

Testing a new multigroup inference approach to reconstructing past environmental conditions

Roy THOMPSON*, Christian KAMENIK¹⁾, Roland SCHMIDT²⁾, Sergi PLA³⁾, Maria RIERADEVALL⁴⁾ and Jordi CATALAN⁵⁾

School of GeoSciences, The University of Edinburgh, West Mains Road, Edinburgh EH93JW, UK

¹⁾Oeschger Centre for Climate Change Research & Institute of Geography, University of Bern, Erlachstrasse 9a, Building 3, 3012 Bern, Switzerland

²⁾Austrian Academy of Science, Institute of Limnology, Mondseestraße 9, 5310 Mondsee, Austria

³⁾Department of Geography, Loughborough University, Loughborough, Leicestershire, LE11 3TU UK

⁴⁾Departament d'Ecologia, Facultat de Biologia, Universitat de Barcelona, Diagonal 645 - 08028 Barcelona, Spain

⁵⁾CSIC-UB Limnology Group, Centre for Advanced Studies of Blanes (CEAB), CSIC, Accés Cala St. Francesc 14, Blanes 17300, Spain

*e-mail corresponding author: roy@ed.ac.uk

ABSTRACT

A new, quantitative, inference model for environmental reconstruction (transfer function), based for the first time on the simultaneous analysis of multigroup species, has been developed. Quantitative reconstructions based on palaeoecological transfer functions provide a powerful tool for addressing questions of environmental change in a wide range of environments, from oceans to mountain lakes, and over a range of timescales, from decades to millions of years. Much progress has been made in the development of inferences based on multiple proxies but usually these have been considered separately, and the different numeric reconstructions compared and reconciled post-hoc. This paper presents a new method to combine information from multiple biological groups at the reconstruction stage. The aim of the multigroup work was to test the potential of the new approach to making improved inferences of past environmental change by improving upon current reconstruction methodologies. The taxonomic groups analysed include diatoms, chironomids and chrysophyte cysts. We test the new methodology using two cold-environment training-sets, namely mountain lakes from the Pyrenees and the Alps. The use of multiple groups, as opposed to single groupings, was only found to increase the reconstruction skill slightly, as measured by the root mean square error of prediction (leave-one-out cross-validation), in the case of alkalinity, dissolved inorganic carbon and altitude (a surrogate for air-temperature), but not for pH or dissolved CO₂. Reasons why the improvement was less than might have been anticipated are discussed. These can include the different life-forms, environmental responses and reaction times of the groups under study.

Key words: cold environment, lake, diatom, chironomid, chrysophyte cyst

1. INTRODUCTION

Tremendous progress has been made in recent years in the use of species assemblages as proxy indicators for the magnitude of past environmental and climatic change. In particular transfer functions have been developed by modelling the response of modern plant or animal assemblages, preserved in surface-sediment calibration sets, to contemporary environmental gradients (see reviews by Birks 1995, 1998; Charles & Smol 1994; Smol & Cumming 2000). These methods derive from the transfer-function methodology first developed by Imbrie & Kipp (1971) in the analysis of the population of species of various fossil groups, including foraminifera, radiolaria, coccoliths and diatoms. A vast body of literature has accumulated concerning the palaeolimnological and palaeoceanographic (cf. Kucera *et al.*, 2005) value of microfossils in assessing environmental changes, including those of palaeotemperature, nutrient concentration, productivity, palaeosalinity and pH. Numerical techniques and taxonomic resolution

have considerably improved, and large calibration, or training data, sets have been developed. Despite the great advances, improvements are still needed. For example, in Broecker's estimation, temperatures that are accurate to 0.5 °C are needed when trying to reconstruct century-long Holocene temperature fluctuations, but few proxies approach that accuracy (Broecker 2001). Some improvements have taken place in the accuracy of the transfer functions since the publication of the Broecker (2001) paper. For example there are now transfer models based on floral and faunal remains that, at least theoretically, operate within error estimates of ±0.6-0.8°C (e.g., Korhola *et al.* 2002; Seppä & Birks 2001; Pla & Catalan 2005). Nevertheless improving lake-sediment-based methodologies, in order to provide the potential for climate reconstruction work within the error bounds set out by Broecker (2001) remains a challenge. It is in the spirit of striving to improve, or to enhance, current methodologies that our multigroup work has been undertaken.

