The use of diatom records to establish reference conditions for UK lakes subject to eutrophication

Authors: [1]Helen Bennion, [1]Gavin L. Simpson

Addresses, emails:

[1] Department of Geography, Environmental Change Research Centre, University College London, Gower Street, London WC1E 6BT, U.K.; hbennion@geog.ucl.ac.uk, gavin.simpson@ucl.ac.uk

Abstract

A knowledge of pre-disturbance conditions is important for setting realistic restoration targets for lakes. For European waters this is now a requirement of the European Council Water Framework Directive (WFD) where ecological status must be assessed based on the degree to which present day conditions deviate from reference conditions. Here, we employ palaeolimnological techniques, principally inferences of total phosphorus from diatom assemblages (DI-TP) and classification of diatom composition data from the time slice in sediment cores dated to ~1850 AD, to define chemical and ecological reference conditions, respectively, for a range of UK lake types. The DI-TP results from 169 sites indicate that reference TP values for low alkalinity lakes are typically < 10 µg L^{-1} and in many cases < 5 µg L^{-1}, whilst those for medium and high alkalinity lakes are in the range 10-30 µg L^{-1} and 20-40 µg L^{-1}, respectively. Within the latter two alkalinity types, the deeper waters (> 3 m mean depth) generally had lower reference TP concentrations than the shallow sites. A small group of shallow marl lakes had concentrations of ~30 µg L^{-1}. Cluster analysis of diatom composition data from 106 lakes where the key pressure of interest was eutrophication identified three clusters, each associated with particular lake types, suggesting that the typology has ecological relevance, although poor cross matching of the diatom groups and the lake typology at type boundaries highlights the value of a site-specific approach to defining reference conditions. Finally the floristic difference between the reference and present day (surface sample) diatom assemblages of each site was estimated using the squared chord distance dissimilarity coefficient. Only 25 of the 106 lakes experienced insignificant change and the findings indicate that eutrophication has impacted all lake types with > 50% of sites exhibiting significant floristic change. The study illustrates the role of the sediment record in determining both chemical and ecological reference conditions, and assessing deviation from the latter. Whilst restoration targets may require modification in the future to account for climate induced alterations, the long temporal perspective offered by palaeolimnology ensures that such changes are assessed against a sound baseline.

Introduction

Eutrophication of aquatic ecosystems, principally caused by phosphorus (P) and nitrogen (N) from agricultural sources, industrial waste and domestic sewage, has been recognised as a global problem since the 1960s (Hutchinson 1969) and remains one of the foremost environmental issues threatening the quality of surface waters (Smith et al. 2006). The symptoms include high algal biomass often accompanied by dense summer blooms of cyanobacteria or green algae, reduced abundance or complete loss of submerged plants, dominance of plankti-benthivorous fish, low water clarity and deoxygenation (Moss 1998). Many lakes have been subject to nutrient loading over relatively long timescales (centuries, decades) (e.g. Bradshaw et al. 2006) and it is therefore likely that few sites in European lowland catchments remain in a pristine condition. Efforts to better manage and restore enriched systems have increased over the last decade or so, ranging from reductions in nutrient loading (e.g. Jeppesen et al. 2007) to biomanipulation (e.g. Søndergaard et al. 2007) and there are now numerous examples of lakes in recovery (Anderson et al. 2005). Nevertheless, whilst control of point sources of pollution, particularly from sewage treatment works, has improved markedly, there are increasing concerns about how to tackle the more challenging problem of diffuse nutrient sources and slow release of P from soils in agricultural regions (Carpenter 2005). Furthermore, we must now consider whether changes in our climate may exacerbate the symptoms of eutrophication and confound recovery efforts (Battarbee et al. 2005; Jeppesen et al. 2007).

Sound management requires a good understanding of ecosystem processes and a knowledge of pre-disturbance conditions to set realistic restoration targets (e.g. Moss et al. 1997; Battarbee 1999). The advent of the European Council Water Framework Directive (WFD) (European Union 2000) in 2000 has made it a requirement for all EU member states to classify their surface waters based on the degree to which present day conditions deviate from those expected in the absence of significant anthropogenic influence, termed reference conditions, and therefore the need for information on the chemical and ecological status of aquatic systems prior to enrichment is greater than ever. Unfortunately long-term datasets are rare for most lakes and, where they exist, monitoring programmes tend to have been initiated once the problem has already become apparent. The WFD states that, in the absence of long-term data, reference conditions may be derived using palaeolimnological methods (Pollard and Huxham 1998; European Union 2000). Indeed, palaeolimnology has now become a standard technique for providing information on ecological status and the timing, rate and magnitude of ecological change (Smol 2008) and has been applied in several recent studies to define reference conditions in the context of the WFD (see review in Bennion and Battarbee 2007).

One of the most widely used biological groups in such studies is the diatoms (Bacillariophyceae), unicellular, siliceous algae (e.g. Stoermer and Smol 1999; Battarbee et al.

2001a). Diatoms are sensitive to water quality and are good indicators of lake trophic status (Hall and Smol 1999). They have been employed both qualitatively, whereby shifts in diatom assemblages over time coupled with ecological information can provide a record of ecological change, and quantitatively, whereby transfer functions have been developed to model the relationship between diatom assemblage composition and water chemistry in a training set of lakes (Birks 1998). With respect to eutrophication, diatom-P transfer functions have been the most commonly developed models for inferring past nutrient concentrations in lakes (Hall and Smol 1999). In most cases, the models reconstruct total P (TP) concentrations as this is the component that is routinely measured in water chemistry surveys given that it is typically the limiting nutrient in fresh waters (Vollenweider 1969), it is traditionally used in lake trophic status classifications (OECD 1982) and it is consequently the nutrient parameter most familiar to lake managers. Numerous diatom-P transfer functions have been developed for European lakes (e.g. Wunsam and Schmidt 1995; Bennion et al. 1996a; Lotter et al. 1998; Bradshaw and Anderson 2001; Kauppila et al. 2002; Chen et al. 2008), several of which were combined in the EU project EDDI ("European Diatom Database") (Battarbee et al. 2001b). The approach is able to provide estimates of baseline (reference) TP concentrations in lakes and, when coupled with dating of sediment cores, the timing, rates and possible causes of enrichment at a particular site (e.g. Bennion et al. 1996b; Anderson 1997; Taylor et al. 2006).

In contrast to many of the methods commonly used to define reference conditions for lakes such as the loss coefficient method (e.g. Moss et al. 1997) and the morphoedaphic index (Vighi and Chiaudani 1985), palaeolimnological techniques can be employed not only to provide a chemical reference condition but to provide an ecological target for management purposes, a concept fundamental to the WFD (European Union 2000). The WFD places emphasis on the ecological structure and function of aquatic ecosystems with biological elements (fish, invertebrates, macrophytes, phytobenthos and phytoplankton) at the centre of the status assessments, and hydromorphology and physico-chemistry as supporting elements. The fossil assemblages in the reference samples of cores can potentially be used to describe the pre-impact communities, i.e. the ecological reference condition, of a particular lake or lake type.

In this paper we evaluate the role of the sedimentary diatom record for defining reference conditions and assessing the degree of ecological change with respect to eutrophication for a set of UK lakes. The lakes have been classified into types based on a scheme designed for WFD implementation which enables the environment agencies in England, Wales and Scotland to report reference conditions on a type-specific basis. We apply diatom transfer functions to establish reference TP concentrations and classify the compositional data to characterise the reference diatom communities for each lake type. Dissimilarity measures are calculated between the reference and surface sediment diatom assemblages to assess the degree of floristic change at each site.

4

Methods

The datasets

A dataset of 169 UK lakes was used for establishing chemical reference conditions. This was derived from existing data held in the database, AMPHORA, at the Environmental Change Research Centre (ECRC), University College London. Data were retrieved for all sites where diatoms had been analysed from a sediment core collected from a UK lake. Each site was assigned to a type according to the lakes reporting typology developed for ecoregion 18 (Great Britain-GB) (Phillips 2003). This was designed to meet the needs of river basin characterisation required by the WFD whereby lakes are classified into types such that type-specific reference conditions can be described for the ecological quality elements. The GB lake typology scheme is based on alkalinity (catchment geology) and lake mean depth data and the 169 lakes represent a broad range of the main lake types with good geographical coverage (Table 1). However, because the diatom data were derived from projects undertaken by the ECRC over the last few decades, many of which have focused on areas sensitive to acidification, there was a strong bias towards low alkalinity and upland systems (Table 1). The numbers of medium and high alkalinity lakes in the diatom database were relatively few. Given that the current paper is concerned with lakes subject to eutrophication, for the definition of ecological reference conditions and assessment of floristic change, a slightly modified version of the dataset was used. This included samples from additional medium or high alkalinity waters in predominantly low lying areas of the UK that were added to the database after 2004. It also screened out all those lakes in acid sensitive regions where the main driver of ecological change was unlikely to be eutrophication. This resulted in a smaller but more appropriate dataset of 106 lakes with more equal representation of the three main alkalinity classes (Table 1, Appendix 1).

A sediment core had been collected from the deepest point of each lake during the original studies which were conducted at some time between the late 1980s and 2003. In most cases the cores were dated, using either radiometric techniques (Appleby et al. 1986) or the spheroidal carbonaceous particle (SCP) profile method (Rose et al. 1995). The sample approximating to ~1850 AD was selected to represent the reference condition, herein termed the reference sample. In the absence of dates (at ~40 sites), the bottom sample of the core was taken to represent the reference sample although it is possible, particularly where cores are short, that this does not extend back as far as 150 years. For the UK, it is generally agreed that ~1850 AD is a suitable date against which to assess impacts for lakes as this represents a period prior to major industrialisation and agricultural intensification (Battarbee 1999; Fozzard et al. 1999). The analysis of reference samples in this way removes the problem inherent in spatial-state schemes in which the lakes have been subject to different pressures and varying degrees of impact (Cardoso et al. 2007). It is accepted, however, that aquatic systems have been subjected to anthropogenic

impacts over much longer time-scales and, therefore, the ~1850 reference conditions are unlikely to equate to the natural or pristine state. The diatoms in the reference samples were counted using standard techniques (Battarbee et al. 2001a) and in all cases a minimum of 300 valves were identified. All diatom data were expressed as percentage relative abundance.

