Biogeosciences Discuss., 11, 15889–15909, 2014
www.biogeosciences-discuss.net/11/15889/2014/
doi:10.5194/bgd-11-15889-2014

Biogeosciences
Discussions
Open Access

Environmental factors controlling lake diatom communities: a meta-analysis of published data

S. Blanco[1,*]

[1]Department of Biodiversity and Environmental Management, University of León, 24071 León, Spain
[*]now at: Institute of the Environment, La Serna, 58, 24007 León, Spain

Received: 17 September 2014 – Accepted: 24 October 2014 – Published: 18 November 2014

Correspondence to: S. Blanco (sblal@unileon.es)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Diatoms play a key role in the development of quantitative methods for environmental reconstruction in lake ecosystems. Diatom-based calibration datasets developed during the last decades allow the inference of past limnological variables such as TP, pH or conductivity and provide information on the autecology and distribution of diatom taxa. However, little is known about the relationships between diatoms and climatic or geographic factors. The response of surface sediment diatom assemblages to abiotic factors is usually examined using canonical correspondence analysis (CCA) and subsequent forward selection of variables based on Monte Carlo permutation tests that show the set of predictors best explaining the distributions of diatom species. The results reported in 40 previous studies using this methodology in different regions of the world are re-analyzed in this paper. Bi- and multivariate statistics (canonical correlation and two-block partial least-squares) were used to explore the correspondence between physical, chemical and physiographical factors and the variables that explain most of the variance in the diatom datasets. Results show that diatom communities respond mainly to chemical variables (pH, nutrients) with lake depth being the most important physiographical factor. However, the relative importance of certain parameters varied along latitudinal and trophic gradients. Canonical analyses demonstrated a strong concordance with regard to the predictor variables and the amount of variance they captured, suggesting that, on a broad scale, lake diatoms give a robust indication of past and present environmental conditions.

1 Introduction

Paleolimnological techniques have been recognized as one of the most powerful methods in the reconstruction of past conditions for continental waters. Lake sediments faithfully record changes that have occurred both within a lake and within its watershed (Chen et al., 2008; Edulnd and Ramstak, 2006). Diatom-based predictive models, using weighted-averaging (WA), partial least squares (PLS) and/or weighted-averaging partial least squares (WA-PLS) regression and calibration, are routinely developed for inferences of water chemistry worldwide. These models are based on the premise that diatoms present in the top sediments represent present-day lake conditions, while bottom samples generally represent pre-industrial conditions (Korsman and Birks, 1996; Dixit et al., 2002). Transfer functions are based on the relationship between diatom assemblages in surface sediments and the environmental variable of interest in a control set of lakes spanning a wide environmental gradient (Kirilova et al., 2008). This allows for reconstruction of past environmental conditions, which can be applied to investigations of recent climate change and human impacts (Saunders et al., 2008). As a first step, autecological parameters of the diatom species must be determined by relating modern limnological variables to surface sediment diatom assemblages (Reavie et al., 1995; Reid, 2005). Currently, a large number of diatom taxa have well-defined optima and tolerances (Beyene et al., 2014).

The multiple factors influencing diatom distributions often make it difficult to obtain strong calibrations for particular limnological variables because interacting effects influence changes in diatom community composition over time. For instance, the role of direct and indirect effects of environmental and climatic factors on diatoms is much debated (Kirilova et al., 2008; Shinneman et al., 2010). Diatom species distributions are expected to display latitudinal gradients, reflecting the crucial role of diatoms in controlling mixing and temperature regimes in lakes (Reid, 2005; Vyverman et al., 2007; Blanco et al., 2014). Although lakes sharing climatic characteristics within a given physiographic region are predicted to change in similar ways, limnological response to external forcing is often spatially complex and depends on local catchment-scale characteristics (Ryves et al., 2004; Perren et al., 2009). A number of studies have used assumptions about the distribution of diatom taxa based on ecological preferences inferred from different world regions, leading to speculative interpretations (Shinneman et al., 2009). This practice is particularly inadvisable with respect to variables related to trophic status (Denys, 2007). Finally, often-neglected biotic interactions may also become important regulators of lake diatom communities (Bennion, 1994).