Palaeolimnological data-modelling tools currently consist of numerical methods for deriving transfer func-

tions from a modern training set of microfossils and an associated dataset of environmental variables. Individual taxonomic groups are matched to individual predictor variables. Improvements to inference models based on multigroup analysis would have a wide applicability in palaeolimnological studies throughout the world and to other disciplines such as palaeoceanography. Kucera *et al.* (2005) in their studies of ocean-sediment cores, for example, address "the challenge that the compilation of a unique multi-proxy" reconstruction currently presents, and deliberate on "possible strategies to do so", given that palaeo-reconstructions of the ocean-surface environment are currently "only feasible on a single-proxy basis". Similarly Charles & Smol (1988) "recommend that ...both chrysophytes and diatoms be analyzed in paleolimnological studies... to quantitatively infer lake water pH", while Rosen *et al.* (2003) urge for "future research to concentrate on ...multiproxy... studies to further understand the relationship between the proxies and climate". Traditionally multi proxy study work involves several proxies being used independently to reconstruct an environmental variable. Many good examples of traditional multi proxy work are summarised in the recent, comprehensive review by Birks & Birks (2006).

Here our aim is to develop an improved transfer-function methodology, with the novel approach of basing the transfer function, for the first time, on the simultaneous analysis of multigroups. We make use of data, including diatoms, chironomids and cysts, previously collected and harmonised as part of large multinational collaborations that involved specialists from Austria, Denmark, England, Finland, Norway, Scotland, Spain, Switzerland and Sweden. We hypothesized that environmental influence on the assemblage data could be better constrained if multiple groups were included in the training sets. To test this idea, we applied the new methodology to two independent mountain-lake datasets, one from the Pyrenees and one from the Central Eastern Alps. In these cold environments we had a particular interest in trying to reconstruct air temperature (which in geographically restricted regions is strongly related to altitude), but also in testing the method on a suite of water chemistry indicators ranging from those that are reasonably well determined by single group reconstructions, e.g., pH, to indicators such as CO₂ that previously have been poorly determined.

2. MATERIALS AND METHODS

2.1. Biological groups considered

The training sets were chosen as being representative of mountain lakes. Major advantages of the Pyrenean and Alpine lake districts were that the lakes had previously been widely studied; and in particular that state-of-the-art, multigroup, assemblage data already existed and that the training sets had already been used in one-group-at-a-time environmental reconstructions.

The taxonomic groups we focus on are (i) diatoms, (ii) chrysophytes and (iii) chironomids, as all are abundant, diverse and essential components of mountain lake biota. Diatom frustules, chrysophycean siliceous cysts, and chitinous larval heads of chironomids are well preserved in lake sediments. They have all been found to be good indicators of past environmental conditions. The selection and combination of these organisms could be expected to allow a holistic view of lake responses to environmental changes, because they represent and span the main habitats (plankton and benthos, littoral and profundal) and trophic levels (primary and secondary producers) found in mountain lakes.

2.2. Laboratory techniques

Standardised laboratory procedures were applied for cleaning, sorting and mounting the groups and for identifying the taxa present.

Diatoms and chrysophytes were cleaned using standard procedures (Battarbee 1986). Diatom slides were prepared for each sediment sample after digestion with hydrogen peroxide and mounted on slides using Naphrax[®], the counting strategy followed the procedure suggested by Battarbee (1986). The taxonomic criteria adopted in the AL:PE (Cameron *et al.* 1999) and SWAP (Stevenson *et al.* 1991) were followed.

For chrysophyte analysis surface sediment samples from the Pyrenees were treated with hydrogen peroxide (33% H₂O₂) and some drops of HCl. The resulting siliceous slurries were mounted in Naphrax (R.I. = 1.7) following Battarbee (1986). Cysts were identified by SEM and then counted under a light microscope. Details on counting, cyst numbering and all cyst descriptions can be found in Pla (2001) and Pla & Catalan (2005) and in Kamenik & Schmidt (2005a). A minimum of 500 cysts were counted per sample. Identification of cysts followed Facher & Schmidt (1996), Pla (2001) and Kamenik *et al.* (2005b). Chrysophyte cysts from the Alps were analysed using a scanning electron microscope (Jeol JSM-35) equipped with an image analyzing system (QUANTEL Crystal). We emphasise that as we never merge the two training sets, there is no problem of using somewhat different counting techniques in the two mountain regions.

Chironomid head capsules were obtained by sieving the sediment through a 90 µm screen after deflocculating with hot 10% KOH for 15 min. The head capsules were hand sorted from a Bolgorov tray with forceps under the magnification of a stereoscopic microscope. The capsules were then mounted in Euparal[®] media, after dehydration with absolute ethanol. Taxa were determined mainly following Wiederholm (1983), Schmid (1993) and Rieradevall & Brooks (2001).