Chemical reference conditions

The technique of weighted averaging (WA) regression and calibration (e.g. ter Braak and van Dam 1989), and its extension to WA partial least squares (WAPLS) (ter Braak and Juggins 1993), have become standard techniques in palaeolimnology for reconstructing past environmental variables (Anderson 1995; Battarbee 1999). Existing diatom transfer functions, derived using these methods, for inferring TP were applied to the diatom data in the reference samples from the 169 lakes, following taxonomic harmonisation between the training sets and the fossil data. Reconstructions of reference DI-TP were produced using either a Northwest European training set of 152 relatively shallow lakes (< 10 m maximum depth) with a median value for the dataset of 104 µg TP L^{-1} and a root mean squared error of prediction (RMSEP) of 0.21 $\log_{10}$ µg TP L^{-1} for the WAPLS two-component (WA-PLS2) model (Bennion et al. 1996a), or a model based on a training set of 56 relatively large, deep lakes (> 10 m maximum depth) from Scotland, Northern Ireland, Cumbria, southern Norway and central Europe, with a median value for the dataset of 22 µg TP L^{-1} (Bennion et al. 2004). For this dataset, the best model was generated with simple WA and inverse deshrinking (RMSEP of 0.25 $\log_{10}$ µg TP L^{-1}). For each reference sample, the training set containing the greatest percentage of the taxa present in the fossil sample was selected, which in most cases resulted in the shallow lake training set being used for the shallow lakes and the deep lake model being used for the deep lakes. All reconstructions were implemented using C^2 (Juggins 2003).

Ecological reference conditions

The smaller dataset of reference samples from 106 UK lakes was used to identify the diatom communities typically associated with reference condition of the main lake types. All 274 taxa present in the samples were included and a Hellinger transformation (Legendre and Gallagher 2001) was applied to the diatom data prior to analysis. k-means cluster analysis was used to partition the diatom data into k = 2, ..., 10 clusters. One hundred random starting locations for the k-means algorithm were generated and the best solution for each k was retained for further analysis. A combination of the Calinski-Harabasz criterion and ordination was used to determine the number of clusters to retain and interpret. The results of the k-means clustering were projected into a two dimensional solution from a non-metric multi-dimensional scaling (nMDS) analysis of the Hellinger pair-wise dissimilarities between sites, following the recommendations of Minchin (1987).

nMDS aims to provide an optimal, low-dimensional mapping of the original dissimilarities by preserving only the rank ordering of pair-wise, between site distances. The indicator species associated with each cluster were identified using the IndVal procedure of Dufrêne and Legendre (1997). The indicator species retained were those taxa which had a significant IndVal score, assessed using 1000 random permutations.

In order to assess how closely the diatom based clusters matched the GB lake typology, the typology classes were converted to dummy (presence-absence) class variables. Each lake was assigned a value of 1 for the dummy class variable representing the GB type the lake belonged to and a 0 for the other types. IndVal was then performed using the diatom-based cluster membership as the grouping variable and the dummy class variables as the 'species' data. As above, 1000 permutations were used to assess the significance of the calculated IndVal scores. The IndVal method combines measures of *fidelity* and *specificity* such that good indicator species are those species that are found predominantly in a single group and are present in the majority of samples within that group, and are not necessarily the most abundant taxa. IndVal can be extended to the use of presence-absence data as described above (Dufrêne and Legendre 1997).

All analyses were performed using R version 2.7.1 (R Core Development Team 2008) using the vegan (version 1.11-5; Oksanen et al 2008) and labdsv (version 1.3-1, Roberts 2007) R packages.

Floristic change

The diatom data from the surface (uppermost 0.5 or 1 cm) sample of each of the 106 lake cores were used to provide information on the current diatom assemblages of the lakes. The surface sediment sample represents the last few years' accumulation of diatoms deposited from a variety of habitats within the lake. However, because this study makes use of existing data from previous studies, some of the core tops represent conditions from up to a decade ago. The diatoms were counted using standard techniques (Battarbee et al. 2001a) in the same way as for the reference samples. The degree of floristic change between the reference and surface sample was assessed using the squared chord distance (SCD) dissimilarity coefficient (Overpeck et al. 1985) computed using R (R Development Core Team 2008) and the analogue package version 0.5-2 (Simpson 2007a, b). SCD values range from 0 to 2 where a value of 0 indicates that two samples are perfectly similar and a value of 2 perfectly dissimilar. Scores less than 0.29, 0.39, 0.48 and 0.58 indicate insignificant floristic change at the 1st, 2.5th, 5th and 10th percentile, respectively (Simpson et al. 2005). The 2.5th percentile is used here to define sites with low floristic change between the reference and surface sample. This is more stringent than the 5th percentile used by

Bennion et al. (2004) and reflects revised thinking by the authors about what constitutes biologically important change at a site based largely on the observation that unimpacted lakes in the AMPHORA database (including lakes analysed for evidence of acidification and not included in this study) typically have SCD scores of <0.4. Where insignificant change is detected, sites can be considered as potential reference lakes and where significant change is observed, the degree of change from site-specific reference conditions is evaluated to assess ecological status.

Results

Chemical reference conditions

The DI-TP reference values for the 169 lakes range from 2 to 50 µg l^{-1}. Typical ranges of reference TP concentrations were determined for the main lake types (Fig. 1). The low alkalinity lakes, both deep and shallow, had reference TP values of < 10 µg L^{-1} and in many cases of < 5 µg L^{-1} reflecting the naturally nutrient-poor status of these waters. The medium alkalinity, deep lakes typically had reference TP concentrations of 10-20 µg L^{-1}, although some lakes had values < 10 µg L^{-1} and some were > 30 µg L^{-1} reflecting the more productive status of many of these waters even prior to the major period of enrichment. The medium alkalinity, shallow lakes had higher values than the deep lakes in this alkalinity band with reference TP concentrations typically in the range 20-30 µg L^{-1}, and the values for the high alkalinity, deep lakes were also in this range. The high alkalinity, shallow lakes had somewhat higher reference TP concentrations than the other lake types, typically of 30-40 µg L^{-1}, reflecting the naturally higher productivity of these systems, although there was a large range of values in this group with a maximum of ~50 µg L^{-1}. Finally, a small group of shallow marl lakes had concentrations of ~30 µg L^{-1}.

Ecological reference conditions

Following removal of five outliers from the full dataset of 106 lakes owing to their unusual species composition, three clusters were identified in the remaining dataset of 101 lakes (Fig. 2a) along with their associated indicator species (Table 2). The outliers included both of the low alkalinity, shallow lakes and therefore this type was omitted from the cluster analysis. Cluster 1 (N=33) was comprised of high alkalinity and marl lakes, most of which were shallow but with a small number of deep lakes also present. The indicator taxa associated with this cluster included the benthic *Fragilaria* species (e.g. *F. construens* var. *venter, F. construens, F. pinnata, F brevistriata*), *Amphora pediculus* and several planktonic taxa including *Stephanodiscus parvus, Stephanodiscus hantzschii* and *Aulacoseira granulata* (Fig. 2b). Cluster 2 (N=36) was comprised of mostly medium alkalinity and several high alkalinity lakes, including both shallow and deep waters. The assemblages of these lakes were typically comprised of *Cyclotella radiosa*, *Cymbella*

microcephala, *Fragilaria virescens* var *exigua*, *Eunotia incisa*, *Navicula radiosa* and *Achnanthes pusilla* (Fig. 2b). Finally, cluster 3 (N=32) was comprised of low to medium alkalinity lakes and was largely restricted to deep waters. The indicator taxa associated with these lake types included planktonic *Cyclotella* taxa (e.g. *C. comensis, C. kuetzingiana*), *Brachysira vitrea, Tabellaria flocculosa, Achnanthes minutissima* and *Hannaea arcus* (Fig. 2b).

The IndVal analysis showed that the diatom clusters map on reasonably well to particular lake types (Table 3). The LA-D and MA-D lakes were significantly associated with cluster 3, MA-Sh lakes were significantly associated with cluster 2, and HA-Sh and Marl-Sh lake types, and to a lesser extent Marl-D types, were significantly associated with cluster 1. Only the HA-D group had an insignificant score being equally distributed between clusters 1 and 2. There were no representatives of the LA-Sh lake type in the cluster analyses.

Floristic change

The SCD scores calculated between the reference and surface sediment samples of the 106 lakes range from 0.13 to 1.77 (Fig. 3), with only 25 lakes experiencing low floristic change (<0.39= 2.5th percentile) (Table 4). Of the 21 low alkalinity, deep lakes, seven have a score below the critical value at the 2.5th percentile whereas neither of the two low alkalinity, shallow lakes have a score below this value (Table 4). The remaining low alkalinity lakes have experienced significant change with nine sites exceeding the critical value at the 10th percentile (> 0.58). Of the 23 medium alkalinity, deep lakes and 13 medium alkalinity, shallow lakes, only five and three sites, respectively, have a score below the critical value at the 2.5th percentile (Table 4). A total of 22 medium alkalinity lakes exceed the critical value at the 10th percentile. Of the eight high alkalinity, deep lakes, three have a score below the critical value at the 2.5th percentile and of the 28 high alkalinity, shallow lakes, only six have a score below this value (Table 4). Many of the lakes in the high alkalinity band have experienced significant floristic change with 21 lakes exceeding the critical value at the 10th percentile. Of the 11 marl lakes, only one deep site has a score below the critical value at the 2.5th percentile (Table 4). Five of the marl lakes have undergone significant floristic change at the 10th percentile.

Discussion

Chemical reference conditions

Application of transfer functions to the fossil diatom assemblages in the 169 lakes enabled typical ranges of reference TP concentrations to be determined for the main lake types. The results indicate that reference concentrations increase with alkalinity and decrease with lake depth. The

low alkalinity lakes have the lowest DI-TP reference values in the dataset with almost all lakes in this group having concentrations < 10 µg L^{-1} and many with values < 5 µg L^{-1}. The similarity in the median and range of DI-TP values for both the deep and shallow, low alkalinity lakes suggests that lake depth has little influence on the low alkalinity systems and it is the low nutrient availability that exerts the strongest control on the diatom populations. In contrast, within the medium and high alkalinity lake types there is a clear relationship with depth such that the shallow lakes within each of these types have higher DI-TP values than their deeper counterparts. For example, in the medium alkalinity deep lakes DI-TP is mostly in the range 10-20 µg L^{-1} compared with 20-30 µg L^{-1} for the medium alkalinity shallow lakes. Likewise the values for the high alkalinity, deep lakes were largely in the range 20-30 µg L^{-1} compared to values typically of 30-40 µg L^{-1} in their shallow equivalent.