Many paleolimnological studies focus on identifying a set of environmental variables that contribute most to determining the structure of diatom assemblages. Ordination in reduced space allows the inference of quantitative information on the relationship among taxa, environmental variables and sampling sites. Ordination methods, including direct and indirect gradient analyses, are generally used to reveal the relationship between diatom and environmental stressors (Yang et al., 2001). In particular, Canonical Correspondence Analysis (CCA) is by far the most common method used in the literature for studies involving many response and predictor variables (Birks and Birks, 1998). Taxa often exhibit unimodal responses to environmental gradients (Birks and Birks, 1998), and CCA assumes a unimodal response and is considered theoretically and ecologically sounder, as well as statistically more robust, than linear-based techniques (Pienitz et al., 1995). CCA assesses the simultaneous influence of multiple environmental variables on diatom assemblages (Dixit and Smol, 1994), identifying a subset of variables that explain statistically significant variation in the diatom species distributions (Ng and King, 1999). The subset is typically selected based on the results of distribution-free Monte Carlo permutation tests. In many cases, an automatic forward selection procedure is used to objectively select a subset of explanatory variables; this method is applicable even when the initial dataset contains more explanatory variables than sampling sites, which is often the case in ecology (Blanchet et al., 2008).

Despite the widespread use of such multivariate methods, little effort has been made to determine the relative importance of different predictors over a large geographical scale. Quantitative models for a many variables are demonstrated to have limited spatial transferability (Álvarez-Blanco et al., 2011; Juggins, 2013). This study explores the relative impact of key abiotic factors on diatom composition by jointly analyzing the results presented in different studies covering a broad range of limnological and climatic conditions. The primary goal is to assess how the amount of variance captured by each predictor changes across different environmental and geographical gradients.

2 Materials and methods

A total of 40 studies on the response of surface sediment diatom assemblages to abiotic factors, which used the CCA-Monte Carlo method, were collected from recent literature using Google Scholar. These included primarily scientific papers and dissertations and reports. The following information was extracted from each reference: (a) the sample size, (b) the geographic coordinates (WGS84) of the study sites (in the case of several sites, the average centroid), (c) the main environmental descriptors (physical, chemical and morphometric) and (d) the variance captured by each of these predictors in the response (diatom) dataset. In most studies, the data had been previously screened to remove (a) environmental variables that had little or no influence on diatom distributions, (b) samples with particularly unusual diatom assemblages and/or environmental characteristics and (c) redundancies in environmental variables (Reavie et al., 1995; Yang et al., 2005). Variables measured in less than 10 studies were excluded from subsequent analyses, so that 18 variables were finally considered (Table 1). These data were arranged in two variable/site (18 × 40) matrices: one with environmental data (hereafter "predictor" matrix), and another with the percentage of variance explained by the corresponding variable in each study (hereafter "criterion" matrix). Missing data were substituted with column medians. Most variables did not follow normal distributions (Kolmogorov–Smirnov test, $p < 0.05$); therefore, the data were Box–Cox transformed to satisfy the assumptions of parametric statistics (Legendre and Legendre, 2012).

The statistical relationship between both matrices was explored using three complementary techniques: (a) simple bivariate Spearman rank correlations, (b) canonical correlation analysis (CCorA) and (c) two-block partial least squares (2BPLS). The latter two methods allow for assessment of the relationship between two sets of variables. CCorA is a multivariate generalization of multiple regression analysis that calculates overall correlations between the variable sets by first extracting the canonical factors from both groups and then finding linear combinations such that the correlation between the variables is maximized (Gould et al., 1986; Muma et al., 2011). A detailed description of this technique is available in Thompson (Thompson, 1984). In contrast, 2BPLS, as described in Rohlf and Corti (2000) is used to determine the combinations of variables in the two matrices that account for most of the covariation present between the variable sets. The method is based on the singular value decomposition of the correlation matrix between the variable matrices which are treated symmetrically. These statistical analyses were conducted using Statistica v. 10.0 (StatSoft, 2011) and PAST v. 3.01 (Harper et al., 2001) software, respectively.

