of prediction for climate reconstructions based on WA

or WA-PLS using computer-intensive bootstrapping procedures, an unresolved problem is how to interpret such errors. All stratigraphical data and resulting reconstructions exhibit strong temporal autocorrelation and thus adjacent samples and their reconstructed values of  $X_f$  are not independent of each other statistically. Related to this problem is our limited ability to compare rigorously temporal series of environmental reconstructions, again because of the statistical problems created by temporal autocorrelation (Clarke 1989).

All biostratigraphical data and hence all climate reconstructions which are, in reality, a transformation (admittedly in many cases a complex, non-linear transformation) of the original stratigraphical data are rather ‘noisy’ due, in part, to the inherent sample-to-sample variation in the biostratigraphical data. Non-parametric regression such as locally weighted regression smoothers (LOESS) provide useful graphical tools for highlighting ‘signal’ or major patterns in temporal series of reconstructed climate variables (Birks 1998). LOESS (Cleveland 1993) models the relationship between a response or dependent variable (e.g. pollen-inferred July temperature) and a predictor or independent variable (e.g. age) when no single functional *a priori* form, such as a linear or quadratic model, is appropriate. LOESS provides a robust graphical summary that helps assess the relationship and detects major trends within ‘noisy’ data such as climate reconstructions. LOESS can also be used to produce a ‘consensus’ reconstruction based on several reconstructions (e.g. different transfer-function models, different data-sets, different records, different proxies) (Birks 1998). More sophisticated smoothers such as SiZer (significance of zero crossing of the derivative) (Chaudhuri & Marron 1999; Holmström & Erästö 2002) and the related BSiZer (Erästö & Holmström 2005) have considerable potential in palaeoclimatology because they can assess which features seen in a range of smoothed data are statistically significant and thus may be environmentally significant (see Korhola *et al.* 2000; Erästö & Holmström 2005, 2006; Weckström *et al.* 2006 for palaeoecological applications). Given the random walk simulations of pollen data presented by Blaauw *et al.* (2010), there is an increasing need to use techniques such as SiZer or BSiZer to distinguish signal from random red noise in stratigraphical time-series.

## EVALUATION AND VALIDATION OF CLIMATE RECONSTRUCTIONS

Evaluation and validation of palaeoecological climate reconstructions receive surprisingly little attention. They are extremely important, as all reconstruction procedures will provide results irrespective of which method or approach is used. The critical question is how reliable are these results?

There are two general approaches – numerical evaluation using reconstruction diagnostic statistics and validation using historical data and instrumental records or independent palaeoclimatic data. In terms of numerical evaluation there are four useful numerical criteria (Birks 1998; Birks 2003; Juggins & Birks 2011). These are only applicable to the assemblage and the calibration-function approaches.

1. Sample-specific RMSEP for individual fossil samples (Birks *et al.* 1990).

2. 'Goodness-of-fit' statistics assessed by fitting fossil samples 'passively' onto the first axis of an ordination of the modern training data constrained by the climate variable being reconstructed and evaluating how well individual fossil samples fit onto this constrained axis in terms of their squared residual distance (Birks *et al.* 1990).
3. Analogue measures for each individual fossil sample in comparison with the modern training set (Simpson 2007, 2011). A reconstructed climate variable is likely to be more reliable if the fossil sample has modern analogues within the training set (ter Braak 1995).
4. The percentages of the total fossil assemblage  $Y_f$  that consist of taxa that (a) are not represented at all in the training set  $Y_m$  or (b) are poorly represented (e.g. only present at low abundance, only present in a few samples) in the training set. Parameters summarising the ecology of these taxa (e.g. WA optima) are more likely to be poorly estimated and have large standard errors associated with them in cross-validation (Birks 1998).

Bigler *et al.* (2002), Heiri *et al.* (2007) Engels *et al.* (2008, 2010), and Lotter *et al.* (2011) illustrate the use of these numerical evaluation criteria in studies on climate change using a range of proxies. These numerical criteria are, however, only a *guide* to the possible reliability of the reconstructions. Even samples with poor evaluation statistics (e.g. no close analogues, a poor fit with the climate variable of interest) may yield accurate reconstructed values, whereas samples with favourable evaluation statistics may be biased, for example if the basic assumptions behind the numerical method used are violated or if the biological system of interest is strongly influenced by environmental or biotic variables unrelated to climate. Examples of palaeoecological reconstructions yielding results in agreement with independent climate records even in instances where evaluation statistics indicate a poor fit with temperature and no close or good analogues include Heiri *et al.* (2003, 2007) and Ilyashuk *et al.* (2010). Not surprisingly the likelihood of encountering non-analogue situations increases with the age of the fossil assemblages (e.g. Birks *et al.* 1990, Heiri *et al.* 2003; Heiri & Millet 2005) and during periods of major climatic change such as the late-glacial period (e.g. Heiri & Millet 2005; Heiri *et al.* 2007). In these instances palaeoecological reconstructions of past climate must be carefully evaluated using other, independent climate records and checked for their palaeoecological and palaeoclimatological plausibility.

It is important to note that although numerical evaluation criteria can help in identifying non-analogue situations, their interpretation is often problematic. Numerical evaluation should therefore be seen not as an end in itself but as a source of information to supplement an interpretation based on an understanding of the ecological processes underlying the observed changes (Juggins & Birks 2011). Korhola *et al.* (2002) provide an elegant example of the use of both numerical and ecological criteria to help to evaluate a chironomid-based climate reconstruction in northern Fennoscandia.

What is required for all these reconstruction approaches is rigorous validation through comparisons with independent palaeoclimate evidence such as stable isotopes analysed in ice-core or tree-ring records, geomorphological evidence (e.g. glacier fluctuations, occurrence of permafrost), or other biological proxies (i.e. a multi-proxy study, Birks & Birks 2006). Work in Norway (S.J. Brooks and H.J.B. Birks, unpublished) in which pollen-based and chironomid-based calibration functions are applied to reconstruct mean July air temperature in the Holocene shows that the pollen-based and chironomid-based reconstructions can sometimes give very different results for the same stratigraphical sequence, with the chironomid-based reconstruction tending to give consistently lower temperature values than the pollen-based climate reconstruction. The critical question is thus, which reconstruction is the more reliable? Additional means of independent validation are needed to help decide which reconstruction should be accepted. One approach is to use plant macrofossil evidence as proof of local presence of, for example, tree *Betula* or *Pinus sylvestris*. On the basis of modern ecological observations (the indicator-species approach), their local presence can set lower limits (e.g. mean July temperature of 11°C or more) for the climate at the times when their macrofossils occur. Another approach is to use the probability density function approach of Kühl *et al.* (2002) for taxa that occur as macrofossils. A novel and as yet untried approach is to divide the fossil pollen assemblages into the numerically abundant and frequent taxa that comprise 75-90% of the assemblage and to use these to reconstruct the past climate using calibration functions and to use the excluded rare and numerically low taxa as 'indicator' taxa for validation purposes in a similar way to the use of plant macrofossils outlined above.

An important question related to the comparisons of chironomid-based and pollen-based climate reconstructions is why should chironomid-based temperatures be lower than the pollen-based temperatures? There are several possible reasons that relate to some of the basic assumptions and limitations of calibration functions.

1. Pollen-climate calibration functions are based on an indirect relationship (Birks & Seppä 2004). Modern pollen is a function, admittedly a complex and non-linear function, of modern vegetation. Modern vegetation is a function of modern climate, at least at a broad scale such as Fennoscandia. There is thus an indirect link between modern pollen and modern climate. An assumption of modern pollen-climate calibration functions is that the bulk of the pollen assemblage comes from the relevant pollen source area of the site and is thus indirectly related to the modern climate at the site. Long-distance dispersed 'extra-regional' pollen can be blown into sites from areas at lower altitudes (e.g. Ortu *et al.* 2006, 2010) or further south. This is a major and well-known problem in pollen analysis of sites in northern, alpine, and arctic areas. Plant macrofossils are a valuable guide to local species presence (Birks & Birks 2000). However, the absence of macrofossils does not mean that the species was absent locally. The net result of far-distance pollen on pollen-based climate reconstructions is that the resulting climatic estimates may be higher than the actual values. This effect is well

- illustrated today by the high maximum bias in the modern pollen-climate models at sites with low summer temperatures today that are above or beyond the present-day tree-line in Norway (Bjune *et al.* 2010).
2. All empirical calibration functions are, by necessity, based on estimates of the parameters (e.g. optimum) of the realised Grinnellian niche or the regeneration niche today (e.g. Smart *et al.* 2010). In the past, the realised niches of taxa may have been different for a variety of ecological, environmental, and historical reasons (Jackson & Overpeck 2000; Jackson & Williams 2004) and thus assumption 3 of calibration functions and other climate reconstruction approaches, namely that the ecological responses of taxa today have not changed over the time represented by the fossil assemblages, may not be valid. Examples suggesting changes in realised niches are provided by comparative studies of reconstructed and simulated Holocene range shifts of *Picea abies* and *Fagus sylvatica*, two common tree species in northern and central Europe. Comparisons of their reconstructed Holocene spread patterns with their modelled range shifts show distinct divergences and thus highlight the difficulties in explaining the species' Holocene range history on the basis of their modern realised climatic niches (Giesecke *et al.* 2007; Miller *et al.* 2008).
  3. An important bias in many modern pollen-climate training sets may result from the fact that much of the Earth's land surface is covered by vegetation that is modified or even produced by human activity, especially in the last few hundred years. In northern Europe, the effects of human impact would generally decrease the percentage of *Quercus*, *Ulmus*, and *Tilia*, all of which are associated with high temperatures, and increase the percentages of non-arboreal pollen such as Poaceae, Ericaceae, *Juniperus*, and *Artemisia*, all of which (in the absence of human activity) are associated with low temperatures (Seppä *et al.* 2004). Holocene climate reconstructions may therefore be difficult to evaluate, especially when there is independent evidence for human activity in the study area and where there is extra-regional far-distance pollen. Climate reconstructions based on late-glacial pollen assemblages with high values of *Artemisia*, Poaceae, and other non-arboreal pollen types have a high chance of giving misleading temperature values, because much of the pollen may be extra-regional. St Jacques *et al.* (2008) develop a pollen-climate calibration set for Minnesota based on 133 pollen assemblages that pre-date European settlement, as well as a set based on modern pollen assemblages. Their study shows serious biases in climate reconstructions based on modern assemblages and less bias when the pre-European assemblages are used.
  4. Although calibration functions for the same climate parameter can be developed for several groups of proxy-indicators, this does not guarantee that these organism groups are necessarily responding to the same aspect of present and past climate. For example, seasonal temperatures and precipitation are strongly

intercorrelated in many available training data-sets. Similarly, a parameter such as summer air temperature will usually be correlated with other thermal parameters representing summer climate such as growing degree days, growing season length (Jackson *et al.* 2009), and, in the case of lake-living indicator groups such as chironomids, cladocerans, or diatoms, water temperature. In chironomid-July air temperature calibration functions, we are assuming that chironomid distribution and abundances today are a function of air temperature, whereas ecologically they are probably in part, or even to a large extent, a function of water temperature. Although there is a monotonic relationship today between air and water temperature (Livingstone & Lotter 1998), this relationship may not have been the same in the past due, for example, to changes in winter precipitation, leading to more and longer lasting snow presence at high elevation or high latitude sites, or due to changes in insolation and cloud abundance which may have affected the thermal regime of lakes. Lower chironomid-based estimates of mean July air temperature compared to pollen-inferred values may thus be a result of the chironomid-July air temperature calibration function being based on an indirect rather than a direct ecological relationship. Similarly, relationships between seasonal temperatures or thermal parameters may have been different in the past than they are at present. Lotter *et al.* (2011) discuss different hypotheses for explaining the offset between chironomid- and pollen-based July air temperature inferences from the sediments of Gerzensee, a small lake on the Swiss Plateau. Both proxies perform well in reconstructing temperatures in the modern environment and produce similar reconstructions during the earliest part of the late-glacial period but diverge in the latter part of the late-glacial interstadial between ca. 13,000 and 13,800 calibrated <sup>14</sup>C years BP. One possible reason for this divergence is that vegetation and chironomid assemblages were responding to different aspects of late-glacial climate which are strongly correlated today but were not necessarily correlated in the late-glacial. Lotter *et al.* (2011) discuss that late-glacial vegetation may have been strongly affected by climatological parameters such as seasonality, winter temperature, or hydrological changes which are expected to have had a lower impact on chironomid distribution and survival.

A general observation from many multi-proxy studies is that while the different climate reconstructions usually agree in terms of highlighting the major trends and change-points, there can be substantial disagreement in the absolute values of the reconstructions between proxies (Birks *et al.* 2000). This is, in some ways, inevitable as different biological proxies have different sensitivity to the climate variable of interest and to any confounding variables. In this way, a partial validation of the basic proxy records follows from an attempt to quantify, understand, and explain the biases inherent in each proxy (Birks & Birks 2006; Juggins & Birks 2011).

Table 2 summarises some of the major problems that can arise in climate reconstructions from biological assemblage

**Table 2. Some Problems that can Arise in Climate Reconstructions from Biological Assemblage Data, the Detection of these Problems, and Potential Solutions**

Problem	Detection	Potential Solution
High amount of noise in reconstruction	Visual examination of time series; calculation of measures of dispersion for time series	Smoothing using locally weighted regression; use simpler model to reduce overfitting
Systematic bias	Comparison with other quantitative reconstructions	Express reconstructed values as differences from mean or modern reconstructed values; use WA with classical deshrinking
Non-analogue situations	Reconstruction diagnostic statistics and evaluation methods	None; try WA or WA-PLS
Quality and availability of instrumental data	Numerical analysis of climate data	Use gridded data and robust interpolation procedures; assess errors associated with interpolation
Variable other than the one of interest driving the biotic changes	Comparison with independent physical, chemical, or biological proxies	None
Poor model performance statistics	High RMSEP, bias, etc	Careful numerical and ecological examination of data
Large sample-specific errors of reconstruction	Large error bars	Examine reconstruction diagnostics and test for non-analogue assemblages

**Table 3. Some Errors that can Arise (and have Arisen) in Developing Modern Organism-Environment Calibration Data-Sets and Possible Solutions**

Errors	Possible Solutions
Modern data from different sedimentary environments with contrasting taphonomies	Data screening
Not all samples are modern samples	Data screening and examination of meta-data
In very large data-sets, same samples may be duplicated by accident	Data screening
Inconsistent taxonomy or low taxonomic resolution	Analytical quality control and recounting
Inconsistent and/or poor environmental data	Data screening, new data
Model overfitting solely to minimise RMSEP	Van der Voet (1994) randomisation test, more rigorous cross-validation
Ignoring spatial autocorrelation in assessing model performance	Telford & Birks (2009) deletion and <i>h</i> -block methods
No estimates of sample-specific errors of reconstruction	Bootstrapping
Model performance not based on cross-validation. Apparent statistics only	Use cross-validation (leave-one-out, split sampling, <i>n</i> -fold cross-validation)
No numerical evaluation of reconstruction	Apply reconstruction diagnostic measures
No validation of reconstruction using independent records	Not always possible but literature or internet searches can be useful
Reconstructing 2 or more variables that may, in reality, not be reconstructable	Careful use of partial constrained ordinations, consider not only marginal effects but also conditional effects (Juggins & Birks 2011), simulations and randomisation tests

data irrespective of the reconstruction method used. The table lists how the problems can be detected and suggests possible solutions. Table 3 lists twelve potential errors that can arise (and have arisen) in developing modern organism-environment calibration data-sets based on our own experiences and those of colleagues. Possible solutions are also given in Table 3. Within the vast literature about calibration data-sets, some papers critically identify and discuss in detail some of these potential errors (e.g. Birks *et al.* 1990; Jones & Juggins 1995; Lotter *et al.* 1997, 1998, 2011; Engels *et al.* 2001, 2008; Bigler *et al.* 2002; Whitmore *et al.* 2005; Finkelstein *et al.* 2006; Heiri *et al.* 2007; Williams & Shuman 2008; Gonzales *et al.* 2009a; Juggins & Birks 2011).

### COMPARISON OF CLIMATE RECONSTRUCTIONS USING DIFFERENT METHODS

The three general approaches (Table 1) and their numerical methods discussed above are all different mathematically and ecologically and they model (implicitly or explicitly) species-climate responses in different ways. These methods should be viewed as complementary and not competitive (Racca *et al.* 2001), especially when there is no clear 'winner' in terms of modern model performance statistics (Juggins & Birks 2011). In these cases, it is useful to compare reconstructions based on a range of techniques. If down-core biological changes are primarily driven by changes in the reconstructed climate variable, one would

expect reconstructions from different models to follow similar trajectories, even if they differ in their absolute values (Birks & Ammann 2000). If different methods produce widely different reconstructions, this suggests that fluctuations in the dominant taxa are not primarily related to changes in the reconstructed variable and the reconstruction must be treated with caution (Juggins & Birks 2011). Similarity in reconstructions based on different methods but the same proxy data does not imply rigorous validation but indicates that the reconstructions are free from method-specific bias (Juggins & Birks 2011). See Walker *et al.* (1997), Lotter *et al.* (2000), Birks (2003), and Köster *et al.* (2004) for examples of such comparisons. When different reconstruction methods produce consistent rather than conflicting reconstructions, these can be combined into a single consensus reconstruction (e.g. Bartlein & Whitlock 1993; Racca *et al.* 2001; Birks 2003; Barrows & Juggins 2005).

Kucera *et al.* (2005) present a conceptual model for assessing the reliability of palaeoclimate reconstructions based on a scatter-plot of dissimilarity  $d$ , against model divergence  $\Delta$ , represented by the standard deviation of the different estimates for each fossil sample. Samples with low  $d$  and  $\Delta$  have convergent estimates and good analogues and are considered reliable. Samples with low  $d$  but high  $\Delta$  also have good analogues but suffer from method-specific bias. Although one is faced with the problem of defining critical values of  $d$  and  $\Delta$ , this simple method provides a tool that focuses attention on the interpretation of the differences between reconstructions based on different methods (Juggins & Birks 2011). This approach warrants wider use in palaeoecology.