These findings agree with other studies on background nutrient conditions which have demonstrated a clear relationship between TP concentrations and alkalinity and mean depth. Indeed the strength of this relationship led Vighi and Chiaudani (1985) to construct the morphoedaphic index (MEI) which predicts background TP concentrations based on the ratio between total dissolved solids, as measured by alkalinity or conductivity, to lake mean depth. In an analysis of data from over 500 European reference lakes, Cardoso et al. (2007) observed that TP increased with alkalinity and declined with depth and a similar conclusion was drawn from the examination of over 1000 TP observations from a range of lake types across Europe (Phillips et al. 2008). Owing to the strong relationship between TP and chlorophyll a in these lakes (Phillips et al. 2008) chlorophyll a similarly varies with lake type such that reference concentrations are highest in high alkalinity, shallow lakes and are lowest in low alkalinity, deep lakes (Carvalho et al. 2008). Alkalinity reflects the fertility and base status of the catchment geology and soils and is therefore likely to be positively correlated with TP. The relationship between mean depth and TP is somewhat more complex but proportionally more nutrients are lost from deep lakes via sedimentation processes. In contrast, most shallow lakes are permanently mixed (polymictic) with a consequent lack of stratification of temperature and oxygen, and increased potential for the recycling of nutrients (Holdren and Armstrong 1980; Carrick et al. 1994). Interestingly, the small group of shallow marl lakes had concentrations of ~30 µg L^{-1}, a similar range to those seen in the medium alkalinity shallow waters. In marl systems, inorganic P is sequestrated from the water column as a result of its co-precipitation with carbonate, thus P concentrations are reduced. Hence, marl lakes have background TP concentrations more akin to medium than to the high alkalinity waters.

Reassuringly the DI-TP values for the main lake types are in the same range as those derived from the palaeo meta-database (Battarbee et al. this issue). Nevertheless, the standard deviations, particularly for the medium and high alkalinity types, are large. The WFD requires reporting at the level of lake type, yet this within-group variability highlights the importance of

defining site-specific reference conditions rather than describing a chemical reference condition on a type-specific basis. It could be argued that some of the variation is associated with errors in the diatom inferences and, like all models, diatom transfer functions are not without their limitations. Light, temperature, flushing rate, substrate and grazing pressure have all been shown to influence diatom distributions in lakes (e.g. Anderson et al. 1993; Bennion 1994) and the problems of using diatoms as indicators of lake trophic status in some shallow systems have been widely acknowledged (e.g. Bennion 1995; Bennion et al. 2001; Sayer 2001). Nevertheless, when the errors in predicted DI-TP values are compared with the inter-annual TP ranges in lakes, the models can be considered to estimate TP with reasonable accuracy (Bennion et al. 2005). Furthermore, when the predictive ability of the diatom transfer functions are compared to OECD-type regression relationships, the errors associated with the former are as good or better (Anderson et al. 1993). Returning to our argument, background TP concentrations depend on a range of factors and are, therefore, site-specific. On the basis of the current study, it is recommended that type-specific values be used only as a guide when determining nutrient reference conditions and assessing ecological status. Carvalho et al. (2008) reached the same conclusion following their analysis of chlorophyll a concentrations in over 500 European reference lakes and highlighted the likelihood of greater errors in type-specific assessments for sites that lie close to type boundaries. However, preliminary analysis of the GB typology TP data appears to suggest that some of the lake types may be distinguishable on the basis of mean TP concentrations (Simpson and Bennion, in prep).

Ecological reference conditions

The classification of the diatom data from the reference samples of 101 of the original 106 lakes accords reasonably well with the GB lake typology scheme, indicating that the latter has some ecological relevance. Cluster 1, comprised of mostly shallow, high alkalinity and marl lakes, is located on the left of the plot, Cluster 2 lakes of intermediate alkalinity but containing both deep and shallow waters are positioned in the centre, and Cluster 3, comprised of predominantly deep, low alkalinity and several medium alkalinity lakes, is located on the far right of axis 1 (Fig. 2a). This indicates that alkalinity and, to a lesser extent, lake depth, the two criteria used in the GB lake typology scheme, are important in explaining the diatom distributions. This allows the diatom assemblages at reference condition to be described for the main lake types. The reference floras of the deep, low to medium alkalinity lakes were characterised by oligotrophic, acidophilous to circumneutral taxa, namely the planktonic *Cyclotella* taxa (*C. comensis, C. kuetzingiana*), *Brachysira vitrea, Tabellaria flocculosa, Achnanthes minutissima* and *Hannaea arcus*. These taxa have been observed in the pre-enrichment assemblages of numerous lakes lying on relatively base poor geology (e.g. Bennion et al. 2004). The reference floras of the lakes with intermediate alkalinity contained taxa commonly found in more productive, circumneutral to slightly alkaline

waters such as *Cyclotella radiosa*, *Cymbella microcephala* and *Fragilaria virescens* var *exigua.* Finally the reference floras of the high alkalinity and marl, predominantly shallow, lakes were characterised by the benthic *Fragilaria* species, *Amphora pediculus* and several planktonic taxa typically found in circumneutral to alkaline, productive waters, such as *Stephanodiscus parvus, S. hantzschii* and *Aulacoseira granulata.* With the exception of some of the West Midlands meres, which are known to have naturally high P concentrations (Moss et al. 1994), the *Stephanodiscus* taxa occur in low relative abundances in the reference samples compared with high percentages seen in the modern assemblages of enriched lakes (e.g. Bennion et al. 2004).

Whilst there is some degree of type-specific discrimination of the diatom floras, there are no clear discontinuities between clusters, and medium alkalinity lakes, for example, are present in all three clusters. Even within each lake typology class, lakes are likely to cover a broad alkalinity range and, therefore, we might expect poor cross matching of the diatom groups and the lake typology at these boundaries. Given that there are no discrete groups of lakes in nature and that variables such as alkalinity and depth are continuous, the emphasis placed on definition of type-specific reference conditions by the WFD does result in the forcing of water bodies into boxes that do not really exist. Our data suggest that whilst the typology can be used to broadly characterise ecological reference conditions, a site-specific approach is likely to be more ecologically meaningful and indeed the sediment record lends itself perfectly to such an approach. However, ongoing work is exploring further the extent to which lake types can be separated on the basis of their diatom communities (Simpson and Bennion, in prep), and the assessment of reference conditions for other biological elements in UK lakes such as macrophytes, phytoplankton and chironomids is also in progress.

Floristic change

According to the WFD, lakes must be classified into one of five ecological status classes (High, Good, Moderate, Poor, Bad) based on the degree to which present day conditions deviate from reference conditions. Whilst there has been considerable output in terms of ecological indicator development (e.g. Solheim and Gulati 2008), with notable exceptions (e.g. Moss et al. 2003; Søndergaard et al. 2005), few studies have attempted ecological classification. Here we have employed a simple dissimilarity measure, the SCD, between core bottom and top samples to calculate the deviation from the reference diatom community. Based on analysis of a much larger unpublished dataset of diatom assemblage change, the critical SCD score of 0.39 (significant change at the 2.5[th] percentile) was selected as the threshold below which a lake can be described as minimally impacted or high status, and hence a potential reference lake. In the current dataset of 106 UK lakes, only 25 sites met this criterion. The critical SCD score of 0.58 (significant change at the 10th percentile) was selected to represent significant floristic change and 57 lakes (> 50% of

the dataset) exceeded this value. Although the surface sediment data are not shown here, the nature of the changes were typical of those associated with nutrient enrichment. The SCD is a useful tool, therefore, for assessing whether lakes have experienced floristic change, but it is very difficult to decide how much change constitutes classification into each of the status bands. A question remains over where one should set the boundaries for defining Good and Moderate status whilst allowing for an element of natural change. If the 2.5th percentile (0.39) is selected to define the High/Good boundary then we might logically select the 5th percentile (0.48) to define the Good/Moderate boundary and the 10th percentile (0.58) to define the Moderate/Poor boundary. Using these criteria, 25 sites would be classified as high status, and 12, 12, and 57 lakes, would be classified as good, moderate or poor status, respectively. Alternatively, to comply strictly with the WFD normative definition of reference or high status as 'no or very minor deviation from undisturbed conditions', we might argue that the 1st percentile value of <0.29 should be used to define true reference lakes. In the current dataset, only 13 sites meet this criterion. Further work is required to refine the method for classifying lakes based on SCD scores.

The proposed scheme seems to provide a sensible framework for ecological classification but, given that diatoms are extremely sensitive to changes in water quality, may be too stringent and thereby the palaeoecological outputs require comparison with the results from contemporary classification tools to assess the appropriateness of the boundaries. Furthermore, it should be noted that whilst the SCD is a useful tool for assessing degree of change, it does not provide any information on the nature of the species shifts, and should therefore be used in combination with the species data themselves to establish the drivers and wider ecological significance of the change.

One limitation of the present study is that not all of the cores have been dated. Several of the cores are relatively short (< 30 cm) and in productive, lowland catchments sediment accumulation rates may be as high as 0.5 to 1 cm per year (Rose et al. this issue) resulting in a record that spans only approximately the last 50 rather than 150 years. Furthermore, poor diatom preservation at a number of sites, particularly the marl lakes, prevented analysis of the deep sediment layers and the lowermost sample for which data were available may not extend back to the pre-enrichment period. It remains uncertain, therefore, whether all of our cores cover comparable time periods.

Summary

The study illustrates that palaeolimnology is a valuable tool for defining reference conditions and ecological status with respect to eutrophication pressures. When datasets are sufficiently well populated, diatom transfer functions can be employed to define reference values of TP for particular lake types. The diatom community composition data themselves, which have been

largely overlooked given the emphasis on transfer function development in recent decades, can provide an assessment of ecological reference conditions. Nevertheless, multi-proxy palaeolimnological studies could define site-specific ecological reference conditions in a more holistic way than can be achieved using a single indicator group and the addition of, for example, Cladocera, plant macrofossil and chironomid data, is recommended.

Simple measures of compositional change between reference and surface samples in cores, such as the SCD coefficient applied here, offer a powerful method for assessing degree of ecological change. These measures could prove useful in determining the extent to which the biological groups return to reference conditions following restoration efforts. Climate change may confound remediation attempts and make it impossible to achieve restoration targets, although presently lake response to the combination of reduced nutrient loading and climate change is poorly understood (Battarbee et al. 2005). It is generally assumed that increasing global temperatures will result in greater phytoplankton biomass and dominance of cyanobacteria but in some regions climate warming may stimulate macrophyte growth and potentially improve water clarity (Jeppesen et al. 2007). Whilst we must accept that restoration targets may require modification to account for such climate induced alterations, the long temporal perspective and direct ecological information offered by palaeolimnology ensures that we have a sound baseline against which to assess future change.