3 Results

The number of samples analyzed in each study varied from 1 to 186, with a median value of 53 and a total of 2490 samples considered in the present meta-analysis. Table 1 presents the descriptive statistics of the predictor and criterion matrices. There is a strong bias in the distribution of lakes analyzed; only 15 % of the papers collected studied Southern Hemisphere lakes. Moreover, 63 % of the studies involved systems situated above parallel 45° N/below parallel 45° S. No reports were found concerning African or Antarctic lakes, and the Neotropical ecozone is represented by a single study (Hassan et al., 2009).

The canonical correlation coefficient between the weighted sums of scores in each matrix was statistically significant ($R = 0.99$, $p = 0.02$ for the first and only significant canonical root). 87.18 % of the variance was extracted from the predictor variables which, on average, accounted for 48.75 % of the variance in the variables of the criterion matrix. Three variables in the criterion set – alkalinity, dissolved oxygen and total nitrogen – correlated highly with this first canonical root (squared canonical factor loadings: 0.54, 0.53 and 0.38, respectively, Fig. 1). With certain exceptions (TP, pH), the direction and relative magnitude of canonical weights were in agreement with those of the canonical loadings (Fig. 1). With respect to 2BPLS results, the first dimension accounted for 40 % of the total covariance between both sets, and variable loadings followed the same trends as CCorA canonical loadings (Fig. 1). Figure 2 presents a joint plot of CCorA and 2BPLS scores, showing for both analyses a linear relationship between predictor and criterion sets and an unbiased distribution of residuals, thus confirming that no major violations of canonical correlation assumptions were present. However, points with scores above two units in both axes could be considered as outliers because they fall outside the 95 % confidence ellipses. A re-examination of the original data confirmed that these points correspond to deep lakes (Zm > 18 m): Sacrower See (Kirilova et al., 2008; Kirilova, 2009) and Lake Kenyir (Rouf et al., 2008), where the variables with high loadings (see above) captured unusually large amounts (> 40 %) of variance in their respective datasets. In this regard, it was observed that the amount of variance explained by alkalinity and dissolved oxygen was positively correlated to lake depth ($\rho = 0.88$ and 0.84, respectively, $p < 0.05$).

pH was selected as a significant explanatory variable in 78 % of the analyzed studies and was the variable that explained most of the variance in the diatom dataset in 29 % of the papers (Table 1). Other important limnological variables included conductivity and total phosphorus and nitrogen concentrations. Mean depth was the only relevant physiographical variable, selected in 44 % of the studies. However, in terms of amount of variance explained, lake area captured a larger proportion than any other forward-selected environmental variable, with a median of 8.10 % of variance explained throughout the studied systems.

The importance of certain predictors varied geographically. The percentage of variance captured by elevation and water temperature was significantly higher ($\rho = 0.86$ and 1.00, respectively, $p < 0.01$, Fig. 3) in high-latitude lakes, whereas pH showed a unimodal response to latitude, peaking at 53° N (Abitibi clay belt area, Canada, Enache and Prairie, 2002). No general trends were observed with respect to total phosphorus (Fig. 4), which reached maximum percentages in NW and NE North America. The relevance of certain variables (pH, silica) decreased along the trophic gradient ($\rho = 0.42$ and 0.95, respectively, $p < 0.05$, Fig. 5).