2.3. Environmental variables

We restrict ourselves to testing our new multigroup approach on just five environmental variables namely

pH, alkalinity, dissolved inorganic carbon (DIC), CO₂ and altitude. We wish to examine a reasonable number of variables that are likely to have significant and independent influences on the biological distributions (i.e., are likely to have causal relationships). But on the other hand we recognise that it is important to refrain from data-dredging, and over-exploiting post-hoc correlations and so we need to guard against any proliferation of associations. pH and alkalinity are chosen as two environmental variables which previously have been found to be good correlates (e.g., Cameron *et al.* 1999; Marchetto & Lami 1994; Marchetto & Schmidt 1992). CO₂ is chosen as a chemically important variable which previously has not yielded good associations. DIC is chosen as an environmental variable with a more intermediate response, while altitude is used as an environmental variable which is likely to bring to light any climatic relationship.

Standard, well-tested measurement methods were used for the chemical variables. Details are provided in Catalan *et al.* (1993) and in Kamenik *et al.* (2001).

2.4. Study areas

2.4.1. Pyrenees

The training set consisted of 83 lakes in the Central and Eastern Pyrenees (NE Spain) and was developed during many seasons of field work. The lakes were chosen from among the *ca* 500 in the area, to provide a broad range of altitude [1,615-2,590 m a.s.l. - this altitudinal range represents a mean annual air temperature change of *ca* 6 °C (Agustí-Panareda & Thompson 2002)], size (0.7-45.4 ha), maximum depth (1-80 m), bedrock (major chemistry) and trophic state. The parameter distribution for the lakes surveyed roughly reflects the distribution for the whole set of lakes in the area. Lake water-chemistry mainly reflects catchment bedrock (Catalan *et al.* 1993). Catchment vegetation was generally poorly developed. Alpine meadows were almost always present and constitute the dominant vegetation in the catchments. *Rhododendron ferrugineum*, along with some sparse pines (*Pinus uncinata*), was common in the lower catchments. Catchments were specifically chosen to be glacier free. Counts of the diatom and chrysophyte assemblage preserved in the surface sediments were available at all 83 lakes. For 27 of these lakes chironomid counts were also available.

2.4.2. Alps - Niedere Tauern

We selected 45 lakes along an altitude gradient (1502-2309 m a.s.l.) which corresponds to a change in air-temperature of *c.* 4.6 °C (Agustí-Panareda & Thompson 2002). The lakes stretch from the sub-alpine forested belt up to the mid-alpine zone. They are located in the Central Eastern Alps (Niedere (Schladminger) Tauern; 47°13' - 47°21' N, 13°36' - 14°04' E (Kamenik & Schmidt 2005b) and were chosen to be as pristine as

possible, while encompassing a wide spread of geographic settings, water chemistries and basin morphologies. The lakes ranged in area from 0.6 to 16.8 ha, and from 5.7 to 43.6 m maximum depth. Their catchments of predominantly crystalline bedrock, schists and metamorphic carbonates all lack glaciers. The oligo- to ultra-oligotrophic lakes are not acidified. Their water chemistry is mainly determined by changes in bedrock, from silicate to carbonate minerals, affecting pH, alkalinity, conductivity, calcium and magnesium, and by trophic variables such as total phosphorus, dissolved organic carbon, potassium, total inorganic nitrogen and total dissolved reactive silica (Kamenik *et al.* 2001). Counts of the diatom and chrysophyte assemblages were available at all 45 lakes.

2.5. Numerical methods

The multivariate method of weighted-averaging, partial, least-squares (WA-PLS) ter Braak & Juggins (1993) is one of the most popular numerical methods of inferring the past environment fossil remains. Weighted averaging is a particularly simple, yet often effective, method of ordination (ter Braak & Prentice 1988). The essential operation is that a set of species weights are used to calculate scores for sites. PLS regression is a technique that generalizes and combines features from principal component analysis and multiple regression. Ter Braak & Juggins (1993) combined WA and PLS and showed how their new weighted averaging partial least squares regression technique was able to add additional regression components and improve the predictive power by utilizing the residual structure in the species data.

If we let x denote the environmental variable to be calibrated, with x_i the values of the environmental variable in site i (e.g., lake i), and the $n \times m$ matrix $Y = (y_{ij})$ with y_{ij} the abundance of taxon j in site i ($i = 1 \dots n$ sites and $j = 1 \dots m$ taxa), then in its first step PLS regression searches for a set of components that explain as much as possible of the covariance between x and Y . It is followed by a regression step where the decomposition is used to predict x . Our basic idea is to combine the multiple group data at the time of this regression step. Thus, first of all, weighted averaging is carried out separately for each species group, exactly as in conventional approaches. Next, at the regression step, the k groups ($k = 1 \dots p$) are combined in the WA-PLS (in exactly the same way as additional PLS components in a conventional WA-PLS analysis), and hence the regression decides on the relative importance given to each group. In this way the algorithm uses the data to decide on the best relative weightings between taxonomic groups. In short our new approach is a direct extension of the basic WA-PLS method, and involves only a small development to the R/Splus computing code employed by Thompson *et al.* (2005).