Acknowledgements

References

Anderson NJ (1995) Using the past to predict the future: lake sediments and the modelling of limnological disturbance. Ecol Model 78: 149-172

Anderson NJ (1997) Historical changes in epilimnetic phosphorus concentrations in six rural lakes in Northern Ireland. Freshwat Biol 38: 427-440

Anderson NJ, Rippey B, Gibson CE (1993) A comparison of sedimentary and diatom-inferred phosphorus profiles: implications for defining pre-disturbance nutrient conditions. Hydrobiologia 253: 357-366

Anderson NJ, Jeppesen E, Søndergaard M (2005) Ecological effects of reduced nutrient loading (oligotrophication) on lakes: an introduction. Freshwat Biol 50: 1589–1593

Appleby PG, Nolan PJ, Gifford DW, Godfrey MJ, Oldfield F, Anderson NJ, Battarbee RW (1986) ^{210}Pb dating by low background gamma counting. Hydrobiologia 141: 21-27

Battarbee RW (1999) The importance of paleolimnology to lake restoration. Hydrobiologia 395/396: 149-159

Battarbee RW, Anderson NJ, Jeppesen E, Leavitt PR (2005) Combining palaeolimnological and limnological approaches in assessing lake ecosystem response to nutrient reduction. Freshwat Biol 50:1772-1780

Battarbee RW, Morley D, Bennion H, Simpson GL, Hughes M, Bauere V (this issue) A palaeolimnological meta-database for assessing the ecological status of lakes. J Paleolimnol (this issue)

Battarbee RW, Jones VJ, Flower RJ, Cameron NG, Bennion H, Carvalho L, Juggins S. (2001a) Diatoms. In: Smol JP, Birks HJB, Last WM (eds) Tracking Environmental Change Using Lake Sediments. Volume 3: Terrestrial, Algal, and Siliceous Indicators. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 155-202

Battarbee RW, Juggins S, Gasse F, Anderson NJ, Bennion H, Cameron NG, Ryves DB, Pailles C, Chalie F, Telford R. (2001b) European Diatom Database (EDDI). An Information System for Palaeoenvironmental Reconstruction. ECRC Research Report No 81, University College London: 94 pp

Bennion H (1994) A diatom-phosphorus transfer function for shallow, eutrophic ponds in south-east England. Hydrobiologia 275/6: 391-410

Bennion H (1995) Surface sediment diatom assemblages in shallow, artificial, enriched ponds, and implications for reconstructing trophic status. Diatom Res 10: 1-19.

Bennion H, Battarbee R (2007) The European Union Water Framework Directive: opportunities for palaeolimnology. J Paleolimnol 38: 285-295

Bennion H, Smith MA (2000) Variability in the water chemistry of shallow ponds in southeast England, with special reference to the seasonality of nutrients and implications for modelling trophic status. Hydrobiologia 436: 145-158

Bennion H, Appleby PG, Phillips GL (2001) Reconstructing nutrient histories in the Norfolk Broads, UK: implications for the role of diatom-total phosphorus transfer functions in shallow lake management. J Paleolimnol 26: 181-204

Bennion H, Fluin J, Simpson GL (2004) Assessing eutrophication and reference conditions for Scottish freshwater lochs using subfossil diatoms. J Appl Ecol 41: 124-138

Bennion H, Juggins S, Anderson NJ (1996a) Predicting epilimnetic phosphorus concentrations using an improved diatom-based transfer function and its application to lake eutrophication management. Environ Sci Technol 30: 2004-2007

Bennion H, Wunsam S, Schmidt R (1995) The validation of diatom-phosphorus transfer functions: an example from Mondsee, Austria. Freshwat Biol 34: 271-283

Bennion, H, Johnes P, Ferrier R, Phillips G, Haworth E (2005). A comparison of diatom phosphorus transfer functions and export coefficient models as tools for reconstructing lake nutrient histories. Freshwat Biol 50: 1651-1670

Bennion H, Duigan CA, Haworth EY, Allott TEH, Anderson NJ, Juggins S, Monteith DM (1996b) The Anglesey lakes, Wales, UK - changes in trophic status of three standing waters as inferred from diatom transfer functions and their implications for conservation. Aquat Conserv Mar: Freshwat Ecosyst 6: 81-92

Birks HJB (1998) Numerical tools in palaeolimnology - progress, potentialities, and problems. J Paleolimnol 20: 307-332

Bradshaw EG, Anderson NJ (2001) Validation of a diatom-phosphorus calibration set for Sweden. Freshwat Biol 46: 1035-1048

Bradshaw EG, Nielsen AB, Anderson NJ (2006) Using diatoms to assess the impacts of prehistoric, pre-industrial and modern land-use on Danish lakes. Reg Environ Change 6: 17–24

Cardoso AC, Solimini A, Premazzi G, Carvalho L, Lyche A, Rekolainen S (2007) Phosphorus reference concentrations in European lakes. Hydrobiol 584: 3-12

Carpenter SR (2005) Eutrophication of aquatic ecosystems: Bistability and soil phosphorus. Proc Natl Acad Sci USA 102: 10002-10005

Carrick HJ, Worth D and Marshall ML (1994) The influence of water circulation on chlorophyll-turbidity relationships in Lake Okeechobee as determined by remote-sensing. J Plankton Res 16: 1117-1135

Carvalho L, Solimini A, Phillips G, van den Berg M, Pietiläinen O-P, Lyche Solheim A, Poikane S and Mischke U (2008) Chlorophyll reference conditions for European lake types used for intercalibration of ecological status. Aquat Ecol 42: 203-211

Chen G, Dalton C, Leira M, Taylor D (2008) Diatom-based total phosphorus (TP) and pH transfer functions for the Irish Ecoregion. J Paleolimnol 40:143-163

Dufrêne M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol Monogr 67:345-366

European Union (2000) Directive 2000/60/EC of the European Parliament and of the Council of 23 October 20000 establishing a framework for Community action in the field of water policy. Official Journal of the European Communities L327: 1-72

Fozzard IR, Doughty CR, Ferrier RC, Leatherland TM, Owen R (1999) A quality classification for management of Scottish standing waters. Hydrobiologia 395/396: 433-453

Gibson CE, Foy RH, Bailey-Watts AE (1996) An analysis of the total phosphorus cycle in some temperate lakes: the response to enrichment. Freshwat Biol 35: 525-532

Hall RI, Smol JP (1999) Diatoms as indicators of lake eutrophication. In: Stoermer EF, Smol JP (eds) The Diatoms: Applications for the Environmental and Earth Sciences. Cambridge University Press, Cambridge, pp 128-168

Holdren GC, Armstrong DE (1980) Factors affecting phosphorus release from intact sediment cores. Environ Sci Technol 14: 79-87

Hutchinson GE (1969) Eutrophication, past and present. In: Rohlich GA (ed.) Eutrophication: Causes, Consequences and Correctives. Washington D.C., National Academy of Sciences, pp 17-26

Jeppesen E, Søndergaard M, Meerhoff M, Lauridsen TL, Jensen JP (2007) Shallow lake restoration by nutrient loading reduction—some recent findings and challenges ahead. Hydrobiologia 584: 239–252

Juggins S (2003) C^2 User guide. Software for ecological and palaeoecological data analysis and visualisation. University of Newcastle, Newcastle upon Tyne, UK, 69 pp.

Kauppila T, Moisio T, Salonen VP (2002) A diatom-based inference model for autumn epilimnetic total phosphorus concentration and its application to a presently eutrophic boreal lake. J Paleolimnol 27: 261-273

Legendre P and Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280

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. J Paleolimnol 19: 443-463

Martens H, Naes T (1989) Methods for calibration. In: Martens H, Naes T (eds) Multivariate Calibration. Wiley, Chichester, pp 73-85

Minchin PR (1987) An evaluation of relative robustness of techniques for ecological ordinations. Vegetatio 71: 145-156

Moss B (1998) Shallow lakes: biomanipulation and eutrophication. SCOPE Newsletter 29: 45

Moss B, Johnes P, Phillips G (1997) New approaches to monitoring and classifying standing waters. In: Boon PJ, Howell CL eds) Freshwater Quality: Defining the Indefinable. HMSO, Edinburgh, pp 118-133

Moss B, McGowan S, Carvalho L (1994) Determination of phytoplankton crops by top-down and bottom-up mechanisms in a group of English lakes, the West Midland Meres. Limnol Oceanogr 39: 1020-1029

Moss B, Stephen D, Alvarezn C, Becares E, Van der Bund W, Collings SE et al. (2003) The determination of ecological status in shallow lakes—a tested system (ECOFRAME) for implementation of the European Water Framework Directive. Aquat Conserv Mar: Freshwat Ecosyst 13:507-549

Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL and Stevens MHH (2008). vegan: Community Ecology Package. R package version 1.11-5. http://cran.r-project.org/ http://vegan.r-forge.r-project.org/

Organisation for Economic Co-operation and Development, OECD. (1982) Eutrophication of waters: monitoring, assessment and control. Technical Report, Environmental Directorate, OECD, Paris: 154 pp.

Overpeck JT, Webb T, Prentice IC (1985) Quantitative interpretation of fossil pollen spectra - dissimilarity coefficients and the method of modern analogs. Quatern Res 23: 87-108

Phillips G (2003) Reporting typology for Ecoregion 18, Great Britain. TAG/LTT 43, March 2003

Phillips G, Pietiläinen O-P, Carvalho L, Solimini A, Lyche Solheim A and Cardoso AC (2008) Chlorophyll–nutrient relationships of different lake types using a large European dataset. Aquat Ecol 42: 213-226

Pollard P, Huxham M (1998) The European Water Framework Directive: a new era in the management of aquatic ecosystem health? Aquat Conserv Mar: Freshwat Ecosyst 8: 773-792

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Rippey B, Anderson NJ, Foy RH (1997) Accuracy of diatom-inferred total phosphorus concentrations and the accelerated eutrophication of a lake due to reduced flushing and increased internal loading. Can J Fish Aquat Sci: 54 2637-2646

Roberts DW (2007) labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.3-1. http://ecology.msu.montana.edu/labdsv/R

Rose NL, Morley D, Appleby PG, Battarbee RW, Alliksaar T, Guilizzoni P, Jeppesen E, Korhola A, Punning JM (this issue) Sediment accumulation rates in European lakes since AD 1850: Trends, reference conditions and exceedence. J Paleolimnol (this issue)

Rose NL, Harlock S, Appleby PG, Battarbee RW (1995) The dating of recent lake sediments in the United Kingdom and Ireland using spheroidal carbonaceous particle concentration profiles. The Holocene 5: 328-335

Sayer C (2001) Problems with the application of diatom-total phosphorus transfer functions: examples from a shallow English lake. Freshwat Biol 46: 743-757

Simpson GL, Shilland EM, Winterbottom JM and Keay J (2005) Defining reference conditions for acidified waters using a modern analogue approach. Environ Pollut 137: 119–133

Simpson GL (2007a) analogue: Analogue matching and Modern Analogue Technique transfer function models. R package version 0.5-2. http://analogue.r-forge.r-project.org/