4 Discussion

Aquatic ecosystem management requires the examination of long-term data to assess past conditions, natural variability and the response of aquatic systems to natural or anthropogenic disturbances (Enache and Prairie, 2002). This approach has largely developed since the creation of modern datasets of biological assemblages and associated environmental data, allowing the quantitative estimation of the past environment (Birks and Birks, 1998). Many studies demonstrate that diatoms can provide valuable paleoenvironment and paleoclimate proxy data because they are sensitive indicators of past lake conditions that are at least indirectly related to climate. Different regression and calibration techniques that relate diatoms in surface sediments to measured lake-water characteristics have been developed to provide quantitative environmental reconstructions required in such studies (Pienitz et al., 1995; Korsman and Birks, 1996). Diatoms have rapid immigration and replication rates, which allow them to respond quickly to changes in the aquatic environment. In many studies, surface sediment diatoms and annual mean water quality data are preferred because the surface sediment record is composed of a number of different ecotypes and was deposited over a long period of time (Dixit and Smol, 1994; Yang et al., 2005).

The analysis presented here indicates that, in general, both physical environmental conditions and lake-water chemistry show strong statistical relationships with diatom community composition. Particularly, it provides evidence of a strong relationship between diatoms and pH – due to the direct physiological influence of pH on diatoms (Chen et al., 2008) – especially in oligotrophic systems (Fig. 5). Nutrients also capture important fractions of variability in diatom assemblages across a large range of trophic levels (Bennion, 1994; Reavie et al., 1995; Blanco et al., 2014). However, the relative importance of nutrients depends on supply rates and on the order in which they are consumed, so that limiting nutrient elements exhibits greater importance in diatom community development (Dong et al., 2012). Particularly, a poor performance of TP may reflect the coexistence of N and P limitation in the studied systems (Reid, 2005).

On the other hand, most diatom assemblages appear to be unresponsive to a number of factors related to trophic status (e.g., nitrates, ammonia). Two alternative explanations can be considered: (a) diatom assemblages from eutrophic lakes are generally more homogeneous (Chen et al., 2008), and the diatom–chemistry relationship is constrained in such low-diversity communities (Denys, 2007); or (b) individual data sets were either too small or did not encompass sufficiently large environmental gradients to accurately predict these variables (Antoniades et al., 2005).

CCorA and 2BPLS results are similar with respect to predictor variables and the amount of variance they explained in the diatom assemblages. Both of the multivariate canonical analyses performed for this study show alkalinity and dissolved oxygen as the variables with the highest loadings and weights (Fig. 1). The relatively low scores of other important predictors such as pH or nutrients may be due to the multicollinearity (Kuylen and Verhallen, 1981) detected in the predictor matrix; in most pairwise comparisons, pH, conductivity, TP and TN were mutually correlated. The observed differences between canonical factor loadings and weights for alkalinity, pH and elevation (Fig. 1) can indicate nonlinear relationships between variables. That is, the relationships among other variables are stronger in cases (studies) scoring high and low on these variables. Finally, TP acts as a "suppressor" variable (sensu Schaninger et al., 1980), removing part of the variance among other predictors uncorrelated with the canonical scores.

Lake depth appears as the most important physiographical variable in terms of average variance captured; this factor is related to the amount of variance explained by alkalinity and dissolved oxygen. According to Shinneman et al. (2010), a strong correlation between depth and alkalinity may lead to spurious results due to unrelated changes in one of these confounding variables. Other studies (Ryves et al., 2004; Tibby, 2004) also indicate that the conditional variance explained by TP, pH, and depth may not be independent. Furthermore, certain diatom species are found predominantly in shallow lakes (Shinneman et al., 2009) where the proportion of non-planktonic taxa will be comparatively larger so that derived species optima may differ from those in deeper water because of greater variability of environmental conditions and stronger interaction of physical or biological factors (Denys, 2006). In shallow lakes, climate can also obscure or exaggerate the nutrient enrichment process because the heat and energy is more easily transferred within the lake than in larger, deeper systems (Dong, 2010).