## NEW APPROACHES TO CLIMATE RECONSTRUCTION

The most successful and widely used approaches to reconstruct past climate changes based on fossil pollen and chironomid assemblage data strongly rely on empirical relationships between assemblages and climatic parameters observed in the modern environment. However, empirical relationships and reconstruction models are not necessarily based on detailed ecological understanding of the mechanisms leading to a relation between fossil and subfossil assemblages and climate variables of interest. Furthermore, they do not usually include an explicit modelling of cause-and-effect relationships between environmental variables and taxon presence and abundance. Forward modelling of proxy responses to environmental change includes the establishment of a model describing the state (e.g. presence, abundance) of the proxy-indicator of interest considering a certain climate state (Haslett & Challenor 2010). Climatic reconstruction is then achieved by inverting the model and assessing the likelihood of different climatic conditions given a certain state (e.g. assemblage composition) of the proxy-indicator (e.g. Guiot *et al.* 2000, 2009). For example, Guiot *et al.* (2000) used a mechanistic model of vegetation expressed as a function of climate. By inverting the model under the constraint that the model output 'fits' the pollen data, Guiot *et al.* (2000) reconstructed past climate under various environmental conditions such as lowered atmospheric CO<sub>2</sub> concentrations. In contrast to the indicator-spe-

cies, assemblage, or multivariate calibration-function approaches described above, this forward modelling approach has a number of advantages. First, as in the Bayesian approaches briefly discussed below it allows the modelling of different sources of variability, whereas the assemblage approach and the multivariate calibration-function approach typically only provide a summary estimate of the reconstruction error. Second, since ecological processes are modelled explicitly, changes in species or assemblage response to changing environments over time can be explicitly included as well. An example includes the study of Wu *et al.* (2007a, 2007b), who use an inversion procedure based on a vegetation model to assess and take into account the effects of different atmospheric carbon dioxide concentrations on plant distribution when developing pollen-based climate reconstructions. Third, multiple proxies and/or multiple sites can be used simultaneously to produce climate reconstructions. The procedure then looks for the most likely climate state considering all the available models and the available proxy data. Fourth, such models are potentially very flexible. If taxa respond to different environmental parameters in different parts of the environmental or geographical gradients examined, this can be taken into account in the modelling approach. Fifth, observations under all environmental conditions are not needed, thereby enabling climate reconstructions for non-analogue conditions. The main disadvantages of this approach are that much more information is needed about the biological system of interest than for the indicator-species, assemblage, or multivariate calibration-function approaches, that many more assumptions are made about the proxy at hand, about the reliability of the applied models, and about the quality of the available proxy data, and that these modelling approaches are very demanding in respect to computing time. Such approaches based on forward modelling have been increasingly used in recent years using Bayesian statistics (e.g. Guiot *et al.* 1999, 2008, 2009; Haslett *et al.* 2006; Hatté *et al.* 2009; Haslett & Challenor 2010) and it is expected that their importance and relevance will continue to increase with increasing computational resources and our understanding about the relationships between climate parameters and proxy responses.

## GENERAL LIMITATIONS OF CLIMATE RECONSTRUCTIONS

The major strengths and weaknesses of the four most commonly used climate reconstruction techniques within the indicator-species approach, of the three most commonly used techniques within the assemblage approach, and of the two most commonly used calibration-function techniques are summarised in Table 4.

The key requirement in all climate-reconstruction methods is the need for high quality, internal consistency of the modern and fossil data-sets. Such data-sets require a detailed and consistent biological taxonomy and, for modern data-sets, reliable and representative climate data. The generation of modern training sets with detailed and consistent taxonomy in the multivariate calibration-function and assemblages approaches ideally requires that all the biological analyses are done by the same analyst who should be skilled in the relevant taxonomy. Many training sets covering broad geographical areas and climate gradients have, by necessity,

Table 4. Strengths and Weaknesses of the Climate Reconstruction Techniques Discussed in this Essay

Approach and Method	Strengths	Weaknesses
<b>Indicator-species approach</b>		
<i>Bio-climate modelling</i> (e.g. Iversen 1944; Conolly & Dahl 1970)	Simple, given distributional and climate data. Good for macrofossils	Presence/ absence data only
<i>Probability density functions</i> (Kühl <i>et al.</i> 2002)	Probability of occurrence within climate space considered; estimates of uncertainty possible. Good for macrofossils and beetles	Presence/ absence data only
<i>Mutual climate range method</i> (Atkinson <i>et al.</i> 1987)	Simple, given reliable distributional and climate data	Presence/ absence data only; assumes uniform probability of occurrence in climate space; wide ranges; no means of deriving model performance statistics; acquiring climate data at appropriate scale difficult for beetles and alpine plants
<i>Modified mutual climate range method</i> (Klotz <i>et al.</i> 2003; 2004)	Simple but elegant way of reducing estimated climate ranges; estimates of uncertainty possible	Presence/ absence data only
<b>Assemblage approach</b>		
<i>Modern analogue technique</i> (Overpeck <i>et al.</i> 1985; Simpson 2007, 2011)	Simple; classical; no assumed response model; deals with quantitative data; ecological plausible	Very sensitive to spatial autocorrelation; local fitting; data demanding; problems of finding appropriate numbers of analogues
<i>Response surfaces</i> (Prentice <i>et al.</i> 1991; Bartlein & Whitlock 1993)	Classical; no assumed response model; deals with quantitative data	Difficult to derived unbiased estimates of model performance; local fitting
<i>Artificial neural networks</i> (e.g. Malmgren <i>et al.</i> 2001; Kucera <i>et al.</i> 2005)	No assumed response model; deals with quantitative data	Easy to overfit; very sensitive to spatial autocorrelation; 'black box'; impossible to interpret
<b>Multivariate calibration-function approach</b>		
<i>Weighted averaging (WA)</i> (ter Braak & van Dam 1989; Birks <i>et al.</i> 1990)	Assumed unimodal response model; ecologically plausible; simple and robust; can extrapolate; robust to autocorrelation; easy to compute prediction errors; global parameter estimation as optima; quantitative data	Edge effects; inverse; ignores other environmental gradients; sensitive to sample distribution in modern data
<i>Weighted averaging partial least squares (WA-PLS)</i> (ter Braak & Juggins 1993)	Assumed unimodal response model; can extrapolate; relatively robust to spatial autocorrelation; global parameter estimation; quantitative data	Edge effects; overfitting easy; inverse; reduced dimensionality; species coefficients not readily interpretable

to be constructed from samples analysed by different analysts. In such cases, taxonomic workshops, standardisation of methodology, quantitative analytical quality control, and agreed taxonomic and nomenclatural conventions are essential. Such harmonisation between data-sets and analysts is time-consuming and unattractive and as a result it is under-funded or even bypassed as national, continental, global data-bases rapidly develop (Lee 2000; Bortolus 2008). Even less attention is given to the quality and representativeness of the modern climate data used. Particular problems arise in deriving reliable site-specific modern climate data in mountainous and high-latitude areas (e.g. Lundquist & Cayan 2007; Pape *et al.* 2009; Tabor & Williams 2010) and standardised procedures of interpolation, lapse-corrections, etc., must be used throughout. Many interpolated modern climate values do not have reported uncertainties or, if reported, are ignored in developing modern organism-climate reconstruction models. In some situations it is difficult to see how to improve the modern climatic data used, given the limited availability and quality of modern climate data in some geographical areas and the complex patterns of climate variability over small areas, especially

those with complex topography (e.g. Lundquist & Cayan 2007; Tabor & Williams 2010).

An important requirement in training-set development is that the modern data should be representative of the likely range of past climate variables (ter Braak 1995). Besides having consistent taxonomy and nomenclature and being of comparable quality, the modern biological data should be from the same sedimentary environment (e.g. lakes of similar size) as the fossil data-sets used for reconstruction purposes (Herzschuh *et al.* 2010a). Recent developments in modelling pollen-source areas (e.g. Prentice 1985; Sugita 1994; Jackson & Lyford 1999) have shown very clearly that moss polsters, small lakes, and large lakes all have very different relevant pollen-source areas (e.g. Zhao & Herzschuh 2009). Pollen spectra from soils, peat bogs, sediments from small enclosed lakes, and sediments from large lakes will also have different taphonomies. Similarly, it can be expected that the relationship of assemblages of lake-dwelling organisms (e.g. diatoms, cladocerans, chironomids) with temperature will be different in shallow, small lake basins compared with deeper, stratified lakes. It is essential to eliminate, when developing modern training sets to use

with fossil data-sets, these sources of variation as far as possible (Zhao *et al.* 2009). Fossil pollen spectra from small lakes can only be reliably compared, whether in the analogue approach or the calibration-function approach, with modern pollen spectra from similar types of lakes. Climate reconstructions from fossil assemblages clearly require modern calibration functions estimated from modern assemblages from the same sedimentary environment as the fossil assemblages. However, Goring *et al.*'s (2010) work in British Columbia suggests that model error declines when samples from different depositional environments are combined for some pollen-based reconstruction models and climate variables.

The other important set of limitations of the multivariate calibration-function approach and the assemblage approach relate to two of the basic assumptions of climate reconstruction (see above), namely that the climate variable to be reconstructed is, or is linearly, or at least monotonically, related to an ecologically important determinant in the system of interest (assumption 2) and other variables than the climate variable of interest have negligible influence or their joint distribution with the climate variable of interest in the past is the same as today (assumption 5). These two assumptions are critical but are not often challenged (Juggins & Birks 2011). Assumption 2 raises the fundamental question of what climate variable can (or cannot) be reconstructed using a particular training set. For example, Fr chette *et al.* (2008) use the same pollen training set to reconstruct July temperature, January temperature, July sunshine, and annual precipitation. Muller *et al.* (2003) use MAT on the same modern pollen data-set in eastern Canada to reconstruct annual precipitation, mean annual temperature, evaporation:potential evaporation, mean January and July temperatures, run-off, and growing degree-days above 0°C and above 5°C. See also Magny *et al.* (2001, 2003) for reconstructions of seven climate variables from one data-set. de Vernal *et al.* (2001; 2005) use the same dinoflagellate cyst assemblages to reconstruct summer and winter sea-surface temperatures, the number of months of ice cover, and summer sea-surface salinity (see Telford 2006 for a critique). The choice of climate variable to be reconstructed is usually done in one of two ways. First, it may be defined *a priori* by the initial project aims and the training-set designed specifically to sample this climate gradient (e.g. Brooks & Birks 2001; Bjune *et al.* 2010). Second, the climate variable may be selected *post hoc* after it appears to be an important biological determinant during a numerical analysis of the training set (e.g. Luoto 2010). In both cases, assumption 2, namely that the climate variable explains, in a statistical sense, a significant portion of the variation in the modern biological data can be tested using constrained ordination methods and associated Monte Carlo permutation tests (Birks 1995; 1998) with the climate variable of interest as the single constraining predictor variable (ter Braak 1987; ter Braak & Juggins 1993). The relative strength or importance of a climate variable of interest can be estimated by comparing the eigenvalue of the first axis of a detrended canonical correspondence analysis using the single variable of interest as the sole constraining variable with that of the first unconstrained axis (e.g. Lotter *et al.* 1997). As ter Braak (1987) notes, for a reconstruction to be useful, the first constrained eigenvalue should be *large* compared to the

other eigenvalues because only then is the variable to be reconstructed determining the modern biological assemblages. Climate variables are inherently inter-related (Jackson *et al.* 2009) and assumption 2 is often qualified by the additional requirement that the variable of interest explains a significant *and independent* portion of the variation in the biological data. This is tested implicitly if variables are stepwise forward-selected in a constrained ordination (ter Braak & Verdonschot 1995) or explicitly using partial constrained ordination (ter Braak & Prentice 1988). There are two important consequences from this. The first is that assumption 2 only requires a correlation between climate and biology, not a causal relationship. For some species there may be experimental evidence that suggests causal effects or demonstrates the physiological basis for the underlying response (e.g. Pigott 1982; Jackson *et al.* 2009). For other taxa there may be observational data that strongly suggest a direct effect (e.g. Brodersen *et al.* 2004, 2008). Where this is not the case, it should be recognised that the model may still have good predictive ability – but it does require an additional assumption that the relationship between the climate variable and the underlying causal ecological gradient has not changed through time. The second consequence is that reconstruction models developed for climate variables selected *post hoc* have an inherently weaker basis than those selected *a priori* in a hypothesis-testing approach. This is because of the circularity in cross-validating a model that has already been found to be significant. Models developed for variables selected using an automated step-wise procedure are especially problematic (Juggins & Birks 2011).

The problem of confounding effects (Table 5) of correlated climate and other environmental variables pervades almost all quantitative reconstructions, starting with Imbrie and Kipp (1971) and their reconstructions of summer and winter sea-surface temperatures and annual salinity. In almost all ecological systems, the composition of sedimentary assemblages is a complex function of multiple climate, edaphic, land-used, and historical factors. The first part of assumption 5, that environmental variables other than the climate variable of interest have negligible influence is therefore almost never met. The second part of this assumption, that the joint distribution of additional variables with the climate variable of interest does not change with time, is also violated in many cases (Jackson *et al.* 2009; Juggins & Birks 2011).

Telford and Birks (2005; 2009) discuss in detail assumption 6, namely the statistical independence of test data-sets, and, as reviewed above in connection with The Assemblage Approach, provide ways of testing and allowing for lack of independence as a result of spatial autocorrelation. Two-way WA is relatively robust to spatial structure in the training set, probably because it estimates global parameters. WA-PLS appears to be less robust to spatial structure (Telford & Birks 2005; 2009) but it is not as sensitive as assemblage-approach methods such as MAT and ANN.

This discussion of the general limitations highlights the need to understand better the limitations of quantitative climate reconstructions, particularly when the critical assumptions are violated and to be aware of the ecological, statistical, and palaeoecological assumptions, strengths, and

**Table 5. Main Types of Biological Proxies Considered in this Essay, Climate Variables to be Reconstructed, Confounding Variables, and Example Studies**

Proxy	Climate Variable	Confounding Variables	Examples
Pollen	Summer temperature	Winter temperature; precipitation; human activity; soil development	Seppä & Birks (2001; 2002) Bjune <i>et al.</i> (2010)
Pollen	Annual precipitation	Summer temperature; winter temperature; human activity; soil development	Birks <i>et al.</i> (2000) Seppä & Birks (2001; 2002)
Pollen	Annual mean temperature	Summer temperature; winter temperature; precipitation; human activity; soil development	Heikkilä & Seppä (2003)
Plant macrofossils	Maximum summer temperature	Winter temperature; precipitation	Conolly & Dahl (1970)
Plant macrofossils	Growing season warmth	Moisture	Thompson <i>et al.</i> (2008)
Chironomids	July air or water temperature	Changes in water chemistry (e.g. oxygen, nutrients) caused by human impact or non-linear ecosystem response to climate forcing; local effects on relationship between surface water and air temperature (e.g. addition of glacial meltwater or diversion of streams to lower altitude lakes)	Brooks & Birks (2001) Heiri & Lotter (2001) Velle <i>et al.</i> (2005)
Diatoms	July air or water temperature	Changes in water chemistry or transparency unrelated, non-linearly or indirectly related to climatic change	Bigler <i>et al.</i> (2002) Anderson (1998)
Cladocerans	July air or water temperature	Limnological changes unrelated to climate, biotic interactions (e.g. presence or absence of fish)	Lotter <i>et al.</i> (2000) Birks & Ammann (2000)

weaknesses of our data (Tables 2-5) (Velle *et al.* 2010; Juggins & Birks 2011).

Assessing and interpreting uncertainties in climate reconstructions is a largely unexplored but critical topic in the reconstruction of past climate from biostratigraphical data. Whilst it is relatively easy, given modern computing power and appropriate software, to derive sample-specific errors of reconstruction for the climate variable of interest by bootstrapping (Birks *et al.* 1990), these errors are solely errors of the reconstruction model itself. It is much more difficult to estimate the inherent errors associated with the modern biological ( $Y_m$ ) or environmental ( $X_m$ ) data or in the fossil biological data ( $Y_f$ ). The errors in the biological data are due to counting procedures, number of fossils counted, and laboratory techniques (Maher *et al.* 2011) and they can generally be estimated. What is much more difficult (and for some sources of error, possibly impossible) is to estimate the errors in the environmental data ( $X_m$ ) due to, for example, interpolation procedures, selection of climate period, local topography, and between-year variation. Moreover, none of the reconstruction methods discussed above can analyse variation in the modern environmental data. Goring *et al.* (2010) explicitly examine model error in pollen data from different sedimentary environments.

One possibility for assessing the influence of a particular error component on the overall error of a transfer function is to quantify the variability of inferred temperatures associated with a particular error source, assume independence between the overall model error and the error source of interest, and estimate the magnitude of the model error if this error source could be eliminated. For example, Heiri *et al.* (2003) obtained replicate samples from five lakes in Norway, assessed the variability of chironomid-inferred temperatures

for the individual lakes and estimated that this kind of variability is responsible for approximately 6-15% of the overall error of the applied calibration function. Heiri *et al.* (2003) concluded that the quantified error component, related to uncertainties associated with estimating the chironomid assemblage of a lake based on a single sediment sample and a limited number of analysed fossils, was too small to warrant more labour-intensive sediment sampling and enumeration schemes.

Environmental reconstructions developed in a Bayesian framework provide the opportunity of explicitly estimating and modelling different error components (see the discussions on New Approaches to Climate Reconstruction above and on Challenges and Future Developments below). A Bayesian framework offers a means of handling in a coherent, logical, and explicit manner the many sources of uncertainty in data and models and this must be considered to be one of the most relevant advantages of this approach compared with more conventional methods for producing climate reconstructions from palaeoecological data. The challenge is, of course, to implement this Bayesian approach in theory and in practice and to model realistically the different error components, given the limited data available for estimating different sources of variability associated with using biological proxies for inferring past climatic conditions. It is an enormous challenge but its successful implementation will be a major contribution to quantitative palaeoecology in the future (e.g. Brewer *et al.* 2008).

## CHALLENGES AND FUTURE DEVELOPMENTS

The major challenges and likely future developments in the reconstruction of climate from fossil assemblages fall into two major groups – numerical and ecological.

The most obvious numerical challenge and future development is to incorporate reconstruction methods into a Bayesian framework. All the reconstruction methods discussed above under the three main approaches have different statistical and ecological bases and assumptions but they all have one thing in common (with the exception of the Kühl *et al.* (2002) probability density function method): they are all so-called frequentist methods and they make the assumption that the model parameters (WA optima or WA-PLS coefficients, etc) are fixed and can be estimated from observations (the observed data in the training set) distributed randomly about the fitted values (Holden *et al.* 2008). Conversely, a Bayesian approach does not rely on an explicit model of the relationships between species and climate but assumes that the model is unknown and is to be estimated from the measured data which are fixed. Specifically, the Bayesian approach uses measured information to modify some prior belief about the environmental values (Robertson *et al.* 1999). This additional information is derived from a training set and expressed as a conditional probability density function, which is combined with the prior probability density function to give a posterior density function using Bayes theorem (Juggins & Birks 2011).

Bayesian approaches have been applied to reconstruct past climate in several palaeoecological studies. Toivonen *et al.* (2001) present a Bayesian model with a conditional probability density function based on a unimodal species-environment response model. Vasko *et al.* (2000) and Korhola *et al.* (2002) extend this to include the more realistic multinomial Gaussian response model and apply it to chironomid-based temperature reconstructions. Haslett *et al.* (2006) develop further these ideas to more generalised modelling of pollen-climate response surfaces, and Holden *et al.* (2008) present a computationally efficient approach based on probability weighting of species response curves. Haslett & Challenor (2010) present a very readable review of the current 'state-of-the-art' of palaeoclimate reconstructions within a Bayesian framework.