Simpson GL (2007b) Analogue Methods in Palaeoecology: Using the analogue Package. J Stat Softw 22: 1-29

Smith VH, Joye SB, Howarth RW (2006) Eutrophication of freshwater and marine ecosystems. Limnol Oceanogr 51: 351–355

Smol JP (2008) Pollution of Lakes and Rivers: A Paleoenvironmental Perspective, 2nd Edition. Blackwell Publishing, Oxford. 383 pp

Solheim AL, Gulati RD (2008) Preface: Quantitative ecological responses for the Water Framework Directive related to eutrophication and acidification of European lakes Aquat Ecol 42: 179-181

Søndergaard M, Jeppesen E, Jensen JP, Amsinck SL (2005) Water Framework Directive: ecological classification of Danish lakes. J App Ecol 42: 616-629

Søndergaard M, Jeppesen E, Lauridsen T, Van Nes SCH, Roijackers R, Lammens E. Portielje R (2007) Lake restoration: successes, failures and long-term effects. J Appl Ecol 44: 1095–1105

Stoermer EF, Smol JP (eds) (1999) The Diatoms: Applications for the Environmental and Earth Sciences. Cambridge University Press, Cambridge

Taylor D, Dalton C, Leira M, Jordan P, Chen G, León-Vintró L, Irvine K, Bennion H, Nolan T (2006) Recent histories of six productive lakes in the Irish Ecoregion based on multiproxy palaeolimnological evidence. Hydrobiologia 571: 237-259

ter Braak CJF, Juggins S (1993) Weighted averaging partial least squares (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia 269/270: 485-502

ter Braak CJF, van Dam H (1989) Inferring pH from diatoms: a comparison of old and new calibration methods. Hydrobiologia 178: 209-223

Vighi M and Chiaudani G (1985) A simple method to estimate lake phosphorus concentrations resulting from natural, background, loadings. Water Res 19: 987–991

Vollenweider RA (1968) Scientific fundamentals of the eutrophication of lakes and flowing waters, with particular reference to nitrogen and phosphorus as factors in eutrophication. Technical Report DAS/CSI/68 27, OECD, Paris: 159 pp

Wunsam S, Schmidt R (1995) A diatom-phosphorus transfer function for Alpine and pre-alpine lakes. Mem Ist Ital Idrobiol 53: 85-99

Table 1 The lake typology for Great Britain indicating the number of lakes of each type in the chemical (N=169) and ecological (N=106) reference conditions datasets

Type description	Catchment geology	Mean depth (m)	No. in chemical reference condition dataset	No. in ecological reference condition dataset
Low alkalinity-Deep (LA-D)	>90% siliceous	> 3.0	77	21
Low alkalinity-Shallow (LA-Sh)	>90% siliceous	≤3.0	25	2
Medium alkalinity-Deep (MA-D)	50-90% siliceous	> 3.0	23	23
Medium alkalinity-Shallow (MA-Sh)	50-90% siliceous	≤3.0	7	13
High alkalinity-Deep (HA-D)	>50% calcareous	> 3.0	9	8
High alkalinity-Shallow (HA-Sh)	>50% calcareous	≤3.0	23	28
Marl-Deep (Marl-D)	>65% limestone	> 3.0	0	3
Marl-Shallow (Marl-Sh)	>65% limestone	≤3.0	5	8

Table 2 The indicator value and significance of indicator taxa associated with each of the clusters. Indicator taxa are identified as those with significant indicator values at the P < 0.01 level

Diatom code	Name	Cluster	Indicator value	Significance (P value)
FR002C	*Fragilaria construens* var. *venter*	1	0.6564	0.001
AM012A	*Amphora pediculus*	1	0.6205	0.001
FR006A	*Fragilaria brevistriata*	1	0.6168	0.001
FR001A	*Fragilaria pinnata*	1	0.5299	0.001
FR002A	*Fragilaria construens*	1	0.5111	0.001
AM001A	*Amphora ovalis*	1	0.3894	0.001
ST010A	*Stephanodiscus parvus*	1	0.3779	0.001
ST001A	*Stephanodiscus hantzschii*	1	0.3694	0.001
AC001A	*Achnanthes lanceolata*	1	0.3641	0.006
AU003A	*Aulacoseira granulata*	1	0.3383	0.001
AC023A	*Achnanthes conspicua*	1	0.3017	0.001
CC001A	*Cyclostephanos dubius*	1	0.2424	0.001
GY005A	*Gyrosigma acuminatum*	1	0.2317	0.009
NA030A	*Navicula menisculus*	1	0.2274	0.002
CO005A	*Cocconeis pediculus*	1	0.2051	0.001
CC002A	*Cyclostephanos invisitatus*	1	0.1818	0.002
FR005D	*Fragilaria virescens* var. *exigua*	2	0.6562	0.001
CY019A	*Cyclotella radiosa*	2	0.4189	0.001
CM004A	*Cymbella microcephala*	2	0.4021	0.003
NA003A	*Navicula radiosa*	2	0.3938	0.001
EU047A	*Eunotia incisa*	2	0.353	0.004
AC035A	*Achnanthes pusilla*	2	0.3393	0.006
NA013A	*Navicula pseudoscutiformis*	2	0.316	0.001
NA002A	*Navicula jaernefeltii*	2	0.3053	0.001
FR009J	*Fragilaria capucina* var. *perminuta*	2	0.25	0.001
NA042A	*Navicula minima*	2	0.2499	0.005
DP067A	*Diploneis modica*	2	0.22	0.001
FR063A	*Fragilaria robusta*	2	0.2103	0.009
EU002E	*Eunotia pectinalis* var. *minor* fo. *Impressa*	2	0.2018	0.002
AC161A	*Achnanthes ventralis*	2	0.1389	0.008
CY9991	*Cyclotella kuetzingiana* agg.	3	0.757	0.001
CY010A	*Cyclotella comensis*	3	0.696	0.001
BR001A	*Brachysira vitrea*	3	0.6294	0.001
TA9996	*Tabellaria flocculosa* agg.	3	0.5823	0.001
HN001A	*Hannaea arcus*	3	0.5291	0.001
AC013A	*Achnanthes minutissima*	3	0.493	0.001
FU002A	*Frustulia rhomboides*	3	0.3896	0.001
CM018A	*Cymbella gracilis*	3	0.3524	0.001
CY007A	*Cyclotella glomerata*	3	0.2913	0.001
PE002A	*Peronia fibula*	3	0.2188	0.001
EU009A	*Eunotia exigua*	3	0.1945	0.004
DT002A	*Diatoma hyemale*	3	0.1875	0.002
DE002A	*Denticula elegans*	3	0.1685	0.004
AC022A	*Achnanthes marginulata*	3	0.1581	0.01
CM013A	*Cymbella helvetica*	3	0.1507	0.006

Table 3 The indicator value and significance of the association of each lake type with the diatom-based clusters. The LA-Sh lakes were outliers and were therefore omitted from the cluster analyses. The cluster column indicates the diatom-based cluster for which the 'species' indicator value was maximal. Level of significance: ****$P \leq 0.001$, ***$P=0.001-0.01$, **$P=0.01-0.05$, *$P=0.05-0.1$ based on 1000 permutations

Lake type	Cluster	Indicator value	Significance (P value)
LA-D	3	0.5938	0.001****
MA-D	3	0.2154	0.01***
MA-Sh	2	0.2821	0.001****
HA-D	1	0.0632	0.413
HA-Sh	1	0.3599	0.001****
Marl-D	1	0.0909	0.057*
Marl-Sh	1	0.1577	0.01***

Table 4 The distribution of squared chord distance dissimilarity coefficient values for each lake type in the ecological reference conditions dataset (N=106)

Lake type	<0.39	0.39-0.48	0.48-0.58	>0.58
LA-D	7	4	3	7
LA-Sh	0	0	0	2
MA-D	5	4	0	14
MA-Sh	3	0	2	8
HA-D	3	0	0	5
HA-Sh	6	3	3	16
Marl-D	1	0	1	1
Marl-Sh	0	1	3	4
Total	25	12	12	57

Figure captions

Fig. 1 Boxplot showing the distribution of the diatom-inferred TP values for the reference samples (~1850 AD) of each lake type (number in parentheses is the number of lakes in that type; total number of lakes is 169). See Table 1 for lake typology codes.

Fig. 2 a) Non-Metric Multi-dimensional scaling (nMDS) ordination (axes 1 and 2) of the between-site (Hellinger) distances in the UK lakes diatom reference assemblage dataset (N=101). The three clusters identified by *k*-means clustering are shown and symbols depict the GB lake types; see Appendix 1 for lake names and Table 1 for lake typology codes, b) Non-Metric Multi-dimensional scaling (nMDS) ordination (axes 1 and 2) of the species scores in the UK lakes diatom reference assemblage dataset. Indicator taxa are shown in bold italic font. See Appendix 2 for taxon names