Tab. 1. Range and basic statistics of each environmental variable. One pH unit means the proton concentration is 10^{-1} moles per liter. PYR = Pyrenees; NT = Niedere-Tauern.

Locality	Variable	Unit	N. of lakes	Minimum	1 st Quantile	Median	Mean	3 rd Quantile	Maximum
PYR	pH	U	82	5.45	6.42	6.77	6.69	7.01	7.82
PYR	alkalinity	$\mu\text{eq L}^{-1}$	82	-3.20	41.2	96.5	116	166	400
PYR	DIC	$\mu\text{eq L}^{-1}$	82	3.29	81.4	139	160	218	447
PYR	CO ₂	$\mu\text{eq L}^{-1}$	82	2.94	26.3	38.5	43.7	54.3	187
PYR	Altitude	m a.s.l.	82	1615	2151	2340	2294	2453	2590
NT	pH	U	41	6.90	7.10	7.20	7.28	7.40	8.20
NT	Altitude	m a.s.l.	41	1502	1863	1962	1969	2111	2309
PYR	Altitude	m a.s.l.	26	1615	2244	2374	2326	2458	2581

The number of components is an essential ingredient of PLS. A good choice is the number that minimizes the prediction error as estimated by cross-validation methods (Wold *et al.* 1984; Martens & Naes 1989). Here predicted value for observations that were not used in the computations for fitting the current model are used to compare the relative performance of different numbers of PLS components. The root mean square error of prediction (RMSEP) can be used to provide realistic estimates of relative performance and may be estimated by bootstrapping (Birks *et al.* 1990) or "leave-one-out" (ter Braak & Juggins 1993). We employ the latter technique in this paper, and use it to assess both the number of PLS components and the number of species groups. The "leave-one-out" technique thus allows us to quantify any improvement of the new multigroup approach over and above the best procedures currently available.

2.6. Data and data manipulation

The power of the transfer function method is very dependent upon the quality of the calibration dataset. The data used in the project had been well checked and harmonised before being converted into our multigroup calibration dataset.

Much of the raw data is taken from data bases drawn up as part of major collaborative ventures such as the internationally based MOLAR, CHILL-10000 and EMERGE EU-projects. Details of the Pyrenean assemblage and environmental data bases can be found in Pla *et al.* (2003) and Pla & Catalan (2005). Details of the Niedere Tauern data are set out in Kamenik *et al.* (2001), Schmidt *et al.* (2004a, b), and Kamenik & Schmidt (2005b).

The data were examined for potential outliers, because these can strongly affect transfer function coefficients and may markedly decrease the predictive ability of inference models. Outliers were identified as samples having an absolute residual (observed - predicted) higher than the SD of the environmental variable of interest and a low influence on the model indicated by Cook's D (Cook's D $< 4/n$), or as lakes with particularly unusual features (e.g., unusual water chemistry). For example, one outlier lake with very unusual water chemistry was removed from the Pyrenean group of 83 lakes to provide an 82-lake training set.

The standard weighted averaging procedure was used to estimate the optima of the dominant taxa (those that occurred in at least 2% relative abundance, in at least two lakes) for each species group, and then a multigroup inference model developed at the calibration step.

Selection of lakes within a training set can be a contentious issue. As training sets expand over time and as our focus is on testing the new multigroup approach we do not use precisely the same lake, and species, selection criteria as used in the original publication. Instead we use the maximum number of lakes available to us, while rigorously maintaining internal consistency, taxonomic quality, and by invoking the same restrained, formalised outlier and taxon removal scheme in the two geographical regions. We reemphasise that the aim of the present study is to test our multigroup approach NOT to find the training set, or sub-set, with the lowest possible prediction errors; consequently we purposely avoid protracted data-massaging and ad-hoc training-set optimization.

3. RESULTS

The range and basic statistics of each environmental variable used, in both of our geographic regions, are set out in table 1. The results of applying the weighted averaging procedures are summarised in tables 2 to 4 where the reconstruction skills of the various combinations of taxonomic groupings are compared. Use of a more complex model (involving two or three species-groups) as opposed to a simpler model (involving just one species-group) inevitably allows a better fit (in terms of residual sum of squares). However, the critical issue is not how well the model provides a description of the data, in terms of goodness of fit, but its parsimony in terms of modelling skill when using out-of-sample data. Here cross-validation provides a satisfactory measure of predictive skill. And so in all cases we report the root mean square error of prediction (RMSEP) as obtained using leave-one-out-at-a-time cross-validation. And so now, as expected, root mean square error of prediction both decreases and increases with model complexity and applicability.