Although the prediction errors for Bayesian models are generally of a similar magnitude to those for non-Bayesian approaches, they have the major advantage that they provide a coherent and explicit handling of uncertainty (Juggins & Birks 2011). Bayesian models offer an elegant solution for modelling multiple sources of evidence and their associated uncertainties within a single coherent framework and are thus a priority for future work (Haslett & Challenor 2010). However, the lack of available computer software and the huge computational burden of some Bayesian reconstruction models (days to weeks for a single reconstruction) currently prevents their more widespread use (Juggins & Birks 2011). Hopefully this situation will change in the future.

Despite continuing development and refinement of reconstruction methods (e.g. Vasko *et al.* 2000; Toivonen *et al.* 2001; Korhola *et al.* 2002; Gersonde *et al.* 2005; Erästä & Holmström 2006; Haslett *et al.* 2006; Holden *et al.* 2008; Hübener *et al.* 2008; Goring *et al.* 2009; Velle *et al.* 2011), we are probably close to the resolution of current data and methods. Decomposition of the variance in model RMSEP shows that the largest part (75% or more) is due to the fact that at a given environmental value today, there is still site-to-site variation in modern assemblages (Birks *et al.* 1990).

No amount of mathematical analysis or modelling is going to minimise nature's inherent variation. We need to pay more attention to the assumptions of the reconstruction approaches and to improve our abilities to interpret and compare temporal series of climate reconstructions. There is an urgent need to develop robust approaches for comparing temporal series and to be able to interpret sample-specific errors of reconstruction in a rigorous manner.

A second future numerical development may be the stronger integration of quantitative climate reconstructions with palaeoclimate models and dynamic vegetation models. In addition to using reconstructions for the validation of palaeoclimate model outputs (e.g. Braconnot *et al.* 2007a, 2007b; Renssen *et al.* 2009), the integration of modelling can provide improvements for quantitative climate reconstructions under environmental conditions that are strongly non-analogous with the present. An example of such a situation is the climate reconstructions from the last glacial period when the atmospheric CO<sub>2</sub> content was 180 ppm, thus nearly 50% lower than at present. Under such conditions, low CO<sub>2</sub> can influence processes of carbon and water uptake by plants and influence their distribution patterns, thus complicating the use of modern climate-vegetation patterns for climate reconstruction. Inverse vegetation modelling has been suggested as a method that can account for the inconstancy of the vegetation-climate relationship under non-analogous environmental conditions, such as during the last glacial period (Guiot *et al.* 2000). In this approach, several most likely climate scenarios are estimated for the time period in question, the vegetation model is run with each scenario and the simulated vegetation types are compared with the fossil pollen data to estimate the most coherent simulations. The past climate is obtained by inverting the climate input from the vegetation model (Guiot *et al.* 1999, 2000, 2009; Garreta *et al.* 2010). Wu *et al.* (2007a) used an inverse modelling approach with modern and glacial CO<sub>2</sub> concentrations to investigate the impact of changing CO<sub>2</sub> concentration on the reconstruction results from Africa and Eurasia. Wu *et al.* (2007b) used inverse vegetation modelling to suggest that the lowering of the altitudinal treelines in tropical Africa during the last glacial maximum was mostly caused by drying but amplified by decreased CO<sub>2</sub> concentration. Garreta *et al.* (2010) improved the method by replacing the static equilibrium model with a dynamic vegetation model LPJ-GUESS. They argue that the use of a dynamic vegetation model helps to bypass the assumption of equilibrium between vegetation and climate change and that the temporal link of the dynamic vegetation model allows a better exploitation of information from fossil pollen records (Garreta *et al.* 2010). On the downside, the approach of inverse modelling incorporates additional sources of error and uncertainty associated with the dynamic vegetation models, including the species parameterisation used in the model, and are likely to lead to large error estimates in the reconstructions of last glacial climate.

The other major set of challenges is ecological. Since the pioneering work on quantitative environmental reconstructions from biological assemblages (e.g. Imbrie & Kipp 1971; Webb & Bryson 1972; ter Braak & van Dam 1989; Birks *et al.* 1990), there have been many methodological developments in the numerical procedures for reconstructing past climates (e.g. MAT, response surfaces, WA-PLS) and in the

creation and application of training sets world-wide. Despite all these developments, surprisingly little work has been done relating to the assumptions and limitations of the reconstruction methods (but see Telford & Birks 2005; 2009).

The development of user-friendly software like C<sup>2</sup> (Juggins 2005) has given all palaeoecologists access to a wide range of reconstruction procedures. As a result, there is a shift from a single method and reconstruction model to a multi-model approach in response to the recognition that in many cases there is no single 'best' method or model (Juggins & Birks 2011). This shift has included an important move away from an obsession with models with the lowest RMSEP or, even worse, the highest  $r^2$  (Birks 1995; 1998) to a more critical consideration of model uncertainty including bias. The use of multi-model reconstructions necessitates the recognition that model selection itself contributes to the overall uncertainty (Johnson & Omland 2004). Rigorous model selection has much to offer palaeoecologists. It offers a way of drawing inferences from a set of multiple competing hypotheses. It is based on likelihood theory, a robust framework that supports most modern statistical approaches (Johnson & Omland 2004). Its major advantages are as follows.

- (1) The investigator is not restricted to evaluating a single model where performance or significance is measured against some arbitrary threshold. Instead, competing models are compared to one another by evaluating the relative support in the observed data and by considering the model complexity (number of parameters, etc) for each model.
- (2) Models can be ranked and weighted, thereby providing a quantitative measure of relative support for each competing model.
- (3) When models have similar levels of support from the data, model averaging can be used to make robust parameter estimates and inferences (Johnson & Omland 2004).

There has also been a trend (Table 3) in recent years to reconstruct several environmental variables from a single training set (e.g. Luoto 2010). Given the very complex set of interacting environmental and biological factors on species distribution and abundance, the robustness of reconstruction methods to the effects of confounding variables (Table 5) becomes very critical, and is often ignored in studies where two or three environmental variables are reconstructed simultaneously from the same data-set. There are statistical tools to help misusing quantitative reconstruction methods to reconstruct variables of little or no quantifiable importance (ter Braak & Verdonschot 1995; Telford 2006; Juggins & Birks 2011), and greater use should be made of these and related statistical modelling techniques. The problem of the confounding effects of correlated environmental variables (Table 5) pervades all quantitative reconstructions (Juggins & Birks 2011), a point emphasised by Anderson (2000) in his discussion of the use of diatoms to reconstruct past climate and by Velle *et al.* (2010) in their discussion of the use of chironomids to reconstruct past climate.

Many of the early studies on quantitative environmental reconstructions involved strong environmental gradients

such as sea-surface temperature, lake-water pH, salinity, or summer temperature. These studies stimulated development of new reconstruction techniques and of high-quality modern training-sets. Attempts at reconstructing weaker, secondary gradients are much more problematical and different methods may give very conflicting reconstructions (Juggins & Birks 2011). Although the problems of confounding variables have barely been addressed, these problems almost certainly mean that reconstructions for some variables based on some proxies are problematic at best (Juggins & Birks 2011). Future developments will hopefully see palaeoecologists being more critical of their inference models and the resulting reconstructions, and be alert to the invidious effects of confounding variables. Such developments will require not only a good understanding of the assumptions and limitations of the numerical methods being used but also a sound knowledge of the organisms and their ecology when used as palaeoenvironmental proxies and of the relevant environmental variables and their interactions.

## CONCLUSIONS

Palaeoecology is the ecology of the past as we discussed in the Introduction. In practice, it has mainly been concerned with reconstruction of past biota, populations, communities, landscapes, environments, and ecosystems (Birks & Birks 1980). In such reconstructions, often all the available palaeoecological data, biotic and abiotic, have been used. The recent upsurge of interest and activity in multi-proxy studies (Birks & Birks 2006) where a range of biotic and abiotic proxies is studied on the same core or set of correlated cores allows palaeoecology to better understand the ecology of the past.

There are at least two major approaches to study the ecology of the past.

1. We study the responses of organisms in the past preserved in sediments to climate change but the palaeoclimate record is *not* based on the fossil group of interest but instead is based on independent palaeoclimate records such as stable isotopes and lake-level changes (Jackson & Booth 2002; Shuman *et al.* 2004), sea-level changes, diatoms, and geochemistry (e.g. Virah-Sawmy *et al.* 2009a; 2009b), calculated orbital solar irradiance and moisture values derived from stable isotopes (e.g. Willis *et al.* 2007), lake-level changes and fire severity (Shuman *et al.* 2009), tree-rings (Helama *et al.* 2009a, 2009b), or palaeolimnological proxies (e.g. Blass *et al.* 2007; Trachsel *et al.* 2010). This approach is very much in the scientific philosophy and methodology of using 'the geological record of ecological dynamics' and 'the geological record as an ecological laboratory' presented by Flessa and Jackson (2005a, 2005b).
2. In a multi-proxy study (e.g. Birks *et al.* 2000), one or more biological proxy can be used to provide an independent palaeoclimatic reconstruction using calibration functions (e.g. chironomids as a temperature proxy (Brooks & Birks 2000) or testate amoebae as a moisture proxy (Booth & Jackson 2003; Booth *et al.* 2004; Booth 2008, 2010). The responses in other proxies, e.g. pollen and, by inference, terrestrial

vegetation, to the inferred climate changes are studied and questions of response times, lags, and rates of turnover can be studied and quantified (e.g. Birks & Birks 2008).

Both approaches give palaeoclimatic records that are independent of the group(s) of fossils of primary ecological interest and permit insights into the ecology of the past. They allow the late-Quaternary palaeoecological record to be used as a long-term ecological observatory or laboratory (Flessa & Jackson 2005a, 2005b) in which long-term ecological dynamics can be studied under a range of environmental conditions, not all of which exist on Earth today (e.g. lower CO<sub>2</sub> concentrations, low human impact) (e.g. Liu *et al.* 2010).

Related to palaeoecological reconstruction but with a greater ecological emphasis than palaeoclimate inferences, are the recent studies that quantify the relationships between modern pollen assemblages and satellite remote-sensing data reflecting tree-cover or leaf-area index (LAI) (e.g. Williams 2002; Williams & Jackson 2003; Tarasov *et al.* 2007; Gonzales *et al.* 2008; Williams *et al.* 2008; Bezrukova *et al.* 2010; Herzschuh *et al.* 2010b). By using modern pollen data and Advanced Very High Resolution Radiometer (AVHRR) or Moderate Resolution Imaging spectroradiometer (MODIS) remote-sensing data, it is possible to establish modern pollen–tree-cover or modern pollen–LAI relationships. Given fossil pollen data and with MAT- (Williams 2002; Tarasov *et al.* 2007; Gonzales *et al.* 2008; Williams *et al.* 2008; Bezrukova *et al.* 2010) or WA-PLS-based (Herzschuh *et al.* 2010b) models, reconstructions of past conifer-forest cover, deciduous-forest cover, vegetation cover, or LAI can be made. Using the same approach, net primary productivity can also be reconstructed which along with LAI is an important index of vegetation greenness (Williams 2002). Such reconstructions provide a valuable link between Earth system models and palaeoecological data.

By more closely integrating palaeoecology with ecology and climate and Earth system modelling (e.g. Brewer *et al.* 2007, 2009), and by changing the emphasis from reconstruction of past environments to the ecology of the past, palaeoecology can provide unique information about past systems such as variations in rates, states, composition, and resilience, responses to external boundary conditions and drivers including those with no modern counterparts, spatial and temporal scaling, and long-term perspectives that are much longer than those provided by ecological observations and long-term monitoring programmes. Palaeoecology can thus provide the ‘missing dimension’ in many aspects of ecology, conservation biology, and global science (Jackson 2006, 2007; Willis *et al.* 2010a).

Quantitative palaeoclimatic reconstructions based on the indicator-species approach, the assemblage approach, and/or the multivariate calibration-function approach will, we predict, continue to play an important role in future Quaternary palaeoecology. However, if the emphasis of Quaternary palaeoecology shifts from reconstructions of the past to the ecology of the past, climate reconstructions may cease to be ends in themselves but hopefully such reconstructions will become important means to new, challenging, and exciting ends in late-Quaternary palaeoecology.

## ACKNOWLEDGEMENTS

We are very grateful to many friends and colleagues for valuable discussions and ideas about palaeoclimate reconstructions, in particular Hilary Birks, Cajo ter Braak, Steve Brooks, Joel Guiot, Ulrike Herzschuh, Steve Juggins, Andy Lotter, Norbert K uhl, Gavin Simpson, Richard Telford, Gaute Velle, Jack Williams, and Kathy Willis. We are greatly indebted to Cathy Jenks for her invaluable help, to two anonymous reviewers for useful comments, and to Valenti Rull for his patience and encouragement. This is publication no. A312 from the Bjerknes Centre for Climate Research.

## REFERENCES

- Ackerly, DD, Loarie, SR, Cornwell, WK, Weiss, SB, Hamilton, H, Branciforte, R & Kraft, NJN (2010) The geography of climate change: implications for conservation biogeography. *Diversity and Distributions*, 16, 476-87.
- Allen, JRM, Huntley, B & Watts, WA (1995) The vegetation and climate of north-west Iberia over the last 14,000 years. *Journal of Quaternary Science*, 11, 125-47.
- Allen, JRM, Watts, WA & Huntley, B (2000) Weichselian palynostratigraphy, palaeovegetation and palaeoenvironment: the record from Lago Grande di Monticchio, southern Italy. *Quaternary International*, 73/74, 91-110.
- Allen, JRM, Watts, WA, McGee, E & Huntley, B (2002) Holocene environmental variability - the record from Lago Grande di Monticchio, Italy. *Quaternary International*, 88, 69-80.
- Anderson, BJ, Akcakaya, HR, Ara ujo, MB, Fordham, DA, Martinez-Meyer, E, Thuiller, W & Brook, BW (2009) Dynamics of range margins to metapopulations under climate change. *Proceedings of the Royal Society B*, 276, 1415-420.
- Anderson, NJ (2000). Diatoms, temperature and climate change. *European Journal of Phycology*, 35, 307-14.
- Anderson, RP, Lew, D & Peterson, AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162, 211-32.
- Andersson, G (1902) Hasseln i Sverige fordom och nu. *Sveriges Geologiska Unders kning Series C*, 3, 1-168.
- Andersson, G (1903) Klimatet i Sverige etter istiden. *Nordisk Tidsskrift*, 1, 1-26.
- Andersson, G (1909) The climate of Sweden in the Late-Quaternary period. Facts and theories. *Sveriges Geologiska Unders kning Series C*,  rbok 3, 1-88.
- Ara ujo, MB & Guisan, A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677-688.
- Ara ujo, MB & Luoto, M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16, 743-53.
- Ara ujo, MB & New, M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22, 42-47.
- Ara ujo, MB & Pearson, RG (2005) Equilibrium of species' distributions with climate. *Ecography*, 28, 693-95.
- Ara ujo, MB, Pearson, RG, Thuiller, W & Erhard, M (2005a) Validation of species-climate impact models under climate change. *Global Change Biology*, 11, 1504-513.
- Ara ujo, MB & Rahbek, C (2006) How does climate change affect biodiversity? *Science*, 313, 1396-397.
- Ara ujo, MB, Thuiller, W & Yoccoz, NG (2009) Reopening the climate envelope reveals macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences of the USA*, 106, E45-46.
- Ara ujo, MB, Whittaker, RJ, Ladle, RJ & Erhard, M (2005b) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, 14, 529-38.
- Arg ez, JA, Christen, JA, Nakamura, M & Sober n, J (2005) Prediction of potential areas of species distributions based on presence-only data. *Environmental and Ecological Statistics*, 12, 27-44.