Fig. 3 Boxplot showing the distribution of the squared chord distance (SCD) dissimilarity scores for each lake type (number in parentheses is the number of lakes in that type; total number of lakes is 106). The SCD critical values at the 2.5th, 5th and 10th percentiles are indicated by horizontal lines. See Table 1 for lake typology codes.
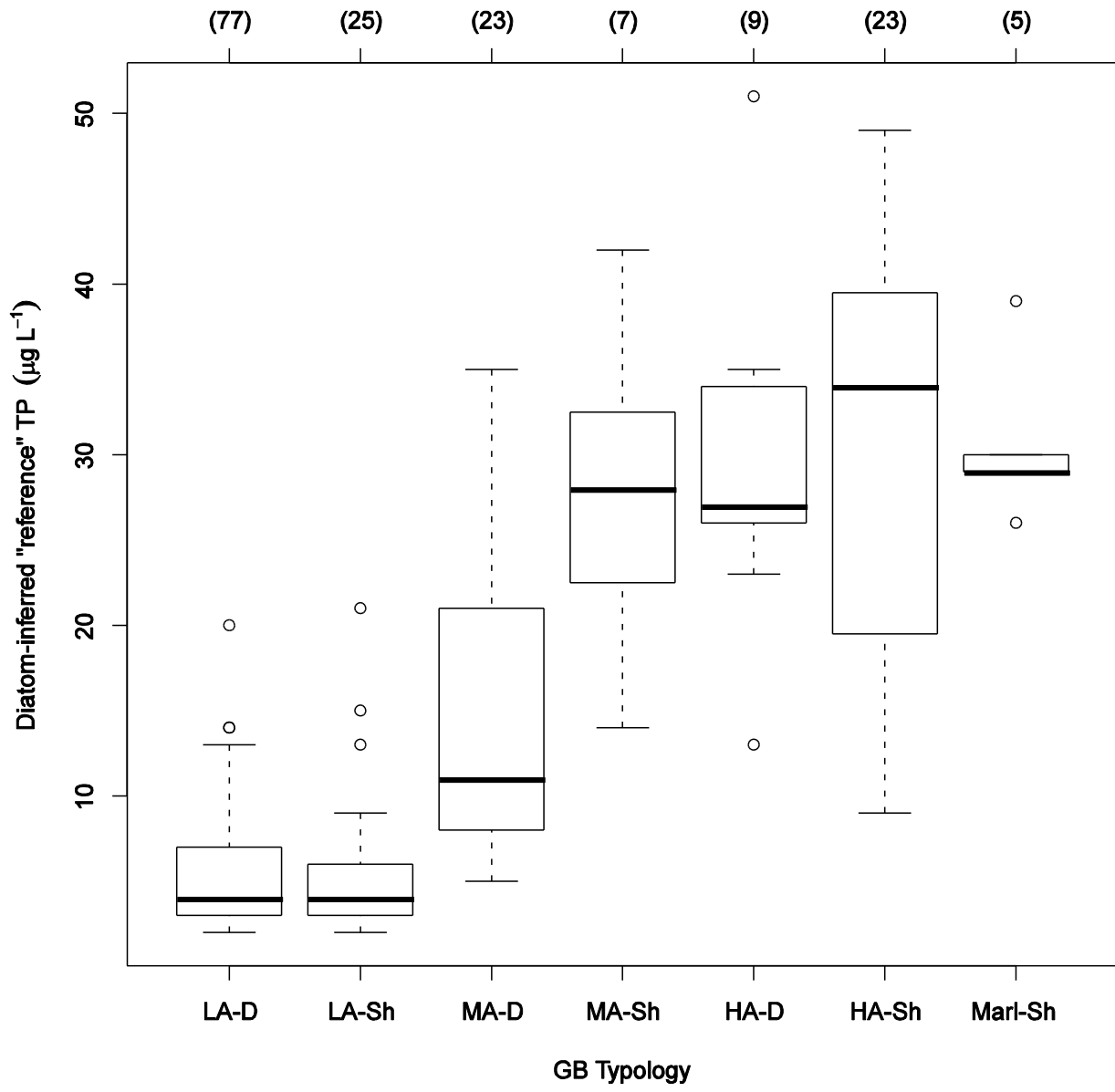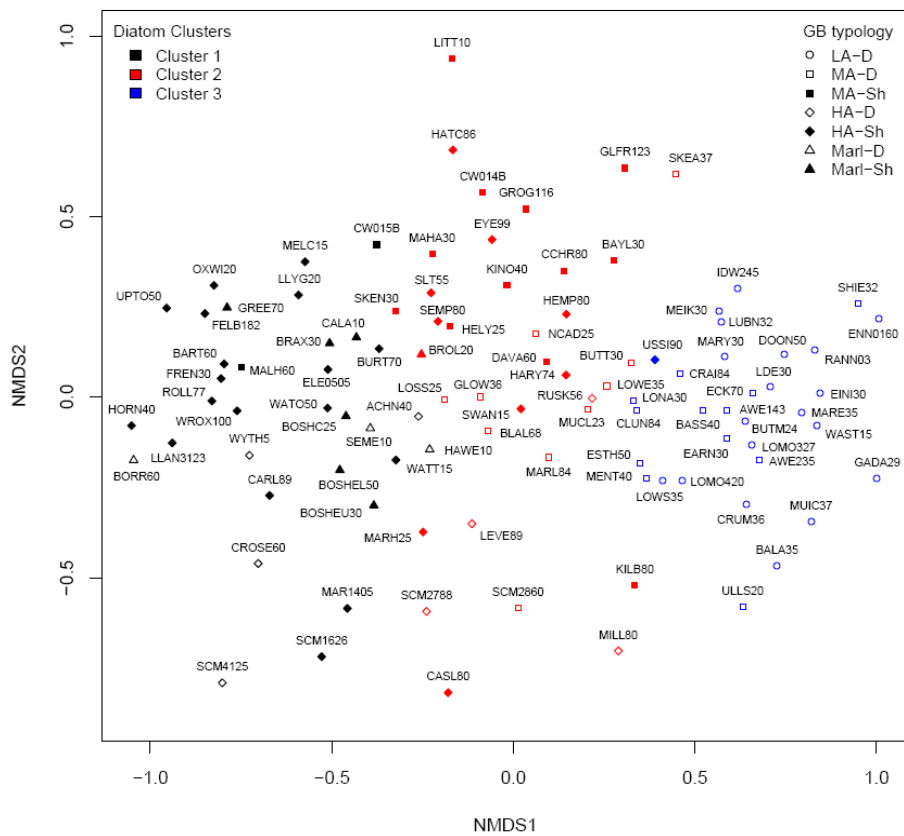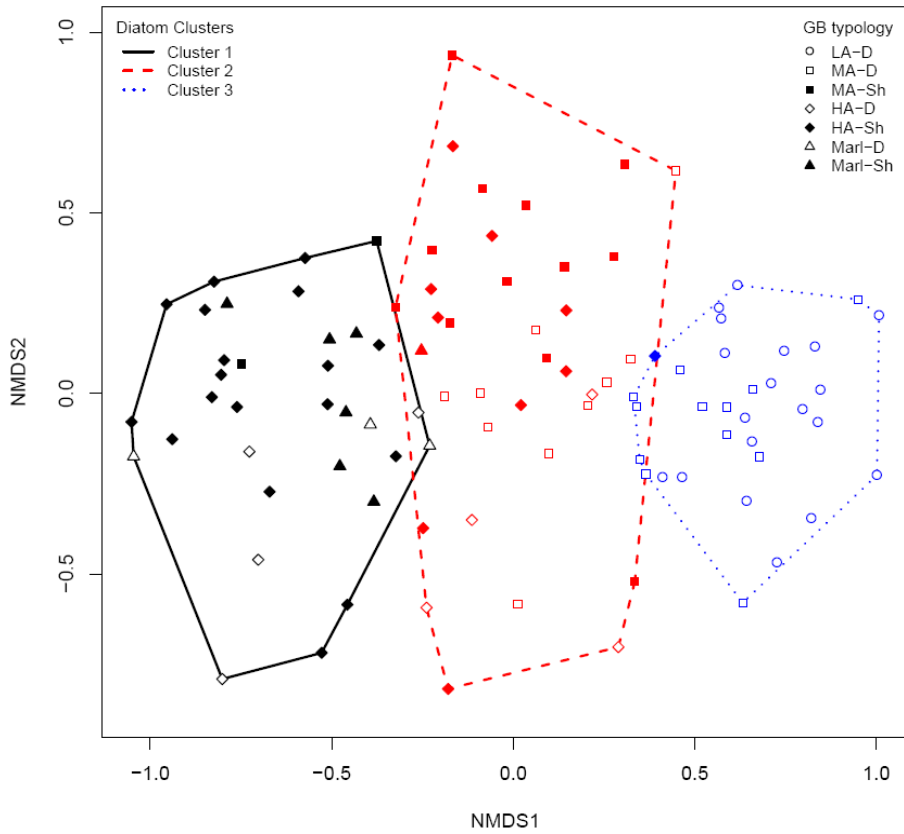
Fig. 1