Although the main statistic we selected as being the most informative when testing the multigroup approach

is RMSEP it is worth mentioning at the outset that all other standard parameters, such as r^2 (the proportion of the variability that is accounted for by the statistical model) were exactly as expected. For example a high r^2 (0.73) of the observed vs. predicted values (when cross-validated) was found for the diatom-pH association, while a low r^2 (0.001) was obtained for the diatom-CO₂ relationship, in our Pyrenean, 82-lake training-set.

While our prime interest is in whether our new methodology produces any improvement when using multiple groups, it is worth first looking at the results for individual groups. In general, our results are in good agreement with those from previous studies. In the Pyrenees diatoms produce the lowest RMSEP values for pH, alkalinity and dissolved inorganic carbon, DIC, while chrysophytes provide better predictive power for altitude (Tab. 2). In the Alps chrysophytes provide better predictive power for both pH and altitude (Tab. 3). Overall we found that the addition of a second PLS component improves the RMSEP in 60% of our 25 tested sets.

Tab. 2. Pyrenean training-set (82 lakes, 2 taxonomic groups). Units: as table 1; RMSEP: Root mean square error of prediction after cross validation; Bold: lowest RMSEP for each subgroup.

Response Group	Classic WA RMSEP	WA PLS-2 RMSEP
pH-chrysophyte	0.393	0.384
pH-diatom	0.268	0.257
pH-multigroup	0.269	0.265
alkalinity-chrysophyte	72.0	72.2
alkalinity-diatom	64.2	61.0
alkalinity-multigroup	63.2	58.2
DIC-chrysophyte	84.7	87.9
DIC-diatom	78.8	78.5
DIC-multigroup	77.8	75.5
CO ₂ -chrysophyte	37.9	36.8
CO ₂ -diatom	35.3	34.3
CO ₂ -multigroup	36.3	35.3
Alt-chrysophyte	148	141
Alt-diatom	165	160
Alt-multigroup	140	131

We turn now to the question of improvement over and above the single-group results when using multiple groups. The biggest improvements in predictive ability are found in the Pyrenees for the diatom-chrysophyte-altitude association (Tab. 2), and also for the diatom-chrysophyte-chironomid altitude association (Tab. 4). Here RMSEP is reduced from 141 m to 131 m, and from 132 m to 120 m respectively. Small improvements are also found for the Pyrenean diatom-chrysophyte-alkalinity and diatom-chrysophyte-DIC associations (Tab. 2). The RMSEP comes down from 61.0 to 58.2 $\mu\text{eq L}^{-1}$ and from 78.5 to 75.5 $\mu\text{eq L}^{-1}$ respectively. Addition of a third PLS component decreased the skill in all cases.

Tab. 3. Niedere-Tauern training-set (41 lakes, 2 taxonomic groups). Units: as table 1; RMSEP: Root mean square error of prediction after cross validation; Bold: lowest RMSEP for each subgroup.

Response Group	Classic WA RMSEP	WA PLS-2 RMSEP
pH-chrysophyte	0.201	0.193
pH-diatom	0.253	0.270
pH-multigroup	0.205	0.207
Alt-chrysophyte	193	196
Alt-diatom	198	213
Alt-multigroup	191	192

Tab. 4. Pyrenean training-set (26 lakes, 3 taxonomic groups). Units: as table 1; RMSEP: Root mean square error of prediction after cross validation; Bold: lowest RMSEP for each subgroup.

Response Group	Classic WA RMSEP	WA PLS-2 RMSEP
Alt-chrysophyte	135	144
Alt-diatom	132	140
Alt-chironomid	149	152
Alt-multigroup	122	120

Turning to the multigroup results for the Alps (Tab. 3) we find no improvement in RMSEP for pH reconstruction and only a very minor improvement from, 193 to 191 m, for altitude (which cannot be regarded as significant).

In summary only minor improvements in cross validation skill are found for certain of our environmental parameters. Indeed for a more complex PLS inference model to be truly useful it should lead to a significant drop in RMSEP. Assessing what is a significant drop is not easy. Birks (1998) uses a rule of thumb that a "useful" additional component is one that gives a reduction in prediction error of 5% or more in a PLS or WA-PLS model. Others suggest a permutation testing approach (e.g., van der Voet 1994). Given the dangers of overfitting PLS models we do not consider the small improvements in WAPLS component seen here, which amount to 9% (for the Pyrenean 82-lake and 26-lake altitude reconstructions), but on average only constitute 3% to be significant. Thus our test failed to discover any overall major improvement in transfer-function performance when using our new multigroup approach.

4. DISCUSSION

The new computer code for our weighted averaging PLS procedure was checked by running the code on the well-known 167 sample SWAP diatom-pH data-set from lowland and upland lakes (Birks *et al.* 1990) as analysed by Juggins (ter Braak & Juggins 1993). Our code reproduces RMSEP to three significant figures (including all the changes in RMSEP as the number of PLS components is increased).