- Ashcroft, MB, Chisholm, LA & French, KO (2008) The effect of exposure on landscape scale soil surface temperatures and species distribution models. *Landscape Ecology*, 23, 211-25.
- Ashcroft, MB, Chisholm, LA & French, KO (2009) Climate change at the landscape scale: predicting fine-grained spatial heterogeneity in warming and potential refugia for vegetation. *Global Change Biology*, 15, 656-67.
- Aspinall, RJ, Miller, JA & Franklin, J (2009) Calculations on the back of a climate envelope: Addressing the geography of species distributions. *Proceedings of the National Academy of Sciences of the USA*, 106, E44.
- Atkinson, TC, Briffa, KR & Coope, GR (1987) Seasonal temperatures in Britain during the past 22,000 years, reconstructed from beetle remains. *Nature*, 325, 587-92.
- Barnekow, L, Loader, NJ, Hicks, S, Froyd, CA & Goslar, T (2007) Strong correlation between summer temperature and pollen accumulation rates for *Pinus sylvestris*, *Picea abies* and *Betula* spp. in a high-resolution record from northern Sweden. *Journal of Quaternary Science*, 22, 653-58.
- Barrows, TT & Juggins, S (2005) Sea-surface temperatures around the Australian margin and Indian Ocean during the Last Glacial Maximum. *Quaternary Science Reviews*, 24, 1017-047.
- Barrows, TT, Juggins, S, De Dekker, P, Thiede, J & Martinez, JI (2000) Sea-surface temperatures of the southwest Pacific Ocean during the Last Glacial Maximum. *Paleoceanography*, 15, 95-109.
- Bartlein, PJ, Anderson, KH, Anderson, PM, Edwards, ME, Mock, CJ, Thompson, RS, Webb, RS & Whitlock, C (1998) Paleoclimatic simulations for North America over the past 21,000 years: features of the simulated climates and comparisons with paleoenvironmental data. *Quaternary Science Reviews*, 17, 549-85.
- Bartlein, PJ, Prentice, IC & Webb, III T (1986) Climatic response surfaces from pollen data for some eastern North American taxa. *Journal of Biogeography*, 13, 35-57.
- Bartlein, PJ & Whitlock, C (1993) Paleoclimatic interpretation of the Elk Lake pollen record. *Geological Society of America Special Paper*, 276, 275-93.
- Baselga, A & Araújo, MB (2009) Individualistic vs. community modelling of species distributions under climate change. *Ecography*, 32, 55-65.
- Baselga, A & Araújo, MB (2010) Do community-level models describe community variation effectively? *Journal of Biogeography*, 37, 1842-850.
- Beale, CM, Lennon, JJ, Elston, DA, Brewer, MJ & Yearsley, JM (2007) Red herrings remain in geographical ecology: a reply to Hawkins *et al.* (2007). *Ecography*, 30, 845-47.
- Beale, CM, Lennon, JJ & Gimona, A (2008) Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences of the USA*, 105.
- Beale, CM, Lennon, JJ & Gimona, A (2009) European bird distributions still show few climate associations. *Proceedings of the National Academy of Sciences of the USA*, 106, E41-43.
- Beale, CM, Lennon, JJ, Yearsley, JM, Brewer, MJ & Elston, DA (2010) Regression analysis of spatial data. *Ecology Letters*, 13, 246-64.
- Bennett, JR, Cumming, BF, Ginn, BK & Smol, JP (2010) Broad-scale environmental response and niche conservatism in lacustrine diatom communities. *Global Ecology and Biogeography*, 19, 724-32.
- Bennie, JJ, Wiltshire, AJ, Joyce, AN, Clark, D, Lloyd, AR, Adamson, J, Parr, T, Baxter, R & Huntley, B (2010) Characterising inter-annual variation in the spatial pattern of thermal microclimate in a UK upland using a combined empirical-physical model. *Agricultural and Forest Meteorology*, 150, 12-19.
- Berzukova, EV, Tarasov, PE, Solovieva, N, Krivonogov, SK & Riedel, F (2010) Last glacial-interglacial vegetation and environment dynamics in southern Siberia: Chronology, forcing and feedbacks. *Palaeogeography Palaeoclimatology Palaeoecology*, 296, 185-98.
- Betts, MG, Diamond, AW, Forbes, GJ, Villard, M-A & Gunn, JS (2006) The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. *Ecological Modelling*, 191, 197-204.
- Betts, MG, Ganio, LM, Huso, MMP, Som, NA, Huettmann, F, Bowman, J & Wintle, BA (2009) Comment on 'Methods to account for spatial autocorrelation in the analysis of species distributional data: a review'. *Ecography*, 32, 374-78.
- Bigler, C, Gavin, DG, Cuning, C & Veblen, TT (2007) Drought induces lagged tree mortality in a subalpine forest in the Rocky Mountains. *Oikos*, 116, 1983-994.
- Bigler, C & Hall, RI (2003) Diatoms as quantitative indicators of July temperature: a validation attempt at century-scale with meteorological data from northern Sweden. *Palaeogeography Palaeoclimatology Palaeoecology*, 189, 147-60.
- Bigler, C, Larocque, I, Peglar, SM, Birks, HJB & Hall, RI (2002) Quantitative multiproxy assessment of long-term patterns of Holocene environmental change from a small lake near Abisko, northern Sweden. *Holocene*, 12, 481-96.
- Bini, LM, Diniz-Filho, JA, Rangel, TF, Akre, TSB & *et al.* (2009) Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography*, 32, 193-204.
- Birks, HH & Ammann, B (2000) Two terrestrial records of rapid climatic change during the glacial-Holocene transition (14,000-9,000 calendar years B.P.) from Europe. *Proceedings of the National Academy of Sciences of the USA*, 97, 1390-394.
- Birks, HH, Battarbee, RW & Birks, HJB (2000) The development of the aquatic ecosystem at Kråkenes Lake, western Norway, during the late glacial and early Holocene - a synthesis. *Journal of Paleolimnology*, 23, 91-114.
- Birks, HH & Birks, HJB (2000) Future uses of pollen analysis must include plant macrofossils. *Journal of Biogeography*, 27, 31-35.
- Birks, HH & Birks, HJB (2006) Multi-proxy studies in palaeolimnology. *Vegetation History and Archaeobotany*, 15, 235-51.
- Birks, HJB (1973a) *Past and Present Vegetation of the Isle of Skye - A Palaeoecological Study*. Cambridge University Press, Cambridge.
- Birks, HJB (1973b) Modern pollen rain studies in some arctic and alpine environments. In: *Quaternary Plant Ecology* (Eds. Birks, HJB & West, RG). Blackwell Scientific Publications, Oxford, pp. 143-68.
- Birks, HJB (1981) The use of pollen analysis in the reconstruction of past climates: a review. In: *Climate and History* (Eds. Wigley TML, Ingram MJ & Farmer G). Cambridge University Press, Cambridge, pp. 111-38.
- Birks, HJB (1994) The importance of pollen and diatom taxonomic precision in quantitative paleoenvironmental reconstructions. *Review of Palaeobotany and Palynology*, 83, 107-17.
- Birks, HJB (1995) Quantitative palaeoenvironmental reconstructions. In: *Statistical Modelling of Quaternary Science Data. Technical Guide 5* (Eds. Maddy D & Brew JS). Quaternary Research Association Cambridge, pp. 161-254.
- Birks, HJB (1996) Contributions of Quaternary palaeoecology to nature conservation. *Journal of Vegetation Science*, 7, 89-98.
- Birks, HJB (1998) Numerical tools in palaeolimnology - Progress, potentialities, and problems. *Journal of Paleolimnology*, 20, 307-32.
- Birks, HJB (2001) Maximum likelihood environmental calibration and the computer program WACALIB - a correction. *Journal of Paleolimnology*, 25, 111-15.
- Birks, HJB (2003) Quantitative palaeoenvironmental reconstructions from Holocene biological data. In: *Global Change in the Holocene* (Eds. Mackay AW, Battarbee RW, Birks HJB & Oldfield F). Arnold London, pp. 342-57.
- Birks, HJB (2008) Holocene climate research - progress, paradigms, and problems. In: *Natural Climate Variability and Global Warming: A Holocene Perspective* (Eds. Battarbee RW & Binney H). Wiley-Blackwell Oxford, pp. 7-57.
- Birks, HJB (2010) Numerical methods for the analysis of diatom assemblage data. In: *The Diatoms - Applications for the Environmental and Earth Sciences* (Eds. Smol JP & Stoermer EF). Cambridge University Press, Cambridge, pp. 23-54.
- Birks, HJB & Birks, HH (1980) *Quaternary Palaeoecology*. (Reprinted 2004 by the Blackburn Press, New Jersey) Edward Arnold, London.
- Birks, HJB & Birks, HH (2008) Biological responses to rapid climate changes at the Younger Dryas-Holocene transition at Kråkenes, western Norway. *The Holocene*, 18, 19-30.
- Birks, HJB, Line, JM, Juggins, S, Stevenson, AC & ter Braak, CJF (1990) Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 327, 263-78.
- Birks, HJB & Seppä, H (2004) Pollen-based reconstructions of late-Quaternary climate in Europe - progress, problems, and pitfalls. *Acta Palaeobotanica*, 44, 317-34.

- Birks, HJB & Seppä, H (2010) Late-Quaternary palaeoclimatic research in Fennoscandia - a historical review. *Boreas*.
- Bjune, AE, Birks, HJB, Peglar, SM & Odland, A (2010) Developing a modern pollen-climate calibration data-set in Norway. *Boreas*, 39, 674-88.
- Blaauw, M, Bennett, KD & Christen, JA (2010) Random walk simulations of fossil proxy data. *The Holocene*, 20, 645-49.
- Blach-Overgaard, A, Svenning, J-C, Dransfield, J, Greve, M & Balslev, H (2010) Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. *Ecography*, 33, 380-91.
- Blass, A, Bigler, C, Grosjean, M & Sturm, M (2007) Decadal-scale autumn temperature reconstruction back to AD 1580 inferred from the varved sediments of Lake Silvaplana (southeastern Swiss Alps). *Quaternary Research*, 68, 184-95.
- Blytt, A (1881) Die Theorie der wechselnden kontinentalen und insularen Klimate. *Englers Botanisch Jahrbüch*, 2, 1-50.
- Bonnet, S, de Vernal, A, Hillaire-Marcel, C, Radi, T & Husum, K (2010) Variability of sea-surface temperature and sea-ice cover in the Fram Strait over the last two millennia. *Marine Micropaleontology*, 74, 59-74.
- Booth, RK (2008) Testate amoebae as proxies for mean annual water-table depth in *Sphagnum*-dominated peatlands of North America. *Journal of Quaternary Science*, 23, 43-57.
- Booth, RK (2010) Testing the climate sensitivity of peat-based paleoclimate reconstructions in mid-continental North America. *Quaternary Science Reviews*, 29, 720-31.
- Booth, RK & Jackson, ST (2003) A high-resolution record of Late Holocene moisture variability from a Michigan raised bog. *The Holocene*, 13, 865-78.
- Booth, RK, Jackson, ST & Gray, CED (2004) Paleocology and high-resolution paleohydrology of a kettle peatland in upper Michigan. *Quaternary Research*, 61, 1-13.
- Bordon, A, Peyron, O, Lézine, A-M, Brewer, S & Fouache, E (2009) Pollen-inferred late-glacial and Holocene climate in southern Balkans (Lake Maliq). *Quaternary International*, 200, 19-30.
- Bortolus, A (2008) Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *Ambio*, 37, 114-15.
- Botkin, DB, Saxe, H, Araujo, MB, Betts, R, Bradshaw, RHW, Cedhagen, T, Chesson, P, Dawson, T, Ettlerson, JR, Faith, DP, Ferrier, S, Guisan, A, Hansen, AS, Hilbert, DW, Loehle, C, Margules, C, New, M, Sobel, MJ & Stockwell, DRB (2007) Forecasting the effects of global warming on biodiversity. *Bioscience*, 57, 227-36.
- Braconnot, P, Otto-Bliesner, B, Harrison, SP, Joussaume, S, Peterschmitt, J-Y, Abe, Ouchi, A, Crucifix, M, Driesschaert, E, Fichet, T, Hewitt, CS, Kageyama, M, Kitoh, A, Loutre, M-F, Marti, O, Merkel, U, Ramstein, G, Valdes, P, Weber, L, Yu, Y & Zhao, Y (2007a) 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, 261-77.
- Braconnot, P, Otto-Bliesner, B, Harrison, SP, Joussaume, S, Peterschmitt, J-Y, Abe, Ouchi, A, Crucifix, M, Fichet, T, Hewitt, CS, Kageyama, M, Kitoh, A, Loutre, M-F, Marti, O, Merkel, U, Ramstein, G, Valdes, P, Weber, L, Yu, Y & Zhao, Y (2007b) Results of PMIP2 coupled simulations of the mid-Holocene and Last Glacial Maximum, Part 2: Feedbacks with emphasis on the location of the ITCZ and mid- and high-latitude heat budget. *Climate of the Past*, 3, 279-96.
- Brady, TJ, Monleon, VJ & Gray, AN (2010) Calibrating vascular plant abundance for detecting future climate changes in Oregon and Washington, USA. *Ecological Indicators*, 10, 657-67.
- Bray, PJ, Blockley, SPE, Coope, GR, Dodswell, LF, Elias, SA, Lowe, JJ & Pollard, AM (2006) Refining mutual climate range (MCR) quantitative estimates of palaeotemperature using ubiquity analysis. *Quaternary Science Reviews*, 25, 1865-876.
- Brewer, S, François, L, Cheddadi, R, Laurent, J-M & Favre, E (2009) Comparison of simulated and observed vegetation for the mid-Holocene in Europe. *Climate of the Past Discussions* 5, 965-1011.
- Brewer, S, Guiot, J & Barboni, D (2007) Use of pollen as climate proxies. In: *Encyclopedia of Quaternary Science* (Ed. Elias SA). Elsevier Oxford, pp. 2497-508.
- Brewer, S, Guiot, J, Sánchez-Goni, MF & Klotz, S (2008) The climate in Europe during the Eemian: a multi-method approach using pollen data. *Quaternary Science Reviews*, 27, 2303-315.
- Brewer, S, Guiot, J & Torre, F (2007) Mid-Holocene climate change in Europe: a data-model comparison. *Climate of the Past*, 3, 499-512.
- Brodersen, KP, Pedersen, O, Walker, IR & Jensen, MT (2008) Respiration of midges (Diptera; Chironomidae) in British Columbian lakes: oxy-regulation, temperature and their role as palaeo-indicators. *Freshwater Biology*, 53, 593-602.
- Brooks, SJ & Birks, HJB (2000) Chironomid-inferred late-glacial and early-Holocene mean July air temperatures for Kråkenes Lake, western Norway. *Journal of Paleolimnology*, 23, 77-89.
- Brooks, SJ & Birks, HJB (2001) Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. *Quaternary Science Reviews*, 20, 1723-741.
- Buckley, LB, Urban, MC, Angilletto, MJ, Crozier, LG, Rissler, LJ & Sears, MW (2010) Can mechanism inform species' distribution models? *Ecology Letters*, 13, 1041-054.
- Burman, P, Chow, E & Nolan, D (1994) A cross-validatory method for dependent data. *Biometrika*, 81, 351-58.
- Calcote, RR (2003) Mid-Holocene climate and the hemlock decline: the range limit of *Tsuga canadensis* in the western Great Lakes region, USA. *The Holocene*, 13, 215-24.
- Carter, RN & Prince, SD (1981) Epidemic models used to explain biogeographical distribution limits. *Nature*, 293, 644-45.
- Caseldine, CJ & Turney, C (2010) The bigger picture: towards integrating palaeoclimate and environmental data with a history of societal change. *Journal of Quaternary Science*, 25, 88-93.
- Caseldine, CJ, Turney, C & Long, AJ (2010) IPCC and palaeoclimate: an evolving story? *Journal of Quaternary Science*, 25, 1-4.
- Chapman, DS (2010) Weak climatic associations among British plant distributions. *Global Ecology and Biogeography*, 19, 831-41.
- Chaudhuri, P & Marron, JS (1999) SiZer for exploration of structures in curves. *Journal of the American Statistical Association*, 94, 807-23.
- Cheddadi, R, Lamb, HF, Guiot, J & van der Kaars, J (1998a) Holocene climatic change in Morocco: a quantitative reconstruction from pollen data. *Climate Dynamics*, 14, 883-90.
- Cheddadi, R, Mamakowa, K, Guiot, J, de Beaulieu, J-L, Reille, M, Andrieu, V, Granoszewski, W & Peyron, O (1998b) Was the climate of the Eemian stable? A quantitative climate reconstruction from seven European pollen records. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 143, 73-85.
- Cheddadi, R, Yu G, Guiot, J, Harrison, SP & Prentice, IC (1997) The climate of Europe 6000 years ago. *Climate Dynamics*, 13, 1-9.
- Christiansen, FB & Fenchel, TM (1977) *Theories of Population in Biological Communities*. Springer, Berlin.
- Clark, RM (1989) A randomization test for the comparison of ordered sequences. *Mathematical Geology*, 21, 429-42.
- Cleveland, WS (1993) *Visualizing Data*. AT&T Bell Laboratories, New Jersey.
- Cole, KL, Fisher, J, Arundel, ST, Cannella, J & Swift, S (2008) Geographical and climatic limits of needle types of one- and two-needled pinyon pines. *Journal of Biogeography*, 35, 257-69.
- Colwell, RK & Rangel, TF (2009) Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences of the USA*, 106, 19651-9658.
- Conolly, AP & Dahl, E (1970) Maximum summer temperatures in relation to modern and Quaternary distributions of certain arctic-montane species in the British Isles. In: *Studies in the Vegetational History of the British Isles* (Eds. Walker D. & West R.G.). Cambridge University Press, Cambridge, pp. 159-223.
- Coope, R, Lehmdahl, G, Lowe, JJ & Walkling, J (1998) Temperature gradients in northern Europe during the last glacial-Holocene transition (14-9 <sup>14</sup>C kyr BP) interpreted from coleopteran assemblages. *Quaternary Science Reviews*, 13, 419-33.
- Currie, DJ (2001) Projected effects of climate change on patterns of vertebrate and tree species richness in the conterminous United States. *Ecosystems*, 4, 216-25.
- Dahl, E (1951) On the relation between summer temperature and the distribution of alpine vascular plants in the lowlands of Fennoscandia. *Oikos*, 3, 22-52.
- Dahl, E (1998) *The Phytogeography of Northern Europe (British Isles, Fennoscandia and adjacent areas)*. Cambridge University Press, Cambridge.