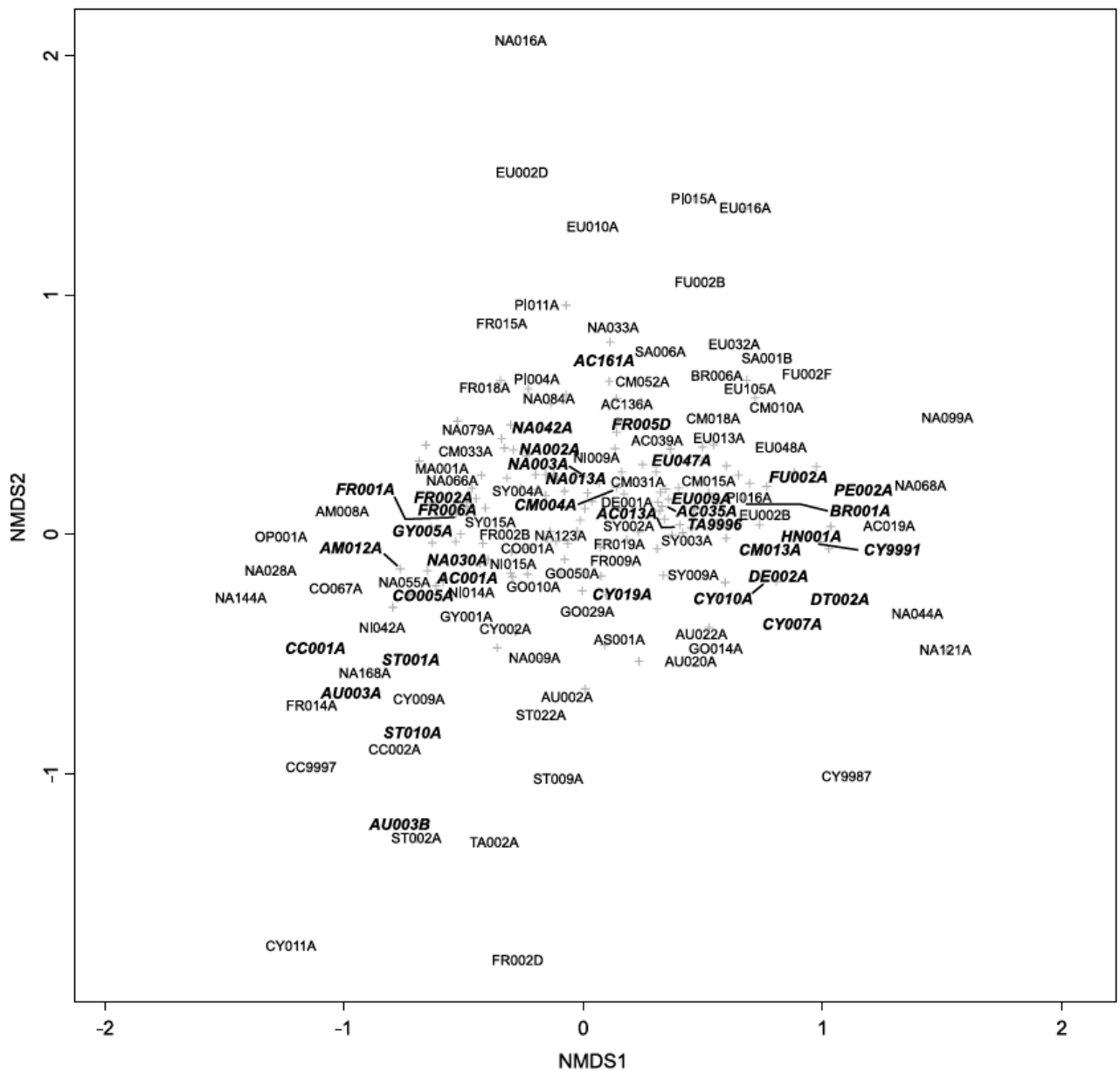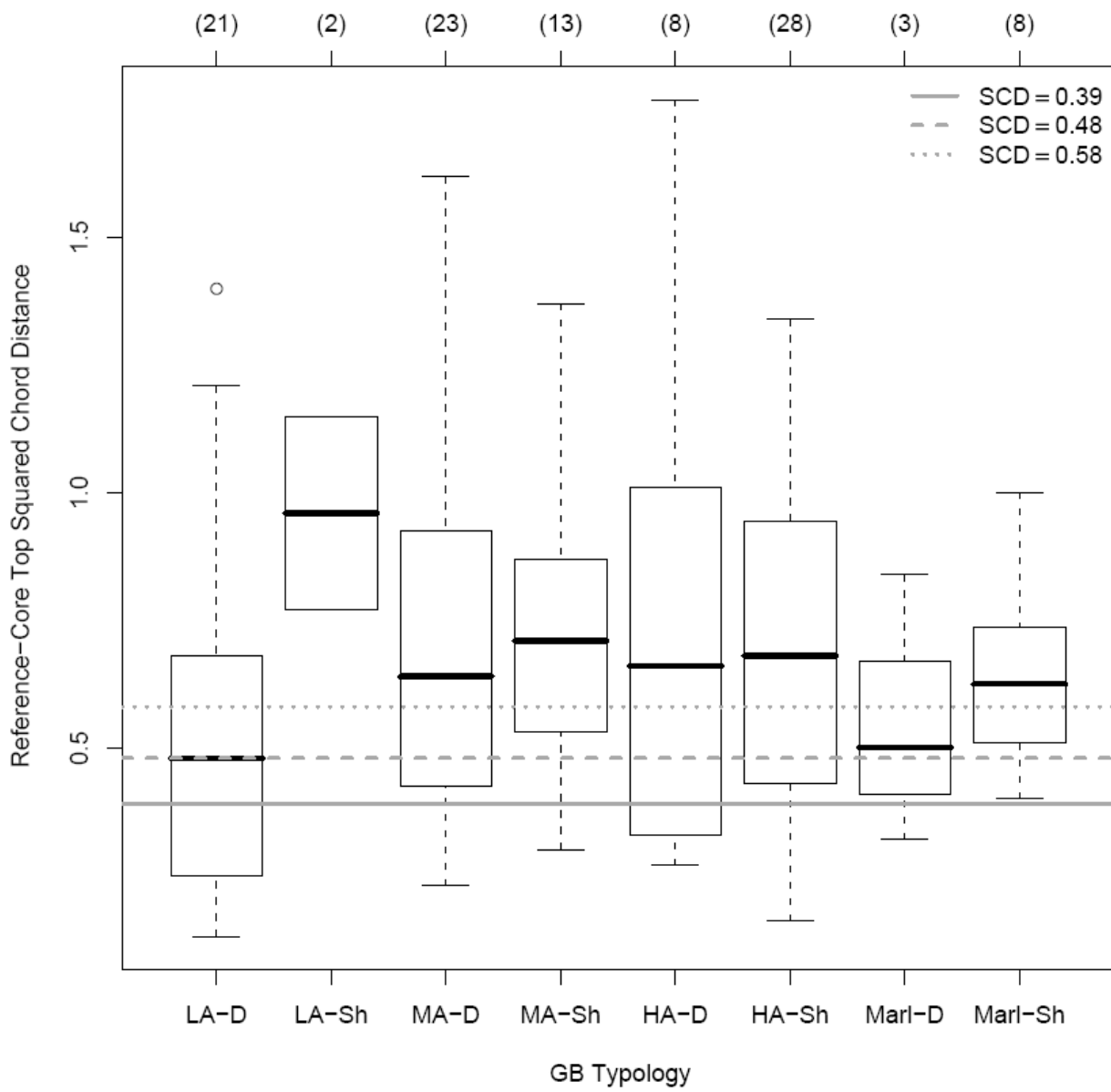**Fig. 2**

a)

Fig. 2

b)

Fig. 3

Appendix 1 List of lake sample codes with full names, location and typology

Sample code	Site name	UK grid reference	GB typology
ACHN40	Loch Achnacloich	NH 665736	HA, D
AWE143	Loch Awe North Basin	NM 930 065	MA, D
AWE235	Loch Awe South Basin	NM 930 065	MA, D
BALA35	Lake Bala or Llyn Tegid	SH 905347	LA, D
BART60	Barton Broad	TG 363215	HA, Sh
BASS40	Bassenthwaite Lake	NY 214296	MA, D
BAYL30	Bayfield Loch	NH821718	MA, Sh
BLAL68	Black Loch	NT 075961	MA, D
BORR60	Loch Borralie	NC 381 668	Marl, D
BOSHC25	Bosherston Lake, Central	SR 976946	Marl, Sh
BOSHEL50	Bosherston Lake, Lower	SR 978952	Marl, Sh
BOSHEU30	Bosherston Lake, Upper	SR 978960	Marl, Sh
BRAX30	Branxhome Easter Loch	NT434118	Marl, Sh
BROL20	Broomlee Lough	NY790697	Marl, Sh
BURT70	Burton Mill Pond	SU 974175	HA, Sh
BUTM24	Buttermere	NY 182157	LA, D
BUTT30	Loch of Butterstone	NO 058449	MA, D
CALA10	Loch Caladail	NC 396666	Marl, Sh
CARL89	Carlingwark Loch	NX 765615	HA, Sh
CASL80	Castle Loch	NY 090815	HA, Sh
CCHR80	Loch Cill Chriosd	NG 611205	MA, Sh
CLUN84	Loch of Clunie	NO 115442	MA, D
CRAI84	Loch of Craiglush	NO 042444	MA, D
CROSE60	Crose Mere	SJ 430305	HA, D
CRUM36	Crummock Water	NY 157188	LA, D
CW014B	Talley Upper	SN 633332	MA, Sh
CW015B	Talley Lower	SN 632337	MA, Sh
CW024B	Llyn Alwen	SH 898567	LA, D
DAVA60	Loch Davan	NJ 442007	MA, Sh
DOON50	Loch Doon	NX 495985	LA, D
EARN30	Loch Earn	NN 640235	MA, D
ECK70	Loch Eck	NS 141939	MA, D
EIDW40	Llyn Eiddwen	SN 606670	LA, Sh
EINI30	Loch Einich	NN 913990	LA, D
ELE0505	Eleven Acre Lake	SP 675369	HA, Sh
ENN0160	Ennerdale Water	NY 110150	LA, D
ESTH50	Esthwaite Water	SD 358969	MA, D
EYE99	Loch Eye	NH 830795	HA, Sh
FELB182	Felbrigg Lake	TG 191388	HA, Sh
FNOD40	Llyn Fanod	SH 603643	LA, D
FOEL32	Llyn y Foel-frech	SH 919593	LA, Sh
FREN30	Frensham Great Pond	SU 845401	HA, Sh
GADA29	Llyn y Gadair	SH 707135	LA, D
GLFR123	Llyn Glasfryn	SH 402422	MA, Sh
GLOW36	Loch Glow	NT 087957	MA, D
GREE70	Greenlee Lough	NY 774698	Marl, Sh
GROG116	Loch Grogary (Croghearraidh)	NF 716711	MA, Sh
HARY74	Loch of Harray	HY 295155	HA, Sh
HATC86	Hatchet Pond	SU 367016	HA, Sh
HAWE10	Hawes Water	SD 477766	Marl, D
HELY25	Llyn Helyg	SJ112772	MA, Sh
HEMP80	Loch Hempriggs	ND 343471	HA, Sh
HORN40	Hornsea Mere	TA 190470	HA, Sh
IDW245	Llyn Idwal	SH 645596	LA, D
KILB80	Kilbirnie Loch	NS 330545	MA, Sh
KINO40	Loch Kinord	NO 442995	MA, Sh
LDE30	Loch Dee	NX 470790	LA, D
LEVE89	Loch Leven	NO 150025	HA, D
LITT10	Little Sea Mere	SZ 029846	MA, Sh
LLAN3123	Llangorse Lake	SO 132265	HA, Sh
LLYG20	Llyn Llygeirian	SH 346898	HA, Sh

LOMO327	Loch Lomond North Basin	NS 365945	LA, D
LOMO420	Loch Lomond South Basin	NS 365945	LA, D
LONA30	Loch Lonachan	NG 626189	MA, D
LOSS25	Loch Lossit	NR408652	MA, D
LOWE35	Loch of Lowes	NO 049439	MA, D
LOWS35	Loweswater	NY 126217	LA, D
LUBN32	Loch Lubnaig	NN 585130	LA, D
MAHA30	Loch Mahaick	NN 706068	MA, Sh
MALH60	Malham Tarn	SD 895667	Marl, Sh
MAR1405	Marsworth Reservoir	SP 922137	HA, Sh
MARE35	Loch Maree	NG 985675	LA, D
MARH25	Martnaham Loch	NS394173	HA, Sh
MARL84	Marlee Loch	NO 145443	MA, D
MARY30	St Marys Loch	NT252229	LA, D
MEIK30	Loch Meiklie	NH434301	LA, D
MELC15	Melchett Mere	SJ750811	HA, Sh
MENT40	Lake of Menteith	NN 580005	MA, D
MILL80	Mill Loch	NY 077833	HA, D
MUCL23	Muckle Water	HY 395300	MA, D
MUIC37	Loch Muick	NO 290830	LA, D
NCAD25	Loch nan Cadhan	NR404668	MA, D
NESS9048	Loch Ness	NH 535295	MA, D
OXWI20	Oxwich Pool	SS503872	HA, Sh
RANN03	Loch Rannoch	NN 610580	LA, D
ROLL77	Rollesby Broad	TG 464142	HA, Sh
RUSK56	Loch Rusky	NN 615034	HA, D
SCM1626	Hanmer Mere	SJ 454395	HA, Sh
SCM2788	Betton Pool	SJ 511079	HA, D
SCM2860	Bomere Pool	SJ 498080	MA, D
SCM4125	Tatton Mere	SJ755801	HA, D
SEME10	Semer Water	SD 918874	Marl, D
SEMP80	Castle Semple Loch	NS 365590	HA, Sh
SHIE32	Loch Shiel	NM 866771	MA, D
SKEA37	Loch Skealtar (Sgealtair)	NF 897686	MA, D
SKEN30	Loch of Skene	NJ 785075	MA, Sh
SLT55	Slapton Ley	SX 824435	HA, Sh
SWAN15	Loch of Swannay	HY 304293	HA, Sh
ULLS20	Ullswater	NY425204	MA, D
UPTO50	Upton Broad	TG 388134	HA, Sh
USSI90	Loch Ussie	NH 505574	HA, Sh
WAST15	Wast Water	NY 165060	LA, D
WATO50	Loch Watston	NN 711003	HA, Sh
WATT15	Loch Watten	ND 229561	HA, Sh
WROX100	Wroxham Broad	TG 308165	HA, Sh
WYTH5	Llyn yr Wyth Eidion	SH 474819	HA, D