Previous studies in the Pyrenees, and in the Niedere Tauern, being based on single taxonomic groups were able to make use of a different selection of lakes and so

were able to achieve slightly better RMSEP values (Pla & Catalan 2005; Schmidt *et al.* 2004a) than obtained here. The precise choice of lakes and taxa included in the analysis can thus be seen to be of great importance. In this study we were especially careful to always use exactly the same lakes and taxa in our comparison of the multi- and single group methodologies, and to keep our methods of outlier detection and data-handling constant.

If significantly improved reconstruction skill could have been demonstrated in any of the training sets then it would have been worth trying the method out on stratigraphic fossil-assemblages in order to make improved reconstructions of past trends in water quality or climate.

What might be the reason why our multigroup analysis produced only minor improvements in environmental reconstructions? Our combination of taxonomic groupings puts together organisms with different life-forms and thus environmental responses, reaction times or sensitivities. In fact, the optimum estimation of species distributions throughout an environmental gradient is rooted in the assumption that phylogenetically related species segregate along it as a result of interspecific competition. The less phylogenetically related the species, the lower the competitive interaction. Therefore, while the appearance of one species interferes with the appearance of another species, within the same taxonomic group, in the same location, or niche, it has a more limited or null influence on the appearance of species of other taxonomic groups. In addition the way species segregate in the environment usually follow distinct factors for each principal taxonomic group. For example, it is well known that diatom species distribution is extremely sensitive to water pH; however, chironomids, although showing some response to pH for extreme cases, mainly follow other factors such as temperature, oxygen and substrate. Our hope had been that in such circumstances the use of multigroup assemblage data would allow any interaction, or confounding, effects which were obscuring the environmental reconstruction to be reduced. Perhaps for the environmental factors that we chose to investigate only a few taxa dominated the environmental gradient and so largely exceeded all the others in their predictive capacity. As a consequence the sensitivity of the weighted averaging to the environmental variables was not able to be improved upon even by the tools of our more elaborated statistics. As a result our multigroup approach was unable to perform much better than current practice allows.

On the other hand perhaps our chosen groups failed to satisfy the prerequisite that sufficient species within the groups were reacting significantly to the same common response variable. For example for the training sets from the Central Eastern Alps Kamenik & Schmidt (2005a) and Schmidt *et al.* (2004b) have previously noted that the different groups do not respond to exactly the same parameter. They found that chrysophytes

respond strongly to the onset of spring mixing whereas, by contrast, diatoms show a strong response to the onset of autumn mixing. Thus both significantly react to temperature change but on different time scales and in different seasons.

Alternatively the problem could be that the main taxonomic groups segregate (or split) the species following similar environmental partitions. Consequently there is little interaction among the species of the main taxonomic groups. So once again our mathematical manipulations were only able to generate slight improvements in reconstruction skill.

A third possibility is that the problem, rather than lying on the right hand side of the calibration equation, as in our two suggestions above, lies on the left hand side; that is the problem could lie with the measurement of the environmental variables at the training sites. ter Braak (1987) set out this suggestion particularly cogently when he noted "*It might be thought easier to measure environmental variables at a site than to infer their values from the species that occur there. But often it is not.*"

How might transfer function work be improved? If the problem with the multigroup work relates to different reaction times then one way to improve the prediction power of transfer functions might be to link environmental data directly with the time of accumulation, or blooming, or lifespan of organisms. The strong prediction power of planktonic organisms, such as chrysophyte cysts (Pla & Catalan 2005; Kamenik & Schmidt 2005b), with shorter reaction times than benthic organisms may be a pointer to this direction (cf. Kamenik *et al.* (2005a) discussion of the reaction times of benthic diatoms *versus* those of planktonic chrysophytes). If this is the case then greater use of sediment traps and data loggers (e.g., Kamenik & Schmidt 2005b) may suggest a practical way forwards for environmental factors with a large seasonal fluctuation, such as water temperature. Finally a different mathematical way forwards might be to try to adapt the co-correspondence analysis approach (ter Braak & Schaffers 2004) to ordination, which relates one environmental data matrix to another data matrix, to address the problem of multigroup inference.

ACKNOWLEDGEMENTS

The programming work was carried out when RT was on sabbatical leave in Blanes and Mondsee. Financial assistance for the visits, from Leverhulme and The Royal Society of Edinburgh, is gratefully acknowledged.

REFERENCES

- Agusti-Panareda, A. & R. Thompson. 2002. Reconstructing air temperature at eleven remote alpine and arctic lakes in Europe from 1781 to 1997 AD. *J. Paleolimnol.*, 28: 7-23.
- Battarbee, R.W. 1986. Diatom analysis. In: Berglund BE (Ed.), *Handbook of Holocene Palaeoecology and Palaeohydrology*. Wiley, Chichester: 527-570.