- Daly, C (2006) Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, 26, 707-21.
- Davis, BAS & Brewer, S (2009) Orbital forcing and role of the latitudinal insolation/temperature gradient. *Climate Dynamics* 32, 143-65.
- Davis, BAS, Brewer, S, Stevenson, AC, Guiot, J & Data Contributors (2003) The temperature of Europe during the Holocene reconstructed from pollen data. *Quaternary Science Reviews*, 22, 1701-16.
- de Knegt, HJ, van Langevelde, F, Coughenour, MB, Skidmore, AK, de Boer, WF, Heitkönig, IMA, Knox, NM, Slotow, R, van der Waal, C & Prins, HHT (2010) Spatial autocorrelation and the scaling of species-environment relationships. *Ecology*, 91, 2455-65.
- de Marco, P, Diniz-Filho, JAF & Bini, LM (2008) Spatial analysis improves species distribution modelling during range expansion. *Biology Letters*, 4, 577-80.
- de Vernal, A, Eynaud, F, Henry, M, Hillaire-Marcel, C, Londeix, L, Mangin, S, Matthiessen, J, Marret, F, Radi, T, Rochon, A, Solignac, S & Turon, J-L (2005) Reconstruction of sea-surface conditions at middle to high latitudes of the Northern Hemisphere during the Last Glacial Maximum (LGM) based on dinoflagellate cyst assemblages. *Quaternary Science Reviews*, 24, 897-924.
- de Vernal, A, Henry, M, Matthiessen, J, Mudie, PJ, Rochon, A, Boessenkool, KP, Eynaud, F, Grosfjeld, K, Guiot, J, Hamel, D, Harland, R, Head, MJ, Kunz-Pirring, M, Levac, E, Loucheur, V, Peyron, O, Pospelova, V, Radi, T, Turon, J-L & Voronina, E (2001) Dinoflagellate cyst assemblages as tracers of sea-surface conditions in the northern North Atlantic, Arctic and sub-Arctic seas: the new 'n = 677' data base and its application for quantitative palaeoceanographic reconstruction. *Journal of Quaternary Science*, 16, 681-98.
- Diekmann, M (2003) Species indicator values as an important tool in applied plant ecology - a review. *Basic and Applied Ecology*, 4, 493-506.
- Dimitriadis, S & Cranston, PS (2001) An Australian Holocene climate reconstruction using Chironomidae from a tropical volcanic maar lake. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 176, 109-31.
- Diniz-Filho, JA, Bini, LM, Rangel, TF, Loyola, RD, Hof, C, Nogués-Bravo, D & Araújo, M (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, 32, 897-906.
- Dormann, CF (2007a) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, 8, 387-97.
- Dormann, CF (2007b) Assessing the validity of autologistic regression. *Ecological Modelling*, 207, 234-42.
- Dormann, CF (2007c) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, 16, 129-38.
- Dormann, CF (2009) Response to Comment on 'Methods to account for spatial autocorrelation in the analysis of species distributional data: a review'. *Ecography*, 32, 379-81.
- Dormann, CF, McPherson, JM, Araújo, MB, Bivand, R, et al. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30, 609-28.
- Duckworth, JC, Kent, M & Ramsay, PM (2000) Plant functional types: an alternative to taxonomic plant community description in biogeography? *Progress in Physical Geography*, 24, 515-42.
- Duncan, RP, Cassey, P & Blackburn, TM (2009) Do climate envelope models transfer? A manipulative test using dung beetle introductions. *Proceedings of the National Academy of Sciences of the USA*, 267, 1449-457.
- Dzwonko, Z (2001) Assessment of light and soil conditions in ancient and recent woodlands by Ellenberg indicator values. *Journal of Applied Ecology*, 38, 942-51.
- Elias, SA (1994) *Quaternary Insects and their Environments*. Smithsonian Institution, Washington DC.
- Elias, SA (1997) The mutual climatic range method of palaeoclimatic reconstruction based on insect fossils: new applications and interhemispheric comparisons. *Quaternary Science Reviews*, 16, 1217-225.
- Elias, SA (2001) Mutual climatic range reconstructions of seasonal temperatures based on Late Pleistocene fossil beetle assemblages in Eastern Beringia. *Quaternary Science Reviews*, 20, 77-91.
- Elias, SA, Andrews, JT & Anderson, KH (1999) Insights on the climatic constraints on the beetle fauna of coastal Alaska, USA, derived from the mutual climatic range methods of paleoclimate reconstruction. *Arctic, Antarctic, and Alpine Research*, 31, 94-98.
- Elith, J, Graham, CH, Anderson, RP, Dudik, M & et al. (2006) Novel methods improve prediction of species distributions from occurrence data. *Ecography*, 29, 129-51.
- Elith, J & Leathwick, JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677-97.
- Ellenberg, H (1948) Unkrautgesellschaften als Mass für den Säuregrad, die Verdichtung und andere Eigenschaften des Ackerbodens. *Berichte über Landtechnik, Kuratorium für Technik und Bauwesen in der Landwirtschaft*, 4, 130-46.
- Ellenberg, H, Weber, HE, Düll, R, Wirth, V, Werner, W & Paulissen, D (1992) Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta Geobotanica*, 18, 1-248.
- Enfield, DB & Cid-Serrano, L (2006) Projecting the risk of future climate shifts. *International Journal of Climatology*, 26, 885-95.
- Engels, S, Bohncke, SJP, Bos, JAA, Brooks, SJ, Heiri, O & Helmens, KF (2008) Chironomid-based palaeotemperature estimates for northeast Finland during Oxygen Isotope Stage 3. *Journal of Paleolimnology*, 40, 49-61.
- Engels, S, Helmens, KF, Välranta, M, Brooks, SJ & Birks, HJB (2010) Early Weichselian (MIS 5d and 5c) temperatures and environmental changes in northern Fennoscandia as recorded by chironomids and macroremains at Sokli, northeast Finland. *Boreas*, 39, 689-704.
- Engler, R, Randin, CF, Vittoz, P, Czaka, T, Beniston, M, Zimmermann, NE & Guisan, A (2009) Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? *Ecography*, 32, 34-45.
- Erasto, P & Holmstrom, L (2005) Bayesian multiscale smoothing for making inferences about features in scatter plots. *Journal of Computational and Graphical Statistics*, 14, 569-89.
- Erasto, P & Holmstrom, L (2006) Selection of prior distributions and multiscale analysis in Bayesian temperature reconstructions based on fossil assemblages. *Journal of Paleolimnology*, 36, 69-80.
- Everitt, BS (Ed.) (2003) *The Cambridge Dictionary of Statistics*. 2nd ed. Cambridge University Press, Cambridge.
- Fægri, K (1950) On the value of palaeoclimatological evidence. *Centenary Proceedings of the Royal Meteorological Society*, 1950, 188-95.
- Fauquette, S, Clauzon, G, Suc, J-P & Zheng, Z (1999) A new approach for palaeoaltitude estimates based on pollen records: example of the Mercantour Massif (southeastern France) at the earliest Pliocene. *Earth and Planetary Science Letters*, 170, 35-47.
- Fauquette, S, Guiot, J & Suc, J-P (1998) A method for climatic reconstruction of the Mediterranean Pliocene using pollen data. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 144, 183-201.
- Feurdean, A, Klotz, S, Brewer, S, Mosbrugger, V, Tamas, T & Wohlfarth, B (2008b) Late glacial climate development in NW Romania - Comparative results from three quantitative pollen-based methods. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 265, 121-33.
- Feurdean, A, Klotz, S, Mosbrugger, V & Wohlfarth, B (2008a) Pollen-based quantitative reconstructions of Holocene climate in NW Romania. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 260, 494-504.
- Finkelstein, SA, Gajewski, K & Viau, AE (2006) Improved resolution of pollen taxonomy allows better biogeographical interpretation of post-glacial forest development: analyses from the North American Pollen Database. *Journal of Ecology*, 94, 415-30.
- Flessa, KW & Jackson, ST (2005a) *The Geological Record of Ecological Dynamics. Understanding the biotic effects of future environmental change*. National Research Council of the National Academies, Washington, D.C.
- Flessa, KW & Jackson, ST (2005b) Forging a common agenda for ecology and paleoecology. *Bioscience*, 55, 1030-031.
- Fletcher, WJ, Sanchez, Goni, MF, Peyron, O & Dormoy, I (2010) Abrupt climate changes of the last deglaciation detected in a Western Mediterranean forest record. *Climate of the Past*, 6, 245-64.

- Foster, DR, Swanson, F, Aber, J, Burke, I, Brokaw, N, Tilman, D & Knaap, A (2003) The importance of land-use and its legacies to ecology and environmental management. *BioScience*, 53, 77-88.
- Franklin, J (2010) *Mapping Species Distributions - Spatial Inference and Prediction*. Cambridge University Press, Cambridge.
- Fréchet, B, de Vernal, A, Guiot, J, Wolfe, AP, Miller, GH, Fredskild, B, Kerwin, MW & Richard, PJH (2008) Methodological basis for quantitative reconstruction of air temperature and sunshine from pollen assemblages in Arctic Canada and Greenland. *Quaternary Science Reviews*, 27, 1197-1216.
- Fridley, JD (2009) Downscaling climate over complex terrain: high finescale (<1000 m) spatial variation of near-ground temperatures in a montane forested landscape (Great Smoky Mountains). *Journal of Applied Meteorology and Climatology*, 48, 1033-1049.
- Froyd, CA & Willis, KJ (2008) Emerging issues in biodiversity and conservation management: the need for a palaeoecological perspective. *Quaternary Science Reviews*, 27, 1723-1732.
- Gajewski, K (2008) The Global Pollen Database in biogeographical and palaeoclimatic studies. *Progress in Physical Geography*, 32, 379-402.
- Garreta, V, Miller, PA, Guiot, J, Hely, C, Brewer, S, Sykes, MT & Litt, T (2010) A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model. *Climate Dynamics*, 35, 371-89.
- Gaston, KJ (2009) Geographic range limits: achieving synthesis. *Proceedings of the Royal Society B*, 276, 1395-406.
- Gavin, DG & Hu, FS (2005) Bioclimatic modelling using Gaussian mixture distributions and multiscale segmentation. *Global Ecology and Biogeography*, 14, 491-501.
- Gavin, DG & Hu, FS (2006) Spatial variation of climatic and non-climatic controls on species distribution: the range limit of *Tsuga heterophylla*. *Journal of Biogeography*, 33, 1384-396.
- Gersonde, R, Crosta, X, Abelmann, A & Armand, L (2005) Sea-surface temperature and sea ice distribution of the Southern Ocean at the EPILOG Last Glacial Maximum - a circum-Antarctic view based on siliceous microfossil records. *Quaternary Science Reviews*, 24, 869-96.
- Giesecke, T, Hickler, T, Kunkel, T, Sykes, MT & Bradshaw, RHW (2007) Towards an understanding of the Holocene distribution of *Fagus sylvatica* L. *Journal of Biogeography*, 34, 118-31.
- Ginn, BK, Cumming, BF & Smol, JP (2007) Diatom-based environmental inferences and model comparisons from 494 northeastern North American lakes. *Journal of Phycology*, 43, 647-61.
- Gonzales, LM, Grimm, EC, Williams, JW & Nordheim, EV (2009a) A modern plant-climate research dataset for modelling eastern North American plant taxa. *Grana*, 48, 1-18.
- Gonzales, LM, Williams, JW & Grimm, EC (2009b) Expanded response-surfaces: a new method to reconstruct paleoclimates from fossil pollen assemblages that lack modern analogues. *Quaternary Science Reviews*, 28, 3315-332.
- Gonzales, LM, Williams, JW & Kaplan, JO (2008) Variations in leaf area index in northern and eastern North America over the past 21,000 years: a data-model comparison. *Quaternary Science Reviews*, 27, 1453-466.
- Goring, S, Lacourse, T, Pellatt, MG, Walker, IG & Mathewes, RW (2010) Are pollen-based climate models improved by combining surface samples from soil and lacustrine substrates? *Review of Palaeobotany and Palynology*, 162, 203-12.
- Goring, S, Pellatt, MG, Lacourse, T, Walker, IR & Mathewes, RW (2009) A new methodology for reconstructing climate and vegetation from modern pollen assemblages: an example from British Columbia. *Journal of Biogeography*, 36, 626-38.
- Gould, SJ (1965) Is uniformitarianism necessary? *American Journal of Science*, 163, 223-28.
- Graham, CH, Ferrier, S, Huettmann, F, Moritz, CK & Peterson, AT (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19, 497-503.
- Green, RE, Collingham, YC, Willis, SG, Gregory, RD, Smith, KW & Huntley, B (2008) Performance of climate envelope models in retrodicting recent changes in bird population size from observed climatic change. *Biology Letters*, 4, 599-602.
- Grichuk, VP (1969) Opyt rekonstruktsii nekotorykh elementov klimata severnogo plusharu v Atlanticheskii period golotsena. In: *Golotsen VIII Kongressu INGQUA Paris* (Ed. Meishtadt M.I.). Izd-vo Nauka Moscow, pp. 41-57.
- Grichuk, VP, Gurtovaya, YY, Zelikson, EM & Borisova, OK (1984) Methods and results of Late Pleistocene reconstructions. In: *Late Quaternary Environments of the Soviet Union* (Ed. Velichko A.A.). Longman London, pp. 251-60.
- Grubb, PJ (1977) The maintenance of species-richness in plant communities: The importance of the regeneration niche. *Biological Reviews*, 52, 107-45.
- Guiot, J (1990) Methodology of the last climatic cycle reconstruction in France from pollen data. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 80, 49-69.
- Guiot, J, Cheddadi, R, Prentice, IC & Jolly, D (1996) A method of biome and land surface mapping from pollen data: Application to Europe 6000 years ago. *Palaeoclimates: Data and Modelling*, 1, 311-24.
- Guiot, J & de Vernal, A (2007) Transfer functions: methods for quantitative paleoceanography based on microfossils. In: *Proxies in Late Cenozoic Paleoclimatology* (Eds. Hillaire-Marcel, C & de Vernal, A). Elsevier Amsterdam, pp. 523-63.
- Guiot, J, Pons, A, de Beaulieu, J-L & Reille, M (1989) A 140,000 year continental climate reconstruction from two European pollen records. *Nature*, 338, 309-13.
- Guiot, J, Reille, M, de Beaulieu, J-L & Pons, A (1992) Calibration of the climate signal in a new pollen sequence from La Grand Pile. *Climate Dynamics*, 6, 259-84.
- Guiot, J, Torre, F, Jolly, D, Peyron, O, Boreux, JJ & Cheddadi, R (2000) Inverse vegetation modeling by Monte Carlo sampling to reconstruct paleoclimates under changed precipitation seasonality and CO<sub>2</sub> conditions: application to glacial climate in Mediterranean region. *Ecological Modelling*, 127, 119-40.
- Guiot, J, Wu, HB, Garreta, V, Hatte, C & Magny, M (2009) A few prospective ideas on climate reconstruction: from a statistical single proxy approach towards a multi-proxy and dynamical approach. *Climate of the Past*, 5, 571-83.
- Guisan, A & Thuiller, W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993-1009.
- Guisan, A & Zimmermann, NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147-86.
- Guo, Q & Liu, Y (2010) ModEco: an integrated software package for ecological niche modeling. *Ecography*, 33, 637-42.
- Hamann, A & Wang, T (2006) Potential effects of climate change on ecosystem and tree species distribution in British Columbia. *Ecology*, 87, 2773-786.
- Hampe, A (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, 13, 469-71.
- Haslett, J, Whitley, M, Bhattacharya, S, Salter-Townsend, M, Wilson, SP, Allen, JRM, Huntley, B & Mitchell, FJG (2006) Bayesian palaeoclimate reconstruction. *Journal of the Royal Statistical Society A*, 169, 395-438.
- Hawkes, JC, Pyatt, DG & White, IMS (1997) Using Ellenberg indicator values to assess soil quality in British forests from ground vegetation: a pilot study. *Journal of Applied Ecology*, 34, 375-87.
- Hawkins, BA, Diniz-Filho, JA, Bini, LM, de Marco, P & Blackburn, TM (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, 30, 375-84.
- Heikkilä, M & Seppä, H (2003) An 11,000 yr palaeotemperature reconstruction from the southern boreal zone in Finland. *Quaternary Science Reviews*, 22, 541-54.
- Heikkilä, M & Seppä, H (2010) Holocene climate dynamics in Latvia, eastern Baltic region: a pollen-based summer temperature reconstruction and regional comparison. *Boreas*, 39, 705-19.
- Heikkinen, RK, Luoto, M, Araújo, MB, Virkkala, R, Thuiller, W & Sykes, MT (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography*, 30, 751-77.
- Heikkinen, RK, Luoto, M, Virkkala, R, Pearson, RG & Körber, J-H (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, 16, 754-63.
- Heiri, C, Bugmann, H, Tinner, W, Heiri, O & Lischke, H (2006) A model-based reconstruction of Holocene treeline dynamics in the Central Swiss Alps. *Journal of Ecology*, 94, 206-16.
- Heiri, O, Cremer, H, Engels, S, Hoek, WZ, Peeters, W & Lotter, AF (2007) Lateglacial summer temperatures in the Northwest European

- lowlands: a chironomid record from Hijkermeer, The Netherlands. *Quaternary Science Reviews*, 26, 2420-437.
- Heiri, O, Lotter, AF, Hausmann, S & Kienast, F (2003) A chironomid-based Holocene summer air temperature reconstruction from the Swiss Alps. *The Holocene*, 13, 477-84.
- Heiri, O, Tinner, W & Lotter, AF (2004) Evidence for cooler European summers during periods of changing meltwater flux to the North Atlantic. *Proceedings of the National Academy of Sciences of the USA*, 101, 15285-5288.
- Helama, S, Meriläinen, J & Tuomenvirta, H (2009) Multicentennial megadrought in northern Europe coincided with a global El Niño-Southern Oscillation drought pattern during the Medieval Climate Anomaly. *Geology*, 37, 175-78.
- Helama, S, Timonen, M, Holopainen, J, Ogurtsov, MG, Mielikäinen, K, Eronen, M, Lindholm, M & Meriläinen, J (2009) Summer temperature variations in Lapland during the Medieval Warm Period and the Little Ice Age relative to natural instability of thermohaline circulation on multi-decadal and multi-centennial scales. *Journal of Quaternary Science*, 24, 450-56.
- Helmens, KF, Bos, JAA, Engels, S, van Meerbeeck, CJ, Bohncke, SJP, Renssen, H, Heiri, O, Brooks, SJ, Seppä, H, Birks, HJB & Wohlfarth, B (2007) Present-day temperatures in northern Scandinavia during the last glaciation. *Geology*, 35, 987-90.
- Hengeveld, R (1990) *Dynamic Biogeography*. Cambridge University Press, Cambridge.
- Herzschuh, U & Birks, HJB (2010) Evaluating the indicator value of Tibetan pollen taxa for modern vegetation and climate. *Review of Palaeobotany and Palynology*, 160, 197-208.
- Herzschuh, U, Birks, HJB, Mischke, S, Zhang, C & Böhner, J (2010a) A modern pollen-climate calibration set based on lake sediments from the Tibetan Plateau and its application to a Late Quaternary pollen record from the Qilian Mountains. *Journal of Biogeography*, 37, 752-66.
- Herzschuh, U, Birks, HJB, Ni, J, Zhao, Y, Liu, H, Liu, X & Grosse, G (2010b) Holocene land-cover changes on the Tibetan Plateau. *The Holocene*, 20, 91-104.
- Hill, MO & Carey, PD (1997) Prediction of yield in the Rothamsted Park Grass Experiment by Ellenberg indicator values. *Journal of Vegetation Science*, 8, 579-86.
- Hill, MO, Preston, CD, Bosanquet, SDS & Roy, DB (2007) *BRYOATT - Attributes of British and Irish mosses, liverworts and hornworts*. Centre for Ecology & Hydrology, Huntingdon.
- Hill, MO, Preston, CD & Roy, DB (2004) *PLANTATT - Attributes of British and Irish Plants: status, size, life history, geography and habitats*. Centre for Ecology & Hydrology, Huntingdon.
- Hill, MO, Roy, DB, Mountford, JO & Bunce, RGH (2000) Extending Ellenberg's indicator values to a new area: an algorithmic approach. *Journal of Applied Ecology*, 37, 3-15.
- Hintikka, V (1963) Über das Grossklima einiger Pflanzenareale in zwei Klimakoordinatensystemen dargestellt. *Annales Botanici Societas Zoologicae Botanicae Fennicae 'Vanamo'*, 34(5), 1-64.
- Hoffman, JD, Aguilar-Amuchastegui, N & Tyre, AJ (2010) Use of simulated data from a process-based habitat model to evaluate methods for predicting species occurrence. *Ecography*, 33, 656-66.
- Holden, PB, Mackay, AM & Simpson, GL (2008) A Bayesian palaeoenvironmental transfer function model for acidified lakes. *Journal of Paleolimnology*, 39, 551-66.
- Holmström, L & Erästö, P (2002) Making inferences about past environmental change using smoothing in multiple time scales. *Computational Statistics and Data Analysis*, 41, 289-309.
- Holtland, WJ, ter Braak, CJF & Schouten, MGC (2010) Iteratio: calculating environmental indicator values for species and relevés. *Applied Vegetation Science*, 13, 369-77.
- Horne, DJ (2007) A mutual temperature range method for Quaternary palaeoclimatic analysis using European non-marine Ostracoda. *Quaternary Science Reviews*, 26, 1398-415.
- Horne, DJ & Mezquita, F (2008) Palaeoclimatic applications of large databases: developing and testing methods of palaeotemperature reconstructions using non-marine ostracods. *Senckenbergiana Lethaea*, 88, 93-112.
- Hotchkiss, SC, Calcote, RR & Lynch, EA (2007) Response of vegetation and fire to Little Ice Age climate change: regional continuity and landscape heterogeneity. *Landscape Ecology*, 22, 25-41.
- Hübener, T, Dressler, M, Schwarz, A, Langner, K & Adler, S (2008) Dynamic adjustment of training set ('moving-window' reconstruction) by using transfer functions in paleolimnology - a new approach. *Journal of Paleolimnology*, 40, 79-95.
- Huntley, B (1993) The use of climate response surfaces to reconstruct palaeoclimate from Quaternary pollen and plant macrofossil data. *Philosophical Transactions of the Royal Society of London B*, 341, 215-24.
- Huntley, B (1996) Quaternary palaeoecology and ecology. *Quaternary Science Reviews*, 15, 591-606.
- Huntley, B (2001) Reconstructing past environments from the Quaternary palaeovegetation record. *Biology and Environment*, 101B, 3-18.
- Huntley, B, Bartlein, PJ & Prentice, IC (1989) Climatic control of the distribution and abundance of beech (*Fagus L.*) in Europe and North America. *Journal of Biogeography*, 16, 551-60.
- Huntley, B, Bernard, P, Altwegg, R, Chambers, L, Coetzee, BWT, Gibson, L, Hockey, PAR, Hole, DG, Midgley, GF, Underhill, LG & Willis, SG (2010) Beyond bioclimatic envelopes: dynamic species' ranges and abundance modelling in the context of climatic change. *Ecography*, 33, 621-26.
- Huntley, B, Berry, PM, Cramer, W & McDonald, AP (1995) Modelling present and potential future ranges of some European higher plants using climate response surfaces. *Journal of Biogeography*, 22, 967-1001.
- Huntley, B, Green, RE, Collingham, YC & Willis, SG (2008) *A Climatic Atlas of European Breeding Birds*. Lynx Editions, Barcelona.
- Huppert, A & Solow, AR (2004) A method for reconstructing climate from fossil beetle assemblages. *Proceedings of the Royal Society B*, 271, 1125-128.
- Hutson, WH (1977) Transfer functions under no-analog conditions: experiments with Indian Ocean planktonic foraminifera. *Quaternary Research*, 8, 355-67.
- Ibáñez, I, Clark, JS & Dietz, MC (2008) Evaluating the sources of potential migrant species: implications under climate change. *Ecological Applications*, 18, 1664-678.
- Ibáñez, I, Clark, JS, LaDeau, S & Lambers, JHR (2007) Exploiting temporal variability to understand tree recruitment response to climate change. *Ecological Monographs*, 77, 163-77.
- Imbrie, J & Kipp, NG (1971) A new micropaleontological method for quantitative paleoclimatology: application to a Late Pleistocene Caribbean core. In: *The Late Cenozoic Glacial Ages* (Ed. Turekian K.K.). Yale University Press, New Haven and London, pp. 71-181.
- Iversen, J (1944) *Viscum, Hedera and Ilex* as climate indicators. A contribution to the study of the post-glacial temperature climate. *Geologiska Föreningens i Stockholm Förhandlingar*, 66, 463-83.
- Iverson, LR, Prasad, AM, Matthews, SN & Peters, M (2008) Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *Forest Ecology and Management*, 254, 390-406.
- Jackson, ST (2006) Vegetation, environment, and time: The origination and termination of ecosystems. *Journal of Vegetation Science*, 17, 549-57.
- Jackson, ST (2007) Looking forward from the past: history, ecology, and conservation. *Frontiers in Ecology and the Environment*, 5, 455.
- Jackson, ST, Betancourt, JL, Booth, RK & Gray, ST (2009) Ecology and the ratchet of events: Climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Sciences of the USA*, 106, 19685-9692.
- Jackson, ST & Booth, RK (2002) The role of late Holocene climate variability in the expansion of yellow birch in the western Great Lakes region. *Diversity and Distributions*, 8, 275-84.
- Jackson, ST & Lyford, ME (1999) Pollen dispersal models in Quaternary plant ecology: Assumptions, parameters, and prescriptions. *Botanical Review*, 65, 39-75.
- Jackson, ST & Overpeck, JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, 26 (Supplement), 194-220.
- Jackson, ST & Williams, JW (2004) Modern analogs in Quaternary paleoecology: Here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences*, 32, 495-537.
- James, FC, Johnston, RF, Wamer, NO, Niemi, GJ & Boecklen, WJ (1984) The Grinnellian niche of the wood thrush. *American Naturalist*, 124, 17-30.