Appendix 2 List of diatom codes with full names and authorities

Code	Name and authority
AC001A	*Achnanthes lanceolata* (Breb. ex Kutz.) Grun. in Cleve & Grun. 1880
AC013A	*Achnanthes minutissima* Kutz. 1833
AC019A	*Achnanthes nodosa* A. Cleve-Euler 1900
AC022A	*Achnanthes marginulata* Grun. in Cleve & Grun. 1880
AC023A	*Achnanthes conspicua* A. Mayer 1919
AC035A	*Achnanthes pusilla* Grun. in Cleve & Grun. 1880
AC039A	*Achnanthes didyma* Hust. 1933
AC136A	*Achnanthes subatomoides* (Hust.) Lange-Bertalot & Archibald in Krammer & Lange-Bertalot 1985
AC161A	*Achnanthes ventralis* (Krasske) Lange-Bertalot 1989
AM001A	*Amphora ovalis* (Kutz.) Kutz. 1844
AM008A	*Amphora thumensis* (Mayer) A. Cleve
AM012A	*Amphora pediculus* (Kutz.) Grun.
AS001A	*Asterionella formosa* Hassall 1850
AU002A	*Aulacoseira ambigua* (Grun. in Van Heurck) Simonsen 1979
AU003A	*Aulacoseira granulata* (Ehrenberg) Simonson 1979
AU003B	*Aulacoseira granulata* var. *angustissima* (O. Mull.) Simonsen 1979
AU020A	*Aulacoseira subarctica* (O.Mull.) Haworth
AU022A	*Aulacoseira subborealis* SWAP 1989
BR001A	*Brachysira vitrea* (Grun.) R. Ross in Hartley 1986
BR006A	*Brachysira brebissonii* R. Ross in Hartley 1986
CC001A	*Cyclostephanos dubius* (Fricke in A. Schmidt) Round 1982
CC002A	*Cyclostephanos invisitatus* Theriot, Stoermer & Hakansson, comb. nov. 1987
CC9997	*Cyclostephanos* [cf. *tholiformis*] H. Bennion 1992
CM004A	*Cymbella microcephala* Grun. in Van Heurck 1880
CM010A	*Cymbella perpusilla* A. Cleve 1895
CM013A	*Cymbella helvetica* Kutz. 1844
CM015A	*Cymbella cesatii* (Rabenh.) Grun. in A. Schmidt 1881
CM018A	*Cymbella gracilis* (Rabenh.) Cleve 1894
CM031A	*Cymbella minuta* Hilse ex Rabenh. 1862
CM033A	*Cymbella hustedtii* Krasske 1923
CM052A	*Cymbella descripta* (Hust.) Krammer & Lange-Bertalot 1985
CO001A	*Cocconeis placentula* Ehrenb. 1838
CO005A	*Cocconeis pediculus* Ehrenb. 1838
CO067A	*Cocconeis neothumensis* Krammer 1991
CY002A	*Cyclotella pseudostelligera* Hust. 1939
CY007A	*Cyclotella glomerata* Bachm. 1911
CY009A	*Cyclotella ocellata* Pant. 1902
CY010A	*Cyclotella comensis* Grun. in Van Heurck 1882
CY011A	*Cyclotella atomus* Hust. 1937
CY019A	*Cyclotella radiosa* (Grunow) Lemmerman 1900
CY9987	*Cyclotella* [cf. *comensis*]
CY9991	*Cyclotella kuetzingiana* agg.
DE001A	*Denticula tenuis* Kutz. 1844
DE002A	*Denticula elegans* Kutz. 1844
DP067A	*Diploneis modica* Hustedt 1945
DT002A	*Diatoma hyemale* (Roth) Heib. 1863
EU002B	*Eunotia pectinalis* var. *minor* (Kutz.) Rabenh. 1864
EU002D	*Eunotia pectinalis* var. *undulata* (Ralfs) Rabenh. 1864
EU002E	*Eunotia pectinalis* var. *minor* fo. *impressa* (Ehr.) Hust.
EU009A	*Eunotia exigua* (Breb. ex Kutz.) Rabenh. 1864
EU010A	*Eunotia faba* (Ehrenb.) Grun. in Van Heurck 1881
EU013A	*Eunotia arcus* Ehrenb. 1837
EU016A	*Eunotia diodon* Ehrenb. 1837
EU032A	*Eunotia serra* Ehrenb. 1837
EU047A	*Eunotia incisa* W. Sm. ex Greg. 1854

EU048A	*Eunotia naegelii*	Migula 1907
EU105A	*Eunotia subarcuatoides*	Alles, Norpel, Lange-Bertalot 1991
FR001A	*Fragilaria pinnata*	Ehrenb. 1843
FR002A	*Fragilaria construens*	(Ehrenb.) Grun. 1862
FR002B	*Fragilaria construens* var. *binodis*	(Ehrenb.) Grun. 1862
FR002C	*Fragilaria construens* var. *venter*	(Ehrenb.) Grun. in Van Heurck 1881
FR002D	*Fragilaria construens* var. *exigua*	(W. Sm.) Schulz 1922
FR005D	*Fragilaria virescens* var. *exigua*	Grun. in Van Heurck 1881
FR006A	*Fragilaria brevistriata*	Grun. in Van Heurck 1885
FR009A	*Fragilaria capucina*	Desm. 1825
FR009J	*Fragilaria capucina* var. *perminuta*	(Grun.) L-B. 1991
FR014A	*Fragilaria leptostauron*	(Ehrenb.) Hust. 1931
FR015A	*Fragilaria lata*	(Cleve-Euler) Renberg 1977
FR018A	*Fragilaria elliptica*	Schum. 1867
FR019A	*Fragilaria intermedia*	Grun. in Van Heurck 1881
FR063A	*Fragilaria robusta*	(Fusey) Manguin
FU002A	*Frustulia rhomboides*	(Ehrenb.) De Toni 1891
FU002B	*Frustulia rhomboides* var. *saxonica*	(Rabenh.) De Toni 1891
FU002F	*Frustulia rhomboides* var. *viridula*	(Breb. ex Kutz.) Cleve 1894
GO010A	*Gomphonema constrictum*	Ehrenb. ex Kutz.1844
GO014A	*Gomphonema intricatum*	Kutz. 1844
GO029A	*Gomphonema clavatum*	Ehr.
GO050A	*Gomphonema minutum*	(Ag.) Ag. 1831
GY001A	*Gyrosigma attenuatum*	(Kutz.) Rabenh. 1853
GY005A	*Gyrosigma acuminatum*	(Kutz.) Rabenh. 1853
HN001A	*Hannaea arcus*	(Ehrenb.) Patr. in Patr. & Reimer 1966
MA001A	*Mastogloia smithii*	Thwaites ex W. Sm. 1856
NA002A	*Navicula jaernefeltii*	Hust. 1942
NA003A	*Navicula radiosa*	Kutz. 1844
NA009A	*Navicula lanceolata*	(Agardh) Kutz.
NA013A	*Navicula pseudoscutiformis*	Hust. 1930
NA016A	*Navicula indifferens*	Hust. 1942
NA028A	*Navicula scutelloides*	W. Sm. ex Greg. 1856
NA030A	*Navicula menisculus*	Schum. 1867
NA033A	*Navicula subtilissima*	Cleve 1891
NA042A	*Navicula minima*	Grun. in Van Heurck 1880
NA044A	*Navicula krasskei*	Hust. 1930
NA055A	*Navicula graciloides*	A. Mayer 1919
NA066A	*Navicula capitata*	Ehrenb. 1838
NA068A	*Navicula impexa*	Hust. 1961
NA079A	*Navicula pseudolanceolata*	Lange-Bertalot 1980
NA084A	*Navicula atomus*	(Kutz.) Grun. 1860
NA099A	*Navicula bremensis*	Hust. 1957
NA121A	*Navicula begeri*	Krasske 1932
NA123A	*Navicula modica*	Hust. 1945
NA144A	*Navicula utermoehlii*	Hust. 1943
NA168A	*Navicula vitabunda*	Hust. 1930
NI009A	*Nitzschia palea*	(Kutz.) W. Sm. 1856
NI014A	*Nitzschia amphibia*	Grun. 1862
NI015A	*Nitzschia dissipata*	(Kutz.) Grun. 1862
NI042A	*Nitzschia acicularis*	(Kutz.) W. Sm. 1853
OP001A	*Opephora martyi*	Herib.1902
PE002A	*Peronia fibula*	(Breb. ex Kutz.) R. Ross 1956
PI004A	*Pinnularia interrupta*	W. Smith
PI011A	*Pinnularia microstauron*	(Ehrenb.) Cleve 1891
PI015A	*Pinnularia abaujensis*	(Pant.) R. Ross in Hartley 1986
PI016A	*Pinnularia divergentissima*	(Grun.in Van Heurck) Cleve 1896

SA001B	*Stauroneis anceps* var. *gracilis* Rabenh. 1864
SA006A	*Stauroneis phoenicenteron* (Nitzsch) Ehrenb. 1943
ST001A	*Stephanodiscus hantzschii* Grun. in Cleve & Grun. 1880
ST002A	*Stephanodiscus tenuis* Hust. 1939
ST009A	*Stephanodiscus alpinus* Hust.
ST010A	*Stephanodiscus parvus* Stoermer & Hakansson 1984
ST022A	*Stephanodiscus neoastraea* Hakansson & Hickel 1986
SY002A	*Synedra rumpens* Kutz. 1844
SY003A	*Synedra acus* Kutz. 1844
SY004A	*Synedra parasitica* (W. Sm.) Hust. 1930
SY009A	*Synedra nana* Meister 1912
SY015A	*Synedra tabulata* (Ag.) Kutz. 1844
TA002A	*Tabellaria fenestrata* (Lyngb.) Kutz. 1844
TA9996	*Tabellaria flocculosa* agg.