- Birks, H.H. & H.J.B. Birks. 2006. Multi-proxy studies in palaeolimnology. *Vegetation History and Archaeobotany*, 15: 235-251.
- Birks, H.J.B. 1995. Quantitative palaeoenvironmental reconstructions. In: Maddy D., Brew J.S. (Eds), *Statistical modelling of Quaternary science data, Technical Guide 5*. Quaternary Research Association, Cambridge: 161-254.
- Birks, H.J.B. 1998. Numerical tools in paleolimnology - progress, potentialities, and problems. *J. Paleolimnol.*, 20: 307-332.
- Birks, H.J.B., S. Juggins & J.M. Line. 1990. Lake surface-water chemistry reconstructions from palaeolimnological data. In: Mason B.J. (Ed.), *The Surface Waters Acidification Programme*. Cambridge University Press, Cambridge: 301-313.
- Broecker, W.S. 2001. Was the Medieval Warm Period Global? *Science*, 291: 1497-1499.
- Cameron, N.G., H.J.B. Birks, V.J. Jones, F. Berge, J. Catalan, R.J. Flower, J. Garcia, B. Kawecka, K. A. Koinig, A. Marchetto, P. Sánchez-Castillo, R. Schmidt, M. Sisko, E. Solovieva, E. Stefkova & M. Toro. 1999. Surface - sediment and epilithic diatom pH calibration sets for remote European mountain lakes (AL:PE project) and their comparison with the Surface Waters Acidification Programme (SWAP) calibration set. *J. Paleolimnol.*, 22: 291-317.
- Catalán, J., E. Ballesteros, E. Gacia, A. Palau & L. Camarero. 1993. Chemical composition of disturbed and undisturbed high-mountain lakes in the Pyrenees: a reference for acidified sites. *Wat. Res.*, 27: 133-141.
- Charles, D.F. & J.P. Smol. 1988. New methods for using diatoms and chrysophytes in reconstructions of lakewater pH in low alkalinity lakes. *Limnol. Oceanogr.*, 33: 1451-1462.
- Charles, D.F. & J.P. Smol. 1994. Long-term chemical changes in lakes: Quantitative inferences from biotic remains in the sediment record. In Baker L. (Ed.), *Environmental Chemistry of Lakes and Reservoirs. Advances in Chemistry Series 237*. American Chemical Society, Washington DC: 3-31.
- Facher, E. & R. Schmidt. 1996. A siliceous chrysophycean cyst-based pH transfer function for Central European lakes. *J. Palaeolimnol.*, 16: 275-321.
- Imbrie, J. & N.G. Kipp. 1971. *A new micropaleontological method for paleoclimatology: Application to a Late Pleistocene Caribbean core. The Late Cenozoic Glacial Ages*. New Haven, Yale University Press: 71-181.
- Kamenik, C. & R. Schmidt. 2005a. Computer-aided SEM analysis of Chrysophyte stomatocysts. *Nova Hedwigia, Beiheft*, 128: 269-274.
- Kamenik, C. & R. Schmidt. 2005b. Chrysophyte resting stages: a tool for reconstructing winter/spring climate from alpine lake sediments. *Boreas*, 34: 477-489.
- Kamenik, C., K.A. Koinig & R. Schmidt. 2005a. Potential effects of pre-industrial lead pollution on algal assemblages from an Alpine lake. *Ver. int. Ver. Limnol.*, 29: 535-538.
- Kamenik, C., R. Schmidt, G. Kum & R. Psenner. 2001. The Influence of Catchment Characteristics on the Water Chemistry of Mountain Lakes. *Arctic Antarctic and Alpine Research*, 33: 404-409.
- Kamenik, C., A. Agustí-Panareda, P.G. Appleby, J.A. Dearing, E.M. Shilland, F. Sporka, E. Stefkova & R. Thompson. 2005b. Palaeolimnological evidence for atmospheric pollution, climate and catchment-related changes in alpine Chrysophyte stomatocyst assemblages (Tatra, Slovakia). *Nova Hedwigia, Beiheft*, 128: 275-293.
- Korhola, A., K. Vasko, H.T.T. Toivonen & H. Olander. 2002. Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modeling. *Quaternary Science Reviews*, 21: 1841-1860.
- Kucera, M., A. Rosell-Melè, R. Schneider, C. Waelbroeck & M. Weinelt. 2005. Multiproxy approach for the reconstruction of the glacial ocean surface (MARGO). *Quaternary Science Reviews*, 24: 813-819.
- Marchetto, A. & A. Lami. 1994. Reconstruction of pH by chrysophycean scales in some lakes of the Southern Alps. *Hydrobiologia*, 274: 83-90.
- Marchetto, A. & R. Schmidt. 1992. A regional calibration data set to infer lakewater pH from sediment diatom assemblages in alpine lakes. *Mem. Ist. ital. Idrobiol.*, 51: 115-125.