- Jeschke, JM & Strayer, DL (2008) Usefulness of bioclimatic models for studying climatic change and invasive species. *Annals of New York Academy of Sciences*, 1134, 1-24.
- Jiménez-Valverde, A, Lobo, JM & Hortal, J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, 14, 885-90.
- Johnson, JB & Omland, KS (2004) Model selection in ecology and evolution. *Trends in Ecology and Evolution*, 19, 101-08.
- Juggins, S (2005) *C2 User Guide. Software for ecological and palaeoecological data analysis and visualisation*. University of Newcastle, Newcastle-upon-Tyne.
- Juggins, S & Birks, HJB (2011) Quantitative environmental reconstructions from biostratigraphical data. In: *Tracking Environmental Change Using Lake Sediments Volume 5: Data Handling and Numerical Techniques* (Eds. Birks, HJB, Lotter, AF, Juggins, S & Smol JP). Springer, Dordrecht (in press).
- Kadmon, R, Farber, O & Danin, A (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13, 853-67.
- Kamenik, C, van der Knaap, WO, van Leeuwen, JFN & Goslar, T (2009) Pollen/climate calibration based on a near-annual peat sequence from the Swiss Alps. *Journal of Quaternary Science*, 24, 529-46.
- Kaspar, F, Kühl, N, Cubasch, U & Litt, T (2005) A model-data comparison of European temperatures in the Eemian interglacial. *Geophysical Research Letters*, 32, L11703.
- Kearney, M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, 115, 186-91.
- Kearney, M & Porter, W (2004) Mapping the fundamental niche: physiology, climate, and the distribution of a nocturnal lizard. *Ecology*, 85, 3119-31.
- Kearney, M & Porter, W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12, 334-50.
- Keith, DA, Akcakaya, HR, Thuiller, W, Midgley, GF, Pearson, RG, Phillips, SJ, Regan, HM, Araújo, M & Rebelo, TG (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, 4, 560-63.
- Kershaw, AP & Nix, HA (1988) Quantitative palaeoclimatic estimates from pollen data using bioclimatic profiles of extant taxa. *Journal of Biogeography*, 15, 589-602.
- Kerwin, MW, Overpeck, JT, Webb, RS & Anderson, KH (2004) Pollen-based summer temperature reconstructions for the eastern Canadian boreal forest, subarctic, and Arctic. *Quaternary Science Reviews*, 23, 1901-924.
- Kitricos, DJ & Leriche, A (2010) The effects of climate data precision on fitting and projecting species niche models. *Ecography*, 33, 115-27.
- Klein, C, Wilson, KA, Watts, ME, Stein, J, Berry, S, Carwardine, J, Smith, MS, Mackey, B & Possingham, HP (2009) Incorporating ecological and evolutionary processes into continental-scale conservation planning. *Ecological Applications*, 19, 206-17.
- Klotz, S, Guiot, J & Mosbrugger, V (2003) Continental European Eemian and early Würmian climate evolution: comparing signals using different quantitative reconstruction approaches based on pollen. *Global and Planetary Change*, 36, 277-94.
- Klotz, S, Müller, U, Mosbrugger, V, de Beaulieu, J-L & Reille, M (2004) Eemian to early Würmian climate dynamics: history and patterns of changes in Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 211, 107-26.
- Klotz, S & Pross, J (1999) Pollen-based climate reconstructions in the European Pleistocene: the modified indicator species approach as a tool for quantitative analysis. *Acta Palaeobotanica*, Suppl 2, 481-86.
- Korhola, A, Vasko, K, Toivonen, HTT & Olander, H (2002) Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modelling. *Quaternary Science Reviews*, 21, 1841-860.
- Korhola, A, Weckström, J, Holmström, L & Erästö, P (2000) A quantitative Holocene climatic record from diatoms in northern Fennoscandia. *Quaternary Science Reviews*, 54, 284-94.
- Köster, D, Racca, JMJ & Pienitz, R (2004) Diatom-based inference models and reconstructions revisited: methods and transformations. *Journal of Paleolimnology*, 32, 233-45.
- Kucera, M, Weinelt, M, Kiefer, T, Pflaumann, U, Hayes, A, Chen, MT, Mix, AC, Barrows, TT, Cortijo, E, Duprat, J, Juggins, S & Waelbroeck, C (2005) Reconstruction of sea-surface temperatures from assemblages of planktonic foraminifera: multi-technique approach based on geographically constrained calibration data sets and its application to glacial Atlantic and Pacific Oceans. *Quaternary Science Reviews*, 24, 951-98.
- Kühl, N (2003) Der Bestimmung botanisch-klimatologischer Transferfunktion und die Rekonstruktion der hodennahen Klimazustandes in Europa während der Eem-Warmzeit. *Dissertationes Botanicae*, 375, 1-149.
- Kühl, N, Gebhardt, C, Litt, T & Hense, A (2002) Probability density functions as botanical-climatological transfer functions for climate reconstruction. *Quaternary Research*, 58, 381-92.
- Kühl, N, Gebhardt, C, Litt, T & Hense, A (2007) Eemian and early Weichselian temperature and precipitation variability in northern Germany. *Quaternary Science Reviews*, 26, 3311-7.
- Kühl, N & Gobet, E (2010) Climate evolution during the Middle Pleistocene warm period of Bilshausen, Germany, compared to the Holocene. *Quaternary Science Reviews*, (submitted).
- Kühl, N & Litt, T (2003) Quantitative time series reconstruction of Eemian temperature at three European sites using pollen data. *Vegetation History and Archaeobotany*, 12, 205-14.
- Kühl, N, Moschen, R, Wagner, S, Brewer, S & Peyron, O (2010) A multiproxy record of late Holocene natural and anthropogenic environmental change from the *Sphagnum* peat bog Dürres Maar, Germany: implications for quantitative climate reconstructions based on pollen. *Journal of Quaternary Science*, 25, 678-88.
- Kumke, T, Hense, A, Schölzel, C, Andreev, AA, Brüchmann, C, Gebhardt, C, Helle, G, Kienel, U, Kühl, N, Litt, T, Neumann, F & Schleser, G (2004a) Transfer functions for paleoclimate reconstructions - applications. In: *The Climate in Historical Times* (Eds. Fischer, H, Kumke, T, Lohmann, G, Flöser, G, Miller, H, von Storch, H & Negendank, JFW). Springer Berlin, pp. 245-62.
- Kumke, T, Schölzel, C & Hense, A (2004b) Transfer functions for paleoclimate reconstructions - theory and methods. In: *The Climate in Historical Times* (Eds. Fischer, H, Kumke, T, Lohmann, G, Flöser, G, Miller, H, von Storch, H & Negendank, JFW). Springer Berlin, pp. 229-43.
- Lamentowicz, M, van der Knaap, WO, Lamentowicz, P, van Leeuwen, JFN, Mitchell, EAD, Goslar, T & Kamenik, C (2010) A near-annual palaeohydrological study based on testate amoebae from a sub-alpine mire: surface wetness and the role of climate during the instrumental period. *Journal of Quaternary Science*, 25, 190-202.
- Larcher, W (2003) *Physiological Plant Ecology*. Fourth edn. Springer, Berlin.
- Larocque, I, Grosjean, M, Heiri, O, Bigler, C & Blass, A (2009) Comparison between chironomid-inferred July temperatures and meteorological data AD 1850-2001 from varved Lake Silvaplana, Switzerland. *Journal of Paleolimnology*, 41, 329-42.
- Lawler, JJ, Shafer, SL, White, D, Kareiva, P, Maurer, EP, Blaustein, AR & Bartlein, PJ (2009) Projected climate-induced faunal change in the Western Hemisphere. *Ecology*, 90, 588-97.
- Lawler, JJ, White, D, Neilson, RP & Blaustein, AR (2006) Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology*, 12, 1568-584.
- Leathwick, JR, Rowe, D, Richardson, J, Elith, J & Hastie, T (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology*, 50, 2034-052.
- Lee, MSY (2000) A worrying systematic decline. *Trends in Ecology and Evolution*, 15, 346.
- Legendre, P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, 74, 1659-673.
- Leith, H & Whittaker, RH (1975) *Primary Productivity of the Biosphere*. Springer, New York.
- Lennon, JJ (2000) Red-shifts and red herrings in geographical ecology. *Ecography*, 23, 101-13.
- Levinsky, I, Skov, F, Svenning, J-C & Rahbek, C (2007) Potential impacts of climatic change on the distributions and diversity patterns of European mammals. *Biodiversity and Conservation*, 16, 3803-816.
- Lichstein, JW, Simons, TR, Shriner, SA & Franzreb, KE (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, 72, 445-63.

- Litt, T, Schölzel, C, Kühl, N & Brauer, A (2009) Vegetation and climate history in the Westeifel Volcanic Field (Germany) during the last 11,000 years based on annually laminated lacustrine maar sediments. *Boreas*, 38, 679-390.
- Liu, Y, Jackson, ST, Brewer, S & Williams, JW (2010) Assessing antiquity and turnover of terrestrial ecosystems in eastern North America using fossil pollen data: a preliminary study. *Earth and Environmental Sciences*, 9, 10.1088/1755-1315/9/1/012005.
- Livingstone, DM & Lotter, AF (1998) The relationship between air and water temperatures in lakes of the Swiss Plateau: a case study with palaeolimnological implications. *Journal of Paleolimnology*, 19, 181-98.
- Lobo, JM, Jiménez-Valverde, A & Hortal, J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33, 103-14.
- Lookingbill, TR & Urban, DL (2003) Spatial estimation of air temperature differences for landscape-scale studies in montane environments. *Agricultural and Forest Meteorology*, 114, 141-51.
- Lotter, AF & Birks, HJB (2003) The Holocene palaeolimnology of Sägistalsee and its environmental history - a synthesis. *Journal of Paleolimnology*, 30, 333-42.
- Lotter, AF, Birks, HJB, Eicher, U, Hofmann, W, Schwander, J & Wick, L (2000) Younger Dryas and Allerød summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. *Palaeogeography Palaeoclimatology Palaeoecology*, 159, 349-61.
- Lotter, AF, Birks, HJB, Hofmann, W & Marchetto, A (1997) Modern diatom, cladocera, chironomid, and chrysophyte cyst assemblages as quantitative indicators for the reconstruction of past environmental conditions in the Alps. I. Climate. *Journal of Paleolimnology*, 18, 395-420.
- Lotter, AF, Birks, HJB, Hofmann, W & Marchetto, A (1998) Modern diatom, cladocera, chironomid, and chrysophyte cyst assemblages as quantitative indicators for the reconstruction of past environmental conditions in the Alps. II. Nutrients. *Journal of Paleolimnology*, 19, 443-63.
- Lotter, AF, Heiri, O, Brooks, SJ, van Leeuwen, JFN, Eicher, U & Ammann, B (2010) Rapid summer temperature changes during Termination 1a: high-resolution multi-proxy climate reconstructions from Gerzensee (Switzerland). *Quaternary Science Reviews*, (in press).
- Lundquist, JD & Cayan, DR (2007) Surface temperature patterns in complex terrain: daily variations and long-term change in the central Sierra Nevada, California. *Journal of Geophysical Research*, 112, D11124.
- Luoto, M & Heikkinen, RK (2008) Disregarding topographical heterogeneity biases species turnover assessments based on bioclimatic models. *Global Change Biology*, 14, 483-94.
- Luoto, M, Pöyry, J, Heikkinen, RK & Saarinen, K (2005) Uncertainty of bioclimate envelope models based on the geographical distribution of species. *Global Ecology and Biogeography*, 14, 575-84.
- Luoto, M, Virkkala, R & Heikkinen, RK (2007) The role of land cover in bioclimatic models depends on spatial resolution. *Global Ecology and Biogeography*, 16, 34-42.
- Luoto, T (2010) Hydrological change in lakes inferred from midge assemblages through the use of an intralake calibration set. *Ecological Monographs*, 80, 303-29.
- Lytle, DE (2005) Palaeoecological evidence of state shifts between forest and barrens on a Michigan sand plain, USA. *The Holocene*, 15, 821-36.
- MacDonald, GM (2010) Global warming and the Arctic: a new world beyond the reach of the Grinnellian niche? *Journal of Experimental Biology*, 213, 855-61.
- MacDonald, GM, Bennett, KD, Jackson, ST, Parducci, L, Smith, FA, Smol, JP & Willis, KJ (2008) Impacts of climate change on species, populations and communities: palaeobiogeographical insights and frontiers. *Progress in Physical Geography*, 32, 139-72.
- Magny, M, Bégeot, C, Guiot, J, Marguet, A & Billaud, Y (2003) Reconstruction and palaeoclimatic interpretation of mid-Holocene vegetation and lake-level changes at Saint-Jorioz, Lake Annecy, French Pre-Alps. *The Holocene*, 13, 265-73.
- Magny, M, Guiot, J & Schoellammer, P (2001) Quantitative reconstruction of Younger Dryas to mid-Holocene paleoclimates at Le Locle, Swiss Jura, using pollen and lake-level data. *Quaternary Research*, 56, 170-80.
- Maher, LJ, Heiri, O & Lotter, AF (2011) Assessment of uncertainties associated with palaeolimnological laboratory methods and microfossil analysis. In: *Tracking Environmental Change Using Lake Sediments Volume 5: Data Handling and Numerical Techniques* (Eds. Birks, HJB, Lotter, AF, Juggins, S & Smol, JP). Springer Dordrecht, (in press).
- Malmgren, BA, Kucera, M, Nyberg, J & Waelbroeck, C (2001) Comparison of statistical and artificial neural network techniques for estimating past sea surface temperatures from planktonic foraminifer census data. *Paleoceanography*, 16, 520-30.
- Malmgren, BA & Nordlund, W (1997) Application of artificial neural networks to paleoceanographic data. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 136, 359-73.
- Markgraf, V, Bradbury, JP & Busby, JR (1986) Palaeoclimates in southwestern Tasmania during the last 13,000 years. *Palaios*, 1, 368-80.
- Marra, MJ, Schulmeister, J & Smith, EGC (2006) Reconstructing temperature during the Last Glacial Maximum from Lyndon Stream, South Island, New Zealand using beetle fossils and maximum likelihood envelopes. *Quaternary Science Reviews*, 25, 1841-49.
- Marra, MJ, Smith, EGC, Schulmeister, J & Leechen, R (2004) Late Quaternary climate change in the Awatere Valley, South Island, New Zealand using a sine model with a maximum likelihood envelope on fossil beetle data. *Quaternary Science Reviews*, 23, 1637-50.
- Martens, H & Næs, T (1989) *Multivariate Calibration*. J. Wiley & Sons, Chichester.
- Mateo, RG, Felicísimo, AM & Muñoz, J (2010) Effects of the number of presences on reliability and stability of MARS species distribution models: the importance of regional niche variation and ecological heterogeneity. *Journal of Vegetation Science*, 21, 908-22.
- McAndrews, JH (1966) Postglacial history of prairie, savanna, and forest in northwestern Minnesota. *Memoirs of the Torrey Botanical Club*, 22, 1-72.
- McKenzie, GM & Busby, JR (1992) A quantitative estimate of Holocene climate using a bioclimatic profile of *Nothofagus cunninghamii* Hook. *Journal of Biogeography*, 19, 531-40.
- Midgley, GF, Davies, ID, Albert, CH, Altwegg, R, Hannah, L, Hughes, GO, O'Halloran, LR, Seo, C, Thorne, JH & Thuiller, W (2010) BioMove - an integrative platform simulating the dynamic response of species to environmental change. *Ecography*, 33, 612-6.
- Millar, CI & Woolfenden, W (1999) The role of climate change in interpreting historic variability. *Ecological Applications*, 9, 1207-16.
- Miller, PA, Giesecke, T, Hickler, T, Bradshaw, RHW, Smith, B, Seppä, H, Valdes, PJ & Sykes, MT (2008) Exploring climatic and biotic controls on Holocene vegetation change in Fennoscandia. *Journal of Ecology* 96, 247-59.
- Minckley, TA, Bartlein, PJ, Whitlock, C, Shuman, BN, Williams, JW & Davis, OK (2008) Associations among modern pollen, vegetation, and climate in western North America. *Quaternary Science Reviews*, 27, 1962-91.
- Minckley, TA & Whitlock, C (2000) Spatial variation of modern pollen in Oregon and southern Washington, USA. *Review of Palaeobotany and Palynology*, 112, 97-123.
- Moine, O, Rousseau, D-D, Jolly, D & Vianey-Liaud, M (2002) Paleoclimatic reconstruction using mutual climatic range on terrestrial mollusks. *Quaternary Research*, 57, 162-72.
- Morin, X, Augspurger, C & Chuine, I (2007) Process-based modeling of species' distributions: what limits temperate tree species' range boundaries? *Ecology*, 88, 2280-91.
- Morin, X & Lechowicz, MJ (2008) Contemporary perspectives on the niche that can improve models of species range shifts under climate change. *Biology Letters*, 4, 573-6.
- Morin, X & Thuiller, W (2009) Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, 90, 1301-13.
- Morrill, C & Jacobsen, RM (2005) How widespread were climate anomalies 8200 years ago? *Geophysical Research Letters*, 32, L19701.
- Mosbrugger, V & Utescher, T (1997) The co-existence approach - a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 134, 61-86.

- Mouton, AM, De Baets, B & Goethals, PLM (2010) Ecological relevance of performance criteria for species distribution models. *Ecological Modelling*, 221, 1995-2002.
- Muller, SD, Richard, PJH, Guiot, J, de Beaulieu, J-L & Fortin, D (2003) Postglacial climate in the St Lawrence lowlands, southern Québec: pollen and lake-level evidence. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 193, 51-72.
- Nakagawa, T, Tarasov, PE, Nishidac, K, Gotanda, K & Yasuda, Y (2002) Quantitative pollen-based climate reconstruction in central Japan: Application to surface and Late-Quaternary spectra. *Quaternary Science Reviews*, 23, 225-44.
- Neumann, F, Schölzel, C, Litt, T, Hense, A & Stein, M (2006) Holocene vegetation history and climate history of the north Golan Heights (Near East). *Vegetation History and Archaeobotany*, 16, 329-46.
- New, M, Hulme, M & Jones, P (1999) Representing twentieth century space-time climate variability. Part 1: Development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate*, 12, 829-56.
- Normand, S, Treier, WA, Randin, CF, Vittoz, P, Guisan, A & Svenning, J-C (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography*, 18, 437-49.
- Norris, JR, Jackson, ST & Betancourt, JL (2006) Classification tree and minimum-volume ellipsoid analyses of the distribution of ponderosa pine in the western USA. *Journal of Biogeography*, 33, 342-60.
- Ortu, E, Brewer, S & Peyron, O (2006) Pollen-inferred palaeoclimate reconstructions in mountain areas: problems and perspectives. *Journal of Quaternary Science*, 21, 615-27.
- Ortu, E, David, F & Peyron, O (2010) Pollen-inferred palaeoclimate reconstruction in the Alps during the Lateglacial and the early Holocene: how to estimate the effect of elevation and local parameters. *Journal of Quaternary Science*, 25, 651-61.
- Ortu, E, Peyron, O, Bordon, A, de Beaulieu, JL, Siniscalco, C & Caramiello, R (2008) Lateglacial and Holocene climate oscillations in the South-western Alps: an attempt at quantitative reconstruction. *Quaternary International*, 190, 71-88.
- Osborn, TJ & Briffa, KR (2006) The spatial extent of 20<sup>th</sup>-century warmth in the context of the past 1200 years. *Science*, 311, 841-4.
- Overpeck, JT, Webb, T & Prentice, IC (1985) Quantitative interpretation of fossil pollen spectra - dissimilarity coefficients and the method of modern analogs. *Quaternary Research*, 23, 87-108.
- Pape, R, Wundram, D & Löffler, J (2009) Modelling near-surface temperature conditions in high mountain environments: an appraisal. *Climate Research*, 39, 99-109.
- Pearman, PB, Guisan, A, Brännimann, O & Randin, CF (2008a) Niche dynamics in space and time. *Trends in Ecology and Evolution*, 23, 149-58.
- Pearman, PB, Randin, CF, Brännimann, O, Vittoz, P, van der Knaap, WO, Engler, R, Le Lay, G, Zimmermann, NE & Guisan, A (2008b) Prediction of plant species distributions across six millennia. *Ecology Letters*, 11, 357-69.
- Pearson, RG & Dawson, TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? *Global Ecology and Biogeography*, 12, 361-71.
- Pearson, RG & Dawson, TP (2004) Bioclimate envelope models: what they detect and what they hide - response to Hampe (2004). *Global Ecology and Biogeography*, 13, 471-3.
- Peterson, AT, Barve, N, Bini, LM & et al (2009) The climate envelope may not be empty. *Proceedings of the National Academy of Sciences of the USA*, 106, E47.
- Peyron, O, Bégeot, C, Brewer, S, Heiri, O, Magny, M, Millet, L, Ruffaldi, P, van Campo, E & Yu, G (2005) Lateglacial climatic changes in Eastern France (Lake Lautrey) from pollen, lake-levels, and chironomids. *Quaternary Research*, 64, 197-211.
- Peyron, O, Guiot, J, Cheddadi, R, Tarasov, P, Reille, M, de Beaulieu, J-L, Bottema, S & Andrieu, V (1998) Climatic reconstruction of Europe for 18,000 yr BP from pollen data. *Quaternary Research*, 49, 183-96.
- Peyron, O, Jolly, D, Bonnefille, R, Vincens, A & Guiot, J (2000) Climate of East Africa 6000 <sup>14</sup>C yr BP as inferred from pollen data. *Quaternary Research*, 54, 90-101.
- Phillips, SJ & Dudik, M (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31, 161-75.
- Phillips, SJ & Elith, J (2010) POC plots: calibrating species distribution models with presence-only data. *Ecology*, 91, 2476-84.
- Phillips, SJ, Dudik, M, Elith, J, Graham, CH, Lehmann, A, Leathwick, JR & Ferrier, S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19, 181-97.
- Pigot, AL, Owens, IPF & Orne, CDL (2010) The environmental limits to geographic range expansion in birds. *Ecology Letters*, 13, 705-15.
- Pigott, CD (1970) The response of plants to climate and climate change. In: *The Flora of a Changing Britain* (Ed. Pering, FH). Claxsey Hampton, pp. 32-44.
- Pigott, CD (1975) Experimental studies on the influence of climate on the geographical distribution of plants. *Weather*, 30, 82-90.
- Pigott, CD (1982) Are the distributions of species determined by failure to set seed? In: *Fruit and Seed Production* (Eds. Marshall, C & Grace, J). Cambridge University Press Cambridge, pp. 203-16.
- Porch, N (2010) Climate space, bioclimatic envelopes and coexistence methods for the reconstruction of past climates: a method using Australian beetles and significance for Quaternary reconstructions. *Quaternary Science Reviews*, 29, 633-47.
- Porch, N, Jordan, GJ, Prince, DM, Barnes, RW, Macphail, MK & Pemberton, M (2009) Last interglacial climates of south-eastern Australia: plant and beetle-based reconstructions from Yarra Creek, King Island, Tasmania. *Quaternary Science Reviews*, 28, 3197-210.
- Prell, WL (1985) The stability of low-latitude sea-surface temperatures: an evaluation of the CLIMAP reconstruction with emphasis on the positive SST anomalies. In: *Technical Report TE025*. Department of Energy Washington DC, pp. 1-60.
- Prentice, IC (1985) Pollen representation, source area, and basin size - toward a unified theory of pollen analysis. *Quaternary Research*, 23, 76-86.
- Prentice, IC, Bartlein, PJ & Webb, T (1991) Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology*, 72, 2038-56.
- Prentice, IC, Guiot, J, Huntley, B, Jolly, D & Cheddadi, R (1996) Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. *Climate Dynamics*, 12, 185-94.
- Proctor, MCF (2009) Desiccation tolerance in some British ferns. *Fern Gazette*, 18, 216-34.
- Pross, J & Klotz, S (2002) Palaeotemperature calculations from the Praetiglian/Tiglian (Plio-Pleistocene) pollen record of Lieth, northern Germany: implications for the climatic evolution of NW Europe. *Global and Planetary Change*, 34, 253-67.
- Pross, J, Klotz, S & Mosbrugger, V (2000) Reconstructing palaeotemperatures for the Early and Middle Pleistocene using the mutual climatic range method based on plant fossils. *Quaternary Science Reviews*, 19, 1785-99.
- Racca, JMJ, Gregory-Eaves, J, Pienitz, R & Prairie, YT (2004) Tailoring paleolimnological diatom-based transfer functions. *Canadian Journal of Fisheries and Aquatic Science*, 61, 2440-54.
- Racca, JMJ, Philibert, A, Racca, R & Prairie, YT (2001) A comparison between diatom-based pH inference models using artificial neural networks (ANN), weighted averaging (WA) and weighted averaging partial least squares (WA-PLS) regressions. *Journal of Paleolimnology*, 26, 411-22.
- Racca, JMJ, Wild, M, Birks, HJB & Prairie, YT (2003) Separating wheat from chaff: Diatom taxon selection using an artificial neural network pruning algorithm. *Journal of Paleolimnology*, 29, 123-33.
- Randin, CF, Dirnböck, T, Dullinger, S, Zimmermann, NE, Zappa, M & Guisan, A (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, 33, 1689-703.
- Randin, CF, Engler, R, Normand, S, Zappa, M, Zimmermann, NE, Pearman, PB, Vittoz, P, Thuiller, W & Guisan, A (2009a) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, 15, 1557-69.
- Randin, CF, Vuissoz, G, Liston, GE, Vittoz, P & Guisan, A (2009b) Introduction of snow and geomorphic disturbances into predictive models of alpine plant distribution in the western Swiss Alps. *Arctic, Antarctic, and Alpine Research*, 41, 347-61.

- Rangel, TF, Diniz-Filho, JA & Bini, LM (2006) Towards and integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography*, 15, 321-7.
- Real, R, Márquez, AL, Olivero, J & Estrada, A (2010) Species distribution models in climate change scenarios are still not useful for informing policy planning: an uncertainty assessment using fuzzy logic. *Ecography*, 33, 304-14.
- Rehfeldt, GE, Crookston, NL, Warwell, MV & Evans, JS (2006) Empirical analyses of plant-climate relationships for the Western United States. *International Journal of Plant Science*, 16, 1123-50.
- Renberg, I & Hellberg, T (1982) The pH history of lakes in southwestern Sweden as calculated from the subfossil diatom flora in the sediments. *Ambio*, 11, 30-3.
- Renssen, H & Isarin, RFB (2001) The two major warming phases of the last deglaciation at ~14.7 and ~ 11.5 ka cal. BP in Europe: climate reconstruction and AGCM experiments. *Global and Planetary Change*, 30, 117-53.
- Renssen, H, Seppä, H, Heiri, O, Roche, DM, Gasse, H & Fichefet, T (2009) The spatial and temporal complexity of the Holocene thermal maximum. *Nature Geoscience*, 2, 411-4.
- Rioual, P, Andrieu-Ponel, V, Rietti-Shati, M, Battarbee, RW, De Beaulieu, JL, Cheddadi, R, Reille, M, Svobodova, H & Shemesh, A (2001) High-resolution record of climate stability in France during the last interglacial period. *Nature*, 413, 293-6.
- Robertson, I, Lucy, D, Baxter, L, Pollard, AM, Aykroyd, RG, Barker, AC, Carter, AHC, Sirotsur, VR & Waterhouse, JS (1999) A kernel-based Bayesian approach to climatic reconstruction. *The Holocene*, 9, 495-500.
- Rolland, C (2003) Spatial and seasonal variations of air temperature lapse rates in alpine regions. *Journal of Climate*, 16, 1032-46.
- Roubicek, AJ, van der Wal, J, Beaumont, LJ, Pitman, AJ, Wilson, P & Hughes, L (2010) Does the choice of climate baseline matter in ecological niche modelling? *Ecological Modelling*, 221, 2280-6.
- Roux, M (1979) Estimation des paléoclimates d'après l'écologie des foraminifères. *Cahiers de l'Analyse des Données*, 4, 61-79.
- Rull, V (2010) Ecology and palaeoecology: two approaches, one objective. *The Open Ecology Journal*, 3, 1-5.
- Rymer, L (1978) The use of uniformitarianism and analogy in palaeoecology, particularly pollen analysis. In: *Biology and Quaternary Environments* (Eds. Walker, D & Guppy, JC). Australian Academy of Sciences, Canberra, pp. 245-57.
- Sagarin, RD, Gaines, SD & Gaylord, B (2006) Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology and Evolution*, 21, 524-30.
- Sakai, A (1971) Freezing resistance of relicts from the Arcto-Tertiary flora. *New Phytologist*, 70, 1199-205.
- Sakai, A & Weiser, CJ (1973) Freezing resistance of trees in North America with reference to tree regions. *Ecology*, 54, 118-26.
- Samuelsson, G (1916) Über den Rückgang der Haselgrenze und anderer pflanzengeographischer Grenzlinien in Skandinavien. *Bulletin of the Geological Institute of Uppsala*, 13, 93-114.
- Sawada, M, Viau, AE, Vettoretti, G, Peltier, WR & Gajewski, K (2004) Comparison of North American pollen-based temperature and global lake-status with CCMA AGCMZ output at 6 ka. *Quaternary Science Reviews*, 23, 225-44.
- Scherrer, D & Körner, C (2010) Infra-red thermometry of alpine landscapes challenges climatic warming projections. *Global Change Biology*, 16, 2602-13.
- Scherrer, D & Körner, C (2011) Topography controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. *Journal of Biogeography*, 38, 406-16.
- Schmidt, GA (2010) Enhancing the relevance of palaeoclimate model/data comparisons for assessments of future climate changes. *Journal of Quaternary Science*, 25, 79-87.
- Scott, GH (1963) Uniformitarianism, the uniformity of nature, and paleoecology. *New Zealand Journal of Geology and Geophysics*, 6, 510-27.
- Seppä, H & Bennett, KD (2003) Quaternary pollen analysis: recent progress in palaeoecology and palaeoclimatology. *Progress in Physical Geography*, 27, 548-79.
- Seppä, H & Birks, HJB (2001) July mean temperature and annual precipitation trends during the Holocene in the Fennoscandian tree-line area: pollen-based climate reconstructions. *The Holocene*, 11, 527-39.
- Seppä, H & Birks, HJB (2002) Holocene climate reconstructions from the Fennoscandian tree-line area based on pollen data from Toskajavri. *Quaternary Research*, 57, 191-9.
- Seppä, H, Birks, HJB, Odland, A, Poska, A & Veski, S (2004) A modern pollen-climate calibration set from northern Europe: developing and testing a tool for palaeoclimatological reconstructions. *Journal of Biogeography*, 31, 251-67.
- Seppä, H & Poska, A (2004) Holocene annual mean temperature changes in Estonia and their relationship to solar insolation and atmospheric circulation patterns. *Quaternary Research*, 61, 22-31.
- Shafer, SL, Bartlein, PJ & Thompson, RS (2001) Potential changes in the distributions of western North America tree and shrub taxa under future climate scenarios. *Ecosystems*, 4, 200-15.
- Sharpe, SE (2002) Constructing seasonal climograph overlap envelopes from Holocene packrat midden contents, Dinosaur National Monument, Colorado. *Quaternary Research*, 57, 306-13.
- Shuman, BN, Henderson, AK, Plank, C, Stefanova, I & Ziegler, SS (2009) Woodland-to-forest transition during prolonged drought in Minnesota after ca. AD 1300. *Ecology*, 90, 2792-807.
- Shuman, BN, Newby, P, Huang, YS & Webb, T (2004) Evidence for the close climatic control of New England vegetation history. *Ecology*, 85, 1297-310.
- Simpson, GL (2007) Analogue methods in palaeoecology: using the analogue package. *Journal of Statistical Software*, 22, 1-29.
- Simpson, GL (2011) Analogue methods in palaeolimnology. In: *Tracking Environmental Change Using Lake Sediments Volume 5: Data Handling and Numerical Techniques* (Eds. Birks, HJB, Lotter, AF, Juggins, S & Smol, JP). Springer Dordrecht, (in press).
- Sinka, KJ & Atkinson, TC (1989) A mutual climatic range method for reconstructing palaeoclimate from plant remains. *Journal of the Geological Society, London*, 156, 381-96.
- Smart, SM, Scott, A, Whitaker, J, Hill, MO, Roy, DB, Critchley, CN, Marini, L, Evans, C, Emmett, BA, Rowe, EC, Crowe, A, Le Duc, M & Marrs, RH (2010) Empirical realised niche models for British higher and lower plants - development and preliminary testing. *Journal of Vegetation Science*, 21, 643-56.
- Šmilauer, P & Birks, HJB (1995) The use of generalised additive models in the description of diatom-environment response surfaces. *Geological Survey of Denmark (DGU) Service Report*, 7, 42-7.
- Smol, JP (2008) *Pollution of Lakes and Rivers*. Second edn. Blackwell, Oxford.
- Smulders, M, Nelson, TA, Jelinski, DE, Nielsen, SE & Stenhouse, GB (2010) A spatially explicit method for evaluating accuracy of species distribution models. *Diversity and Distributions*, 16, 996-1008.
- Soberón, J (2007) Grinnellian and Eltonian niches and geographic distribution of species. *Ecology Letters*, 10, 1115-23.
- Soberón, J & Nakamura, M (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the USA*, 106, 19644-50.
- Soberón, J & Peterson, AT (2005) Interpretation of models of fundamental ecological niches and species distributional areas. *Biodiversity Informatics*, 2, 1-10.
- St, Jacques, J-M, Cumming, BF & Smol, JP (2008) A pre-European settlement pollen-climate calibration set for Minnesota, USA: developing tools for palaeoclimatic reconstructions. *Journal of Biogeography*, 35, 306-24.
- Sugita, S (1994) Pollen representation of vegetation in Quaternary sediments - theory and method in patchy vegetation. *Journal of Ecology*, 82, 881-97.
- Svenning, J-C, Normand, S & Skov, F (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, 31, 316-26.
- Svenning, J-C, & Skov, F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, 7, 565-73.
- Svenning, J-C, & Skov, F (2005) The relative roles of environment and history as controls of tree species composition and richness in Europe. *Journal of Biogeography*, 32, 1019-33.
- Svenning, J-C, & Skov, F (2007) Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, 10, 453-60.
- Swetnam, TW, Allen, CD & Betancourt, JL (1999) Applied historical ecology: using the past to manage the future. *Ecological Applications*, 9, 1189-206.