- Martens, H. & T. Næs. 1989. *Multivariate calibration*. Wiley, Chichester: 419 pp.
- Pla, S. 2001. *Chrysophycean cyst from the Pyrenees*. J. Cramer, Berlin, Germany: 179 pp.
- Pla, S., & J. Catalan. 2005. Chrysophyte cysts from lake sediments reveal the submillennial winter/spring climate variability in the northwestern Mediterranean region throughout the Holocene. *Climate Dynamics*, 24: 264-278.
- Pla, S., L. Camarero & J. Catalan. 2003. Chrysophycean cysts relationships to water chemistry in Pyrenean lakes (NE Spain) and their potential for environmental reconstruction. *J. Paleolimnol.*, 30: 21-34.
- Rieradevall, M. & S.J. Brooks. 2001. An identification guide to subfossil Tanypodinae larvae (Insecta: Diptera: Chironomidae) based on cephalic setation. *J. Paleolimnol.*, 25: 81-99.
- Rosén, P., U. Segerström, I. Renberg & L. Eriksson. 2003. Do diatom, chironomid and pollen records consistently infer Holocene July air temperature? A comparison using sediment cores from four alpine lakes in northern Sweden. *Arctic, Antarctic and Alpine Research*, 35: 279-290.
- Schmid, P.E. 1993. A key to the larval Chironomidae and their instars from Austrian Danube region streams and rivers with particular reference to a numerical taxonomic approach. Part 1. Diamesinae, Prodiamesinae and Orthoclaudiinae. *Wasser und Abwasser*, Supplement 3: 1-514.
- Schmidt, R., C. Kamenik, C. Kaiblinger & M. Hetzel. 2004a. Tracking Holocene environmental changes in an alpine lake sediment core: application of regional diatom calibration, geochemistry, and pollen. *J. Paleolimnol.*, 32: 177-196.
- Schmidt, R., C. Kamenik, H. Lange-Bertalot & R. Klee. 2004b. *Fragilaria* and *Staurosira* taxa (Bacillariophyceae) from surface sediments of 40 lakes in the central Austrian Alps (Niedere Tauern) in relation to environmental variables. *J. Paleolimnol.*, 63: 171-189.
- Seppä, H. & H.J.B. Birks. 2001. July mean temperature and annual precipitation trends during the Holocene in the Fennoscandian tree-line area: pollen-based climate reconstructions. *The Holocene*, 11: 527-539.
- Smol, J.P. & B.F. Cumming. 2000. Tracking long-term changes in climate using algal indicators in lake sediments. *J. Phycol.*, 36: 986-1011.
- Stevenson, A.C., S. Juggins, H.J.B. Birks, D.S. Anderson, N.J. Anderson, R.W. Battarbee, F. Berge, R.B. Davis, R.J. Flower, E.Y. Haworth, V.J. Jones, J.C. Kingston, A.M. Kreiser, J.M. Line, M.A.R. Munro & I. Renberg. 1991. *The surface waters acidification project palaeolimnology programme: modern diatom/lake water chemistry dataset*. ENSIS Publishing Ltd., London.
- ter Braak, C.J.F. 1987. Calibration. In: Jongman R.H.G., ter Braak C.J.F., van Tongeren O.F.R. (Eds), *Data analysis in community and landscape ecology*. Pudoc, Wageningen: 78 pp.
- ter Braak, C.J.F. & S. Juggins. 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia*, 269/270: 485-502.
- ter Braak, C.J.F. & I.C. Prentice. 1988. A theory of Gradient Analysis. *Advances in Ecological Research*, 18: 272-317.

- ter Braak, C.J.F. & A.P. Schaffers. 2004. Co-Correspondence Analysis: a new ordination method to relate two community compositions. *Ecology*, 85(3), 834–846.
- Thompson, R., D. Price, N. Cameron, V. Jones, C. Bigler, P. Rosen, R.I. Hall, J. Catalan, J. Garcia, J. Weckstrom & A. Korhola. 2005. Quantitative Calibration of Remote Mountain-lake Sediments as Climatic Recorders of Air Temperature and Ice-cover duration. *Arctic, Antarctic, and Alpine Research*, 37: 626–635.
- van der Voet, H. 1994. Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems*, 25: 313–323.
- Wiederholm, T. 1983. *Chironomidae of the Holarctic region. Keys and diagnoses. Part 1: larvae*. Entomologica Scandinavica Supplement: 457 pp.
- Wold, S., A. Ruhe, H. Wold & W.J. Dunn. 1984. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, 5: 735–743.

Received: June 2008

Accepted: August 2008