- Tabor, K & Williams, JW (2010) Globally downscaled climate projections for assessing the conservation impacts of climate change. *Ecological Applications*, 20, 554-65.
- Tarasov, P, Bezrukova, E, Karabanov, E, Nakagawa, T, Wagner, M, Kulagina, N, Letunova, P, Abzaeva, A, Granoszewski, W & Reidel, F (2007) Vegetation and climate dynamics during the Holocene and Eemian interglacials derived from Lake Baikal pollen records. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 252, 440-57.
- Tarasov, P, Guiot, J, Cheddadi, R, Andreev, AA, Bezusko, LG, Blyakharchuk, TA, Dorofeyuk, NI, Filimonova, LV, Volkova, VS & Zernitskaya, VP (1999b) Climate in northern Eurasia 6000 years ago reconstructed from pollen data. *Earth and Planetary Science Letters*, 171, 635-45.
- Tarasov, P, Peyron, O, Guiot, J, Brewer, S, Volkova, VS, Bezusko, LG, Dorofeyuk, NI, Kvavadze, EV, Osipova, IM & Panova, NK (1999a) Late glacial maximum climate of the former Soviet Union and Mongolia reconstructed from pollen and plant macrofossil data. *Climate Dynamics*, 15, 227-40.
- Tarasov, P, Williams, JW, Andreev, A, Nakagawa, T, Bezrukova, E, Herzschuh, U, Igareshi, Y, Müller, S, Werner, K & Zheng, Z (2007) Satellite- and pollen-based quantitative woody cover reconstructions for northern Asia: verification and application to late-Quaternary pollen data. *Earth and Planetary Science Letters*, 264, 284-98.
- Tarrosó, P & Rebelo, H (2010) E-Clic - easy climate data converter. *Ecography*, 33, 617-20.
- Telford, RJ (2006) Limitations of dinoflagellate cyst transfer functions. *Quaternary Science Reviews*, 25, 1375-82.
- Telford, RJ, Andersson, C, Birks, HJB & Juggins, S (2004) Biases in the estimation of transfer function prediction errors. *Paleoceanography*, 19, PA4014.
- Telford, RJ & Birks, HJB (2005) The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance. *Quaternary Science Reviews*, 24, 2173-9.
- Telford, RJ & Birks, HJB (2009) Design and evaluation of transfer functions in spatially structured environments. *Quaternary Science Reviews*, 28, 1309-16.
- ter Braak, CJF (1987) *CANOCO - a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis*. TNO Institute of Applied Computer Science, Wageningen, The Netherlands, p. 95.
- ter Braak, CJF (1995) Non-linear methods for multivariate statistical calibration and their use in palaeoecology: a comparison of inverse ( $k$ -nearest neighbours, partial least squares and weighted averaging partial least squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems*, 28, 165-80.
- ter Braak, CJF (1996) *Unimodal Models to Relate Species to Environment*. DLO-Agricultural Mathematics Group, Wageningen.
- ter Braak, CJF & Barendregt, LG (1986) Weighted averaging of species indicator values: its efficiency in environmental calibration. *Mathematical Biosciences*, 78, 57-72.
- ter Braak, CJF & Gremmen, NJM (1987) Ecological amplitudes of plant species and the internal consistency of Ellenberg's indicator values for moisture. *Plant Ecology*, 69, 79-87.
- ter Braak, CJF & Juggins, S (1993) Weighted Averaging Partial Least-Squares Regression (WA-PLS) - an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia*, 269, 485-502.
- ter Braak, CJF, Juggins, S, Birks, HJB & van der Voet, H (1993) Weighted averaging partial least squares regression (WA-PLS): definition and comparison with other methods for species-environment calibration. In: *Multivariate Environmental Statistics* (Eds. Patil, GP & Rao, CR). Elsevier, Amsterdam, pp. 529-60.
- ter Braak, CJF & Looman, CWN (1986) Weighted averaging, logit regression and the Gaussian response model. *Vegetatio*, 65, 3-11.
- ter Braak, CJF & Looman, CWN (1987) Regression. In: *Data Analysis in Community and Landscape Ecology* (Eds. Jongman, RHG, ter Braak, CJF & Tongeren, OFR). Pudoc Press, Wageningen, pp. 29-90.
- ter Braak, CJF & Prentice, IC (1988) A theory of gradient analysis. *Advances in Ecological Research*, 18, 271-317.
- ter Braak, CJF & van Dam, H (1989) Inferring pH from diatoms - a comparison of old and new calibration methods. *Hydrobiologia*, 178, 209-23.
- ter Braak, CJF, van Dobben, H & di Bella, G (1996) On inferring past environmental change from species composition data by nonlinear reduced rank models. In: *XVIII International Biometric Conference, July 1-5 1996*, pp. 65-70.
- ter Braak, CJF & Verdonschot, PFM (1995) Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquatic Sciences* 57, 255-89.
- Thompson, RS, Anderson, KH & Bartlein, PJ (1999a) Atlas of relations between climatic parameters and distribution of important trees and shrubs in North America - Introduction and Conifers. *United States Geological Survey Professional Paper*, 1650-A, pp. 269.
- Thompson, RS, Anderson, KH & Bartlein, PJ (1999b) Atlas of relations between climatic parameters and distribution of important trees and shrubs in North America - Hardwoods. *United States Geological Survey Professional Paper*, 1650-B, pp. 423.
- Thompson, RS, Anderson, KH & Bartlein, PJ (1999c) Quantitative paleoclimatic reconstructions from Late Pleistocene plant macrofossils of the Yucca Mountain Range. *United States Geological Survey Open-File Report*, 99-338, 1-38.
- Thompson, RS, Anderson, KH & Bartlein, PJ (2000) Atlas of relations between climatic parameters and distribution of important trees and shrubs in North America. Additional Conifers, Hardwoods, and Monocots. *United States Geological Survey Professional Paper*, 1650-C, pp. 386.
- Thompson, RS, Anderson, KH & Bartlein, PJ (2001) Atlas of relations between climatic parameters and distribution of important trees and shrubs in North America. Alaska Species, Ecoregions, and Hardiness Zones. *United States Geological Survey Professional Paper*, 1650-D.
- Thompson, RS, Anderson, KH & Bartlein, PJ (2008) Quantitative estimation of bioclimatic parameters from presence/absence vegetation data in North America by the modern analog technique. *Quaternary Science Reviews*, 27, 1234-54.
- Thuiller, W, Albert, C, Araújo, MB, Berry, PM, Cabeza, M, Guisan, A, Hickler, T, Midgley, GF, Paterson, J, Schurr, FM, Sykes, MT & Zimmermann, NE (2008) Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology, Evolution, and Systematics*, 9, 137-52.
- Thuiller, W, Araújo, MB & Lavorel, S (2003) Generalized models vs classification tree analysis: predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, 14, 669-80.
- Thuiller, W, Lafourcade, B, Engler, R & Araújo, M (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-73.
- Thuiller, W, Lavorel, S, Araújo, M, Sykes, MT & Prentice, IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences USA*, 102, 8245-50.
- Thuiller, W, Lavorel, S, Sykes, MT & Araújo, MB (2006) Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions*, 12, 49-60.
- Titeux, N, Maes, D, Marmion, M, Luoto, M & Heikkinen, RK (2009) Inclusion of soil data improves the performance of bioclimatic envelope models for insect species distributions in temperate Europe. *Journal of Biogeography*, 36, 1459-73.
- Toivonen, HTT, Mannila, H, Korhola, A & Olander, H (2001) Applying Bayesian statistics to organism-based environmental reconstructions. *Ecological Applications*, 11, 618-30.
- Tonello, MS, Mancini, MV & Seppä, H (2009) Quantitative reconstruction of Holocene precipitation changes in southern Patagonia. *Quaternary Research*, 72, 410-20.
- Trachsel, M, Grosjean, M, Larocque-Tobler, I, Schwikowski, M, Blass, A & Sturm, B (2010) Quantitative summer temperature reconstruction derived from a combined biogenic Si and chironomid record from varved sediments of Lake Silvaplana (south-eastern Swiss Alps) back to AD 1177. *Quaternary Science Reviews*, 29, 2719-30.
- Trenbarth, KE & Otto-Bliesner, BL (2003) Toward integrated reconstructions of past climates. *Science*, 300, 589-91.

- Tweiten, MA, Hotchkiss, SC, Booth, RK, Calcote, RR & Lynch, EA (2009) The response of a jack pine forest to late-Holocene climate variability in northwestern Wisconsin. *The Holocene*, 19, 1049-61.
- van der Voet, H (1994) Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems*, 25, 313-23.
- Vasko, K, Toivonen, HTT & Korhola, A (2000) A Bayesian multinomial Gaussian response model for organism-based environmental reconstructions. *Journal of Paleolimnology*, 24, 243-50.
- Velle, G, Brodersen, KP, Birks, HJB & Willassen, E (2010b) Midges as quantitative temperature indicator species: lessons for palaeoecology. *The Holocene*, 10.1177/0959683610365933.
- Velle, G, Kongshavn, K & Birks, HJB (2010a) Minimizing the edge-effect in environmental reconstructions by trimming the calibration-set: chironomid-inferred temperatures from Spitsbergen. *The Holocene*, (in press).
- Verbruggen, F, Heiri, O, Meriläinen, JJ & Lotter, AF (2010) Subfossil chironomid assemblages in deep, stratified European lakes: relationships with temperature, trophic state and oxygen. *Freshwater Biology*, 10.1111/j.1365-2427.2010.02508.x.
- Virah-Sawmy, M, Bonsall, MB & Willis, KJ (2009c) 'Tales of *Symphonia*': extinction dynamics in response to past climate change in Madagascan rainforests. *Biology Letters*, 5, 821-5.
- Virah-Sawmy, M, Gillson, L & Willis, KJ (2009a) How does spatial heterogeneity influence resilience to climatic change? Ecological dynamics in southeast Madagascar. *Ecological Monographs*, 79, 557-74.
- Virah-Sawmy, M, Willis, KJ & Gillson, L (2009b) Threshold response of Madagascar's littoral forest to sea-level rise. *Global Ecology and Biogeography*, 18, 98-110.
- Virah-Sawmy, M, Willis, KJ & Gillson, L (2010) Evidence for drought and forest declines during the recent megafaunal extinctions in Madagascar. *Journal of Biogeography*, 37, 506-19.
- von Grafenstein, U, Erlenkauer, H & Trumborn, P (1999) Oxygen and carbon isotopes in modern fresh-water ostracod valves: assessing vital offsets and autecological effects of interest for palaeoclimate studies. *Palaeogeography Palaeoclimatology Palaeoecology*, 148, 133-52.
- von Post, L (1946) The prospect for pollen analysis in the study of the Earth's climatic history. *New Phytologist*, 45, 193-217.
- Waelbroeck, C, Labeyrie, L, Duplessy, J-C, Guiot, J, Labracherie, M, Leclaire, H & Duprat, J (1998) Improving past sea surface temperature estimates based on planktonic fossil faunas. *Paleoceanography*, 13, 272-83.
- Walker, IR, Levesque, AJ, Cwynar, LC & Lotter, AF (1997) An expanded surface-water palaeotemperature inference model for use with fossil midges from eastern Canada. *Journal of Paleolimnology*, 18, 165-78.
- Walther, G-R, Berger, S & Sykes, MT (2005) An ecological 'footprint' of climate change. *Proceedings of the Royal Society B*, 272, 1427-32.
- Ward, G, Hastie, T, Barry, S, Elith, J & Leathwick, JR (2009) Presence-only data and the EM algorithm. *Biometrics*, 65, 554-63.
- Warren, DL, Glor, RE & Turelli, M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, 33, 607-11.
- Watts, WA, Allen, JRM & Huntley, B (1996a) Vegetation history and palaeoclimate of the last glacial period at Lago Grande di Monticchio, southern Italy. *Quaternary Science Reviews*, 15, 133-53.
- Watts, WA, Allen, JRM, Huntley, B & Fritz, SC (1996b) Vegetation history and climate of the last 15,000 years at Lago Grande di Monticchio, southern Italy. *Quaternary Science Reviews*, 15, 133-53.
- Webb, RS, Anderson, KH & Webb III, T (1993a) Pollen response-surface estimates of late-Quaternary changes in the moisture balance of the northeastern United States. *Quaternary Research*, 40, 213-27.
- Webb, T, Anderson, KH, Bartlein, PJ & Webb, RS (1998) Late Quaternary climate change in eastern North America: a comparison of pollen-derived estimates with climate model results. *Quaternary Science Reviews*, 17, 587-606.
- Webb, T, Bartlein, PJ, Harrison, SP & Anderson, KH (1993b) Vegetation, lake levels, and climate in eastern North America for the past 18,000 years. In: *Global Climates since the Last Glacial Maximum* (Eds. Wright, HE, Kutzbach, J, Webb III T, Ruddiman, WF, Street-Perrot, FA & Bartlein, PJ). University of Minnesota Press, Minneapolis, pp. 415-67.
- Webb, T, Bartlein, PJ & Kutzbach, J (1987) Climatic change in eastern North America during the past 18,000 years: comparisons of pollen data with model results. In: *The Geology of North America. North America and Adjacent Oceans during the Last Deglaciation* (Eds. Ruddiman, WF & Wright, HE). Geological Society of America, Boulder, Colorado, pp. 447-62.
- Webb, T & Bryson, RA (1972) Late- and post-glacial climate change in Northern Midwest, USA: Quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. *Quaternary Research* 2, 70-115.
- Weckström, J, Korhola, A, Erästä, P & Holmström, L (2006) Temperature patterns over the last eight centuries in northern Fennoscandia inferred from sedimentary diatoms. *Quaternary Research*, 66, 78-86.
- Whitmore, J, Gajewski, K, Sawada, M, Williams, JW, Shuman, PJ, Minchley, T, Viau, AE, Webb III, T, Shafer, SL, Anderson, P & Brubaker, L (2005) Modern pollen data from North America and Greenland for multi-scale paleoenvironmental reconstructions. *Quaternary Science Reviews*, 24, 1828-48.
- Wiederholm, T (1980) Use of benthos in lake monitoring. *Journal of the Water Pollution Control Federation*, 52, 537-47.
- Wiens, JJ, Ackerly, DD, Allen, AP, Anacker, BL, Buckley, LB, Cornell, HV, Damschen, EJ, Davies, TJ, Grytnes, J-A, Harrison, SP, Hawkins, BA, Holt, RD, McCain, C & Stephens, PR (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, 13, 1310-24.
- Wiens, JJ & Graham, CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, 36, 519-39.
- Williams, JW (2002) Variations in tree cover in North America since the last glacial maximum. *Global and Planetary Change*, 35, 1-23.
- Williams, JW, Gonzales, LM & Kaplan, JO (2008) Leaf area index for northern and eastern North America at the Last glacial Maximum - a data-model comparison. *Global Ecology and Biogeography*, 17, 122-34.
- Williams, JW & Jackson, ST (2003) Palynological and AVHRR observations of modern vegetational gradients in eastern North America. *The Holocene*, 13, 485-97.
- Williams, JW & Jackson, ST (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, 5, 475-82.
- Williams, JW & Shuman, BN (2008) Obtaining accurate and precise environmental reconstructions from the modern analog technique and North American surface pollen dataset. *Quaternary Science Reviews*, 27, 669-87.
- Willis, KJ, Bailey, RM, Bhagwat, SA & Birks, HJB (2010a) Biodiversity baselines, thresholds, and resilience: testing predictions and assumptions using palaeoecological data. *Trends in Ecology and Evolution*, 25, 583-91.
- Willis, KJ, Bennett, KD, Bhagwat, SA & Birks, HJB (2010b) 4°C and beyond: what did this mean for biodiversity in the past? *Systematics and Biodiversity*, 8, 3-9.
- Willis, KJ & Bhagwat, SA (2009) Biodiversity and climate change. *Science*, 326, 806-7.
- Willis, KJ & Bhagwat, SA (2010) Questions of importance to the conservation of global biological diversity: Answers from the past. *Climate of the Past Discussions*, 6, 1139-62.
- Willis, KJ & Birks, HJB (2006) What is natural? The need for a long-term perspective in biodiversity conservation. *Science*, 314, 1261-5.
- Willis, KJ, Kleczkowski, A, New, M & Whittaker, RJ (2007) Testing the climate variability on European plant diversity: 320,000 years of water-energy dynamics and its long-term influence on plant taxonomic richness. *Ecology Letters*, 10, 673-9.
- Wold, S (1992) Nonlinear partial least squares modelling II. Spline inner relation. *Chemometrics and Intelligent Laboratory Systems*, 14, 71-84.
- Wright, HE (1967) The use of surface samples in Quaternary pollen analysis. *Review of Palaeobotany and Palynology*, 2, 321-30.
- Wright, HE, McAndrews, JH & van Zeist, W (1967) Modern pollen rain in West Iran, and its relation to plant geography and Quaternary vegetational history. *Journal of Ecology*, 55, 415-43.

- Wu, H, Guiot, J, Brewer, S & Guo, Z (2007a) 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, 211-29.
- Wu, H, Guiot, J, Brewer, S, Guo, Z & Peng, C (2007b) Dominant factors controlling glacial and interglacial variations in the treeline elevation in tropical Africa. *Proceedings of the National Academy of Sciences of the USA*, 104, 9720-4.
- Yates, CJ, Elith, J, Latimer, AM, Le Maitre, D, Midgley, GF, Schurr, FM & West, AG (2010) Projecting climate change impacts on species distributions in megadiverse South African Cape and Southwest Australian floristic regions: opportunities and challenges. *Austral Ecology*, 35, 374-91.
- Yee, TW (2004) A new technique for maximum-likelihood canonical Gaussian ordination. *Ecological Monographs*, 74, 685-701.
- Yee, TW (2006) Constrained additive ordination. *Ecology*, 97, 203-13.
- Yee, TW & Mitchell, ND (1991) Generalized additive models in plant ecology. *Journal of Vegetation Science*, 2, 587-602.
- Zhao, Y & Herzschuh, U (2009) Modern pollen representation of source vegetation in the Qaidam Basin and surrounding mountains, north-eastern Tibetan Plateau. *Vegetation History and Archaeobotany*, 18, 245-60.
- Zhao, Y, Xu, Q, Huang, X, Guo, X & Tao, S (2009) Differences of modern pollen assemblages from lake sediments and surface soils in arid and semi-arid China and their significance for pollen-based quantitative climate reconstructions. *Review of Palaeobotany and Palynology*, 156, 519-24.

---

Received: November 01, 2010

Revised: December 12, 2010

Accepted: December 16, 2010

© Birks *et al.*; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted, non-commercial use, distribution & reproduction in any medium, provided the work is properly cited.