The analyzed dataset presented an uneven distribution of lakes along the latitudinal gradient, thus preventing statistically accurate interpretations of the results. This confirms that, with the exceptions of Europe and North America, the relationship between diatoms and environmental data has been poorly documented (Hassan et al., 2009). Rühland and Smol (2002) highlight the need to define environmental optima and tolerances of diatom taxa in new geographical regions to build robust and meaningful inference models. In general, although diatom assemblage composition is related to climatic patterns (Weckström and Korhola, 2001), the potential influence of geographical factors has been underestimated (Enache and Prairie, 2002; Yang et al., 2003; Álvarez-Blanco et al., 2011). Climate has a profound effect on diatom dynamics, affecting diatom communities via its links with nutrient availability (Kirilova et al., 2008). In fact, geographical features have been found to explain significant amounts of the species variance (Reid, 2005). Contrary to our observations (Fig. 3), it has been noted (Schmidt et al., 2004) that the percentage of diatom distribution explained by water temperature is relatively lower in high latitude lakes. The effect of such confounding gradients on model performance has received surprisingly little attention (Juggins, 2013), and the interaction between limnological and geographical factors on lake ecosystems cannot be resolved until long-term monitoring datasets that include ecological communities and environmental variables become available (Dong, 2010). Additionally, many diatom datasets covering broad geographical areas and ecological gradients have, by necessity, been constructed by combining samples analyzed by different analysts (Birks and Birks, 1998), thus introducing a source of bias in subsequent inferences.

By compiling published data at a pan-continental scale, it has been possible in this paper to confirm the major predictors configuring lake diatom assemblages worldwide. However, often-neglected factors such as stratification or silica cycling patterns

(Ryves et al., 2004) may also be important. Attempts at precise predictions of diatom responses to climatic variables would require complex models incorporating information on species interactions and dispersal (Weckström and Korhola, 2001). Data gained from monitoring surveys would also be a useful complement to paleolimnological records in understanding how diatom communities respond to environmental changes (Kirilova et al., 2008). Regardless, future attempts to perform comparable meta-analyses would benefit from larger sample sizes and increased variable gradient ranges.

References

Álvarez-Blanco, I., Cejudo-Figueiras, C., Bécares, E., and Blanco, S.: Spatiotemporal changes in diatom ecological profiles: implications for biomonitoring, Limnology, 12, 157–168, doi:10.1007/s10201-010-0333-1, 2011.

Antoniades, D., Douglas, M. S. V., and Smol, J. P.: Benthic diatom autecology and inference model development from the Canadian High Arctic Archipelago, J. Phycol., 41, 30–45, doi:10.1111/j.1529-8817.2005.04049.x, 2005.

Bennion, H.: A diatom-phosphorus transfer function for shallow, eutrophic ponds in southeast England, Hydrobiologia, 275–276, 391–410, doi:10.1007/BF00026729, 1994.

Beyene, A., Awoke, A., and Triest, L.: Estimation of environmental optima and tolerances of diatoms using multifactor multiplicative modeling, Ecol. Inform., 19, 53–61, doi:10.1016/j.ecoinf.2013.12.007, 2014.

Birks, H. J. B. and Birks, H. J. B.: D. G. Frey and E. S. Deevey Review 1: Numerical tools in palaeolimnology – progress, potentialities, and problems, J. Paleolimnol., 20, 307–332, doi:10.1023/A:1008038808690, 1998.

Blanchet, F. G., Legendre, P., and Borcard, D.: Forward selection of explanatory variables, Ecology, 89, 2623–2632, doi:10.1890/07-0986.1, 2008.

Blanco, S., Cejudo-Figueiras, C., Álvarez-Blanco, I., van Donk, E., Gross, E. M., Hansson, L.-A., Irvine, K., Jeppesen, E., Kairesalo, T., Moss, B., Nõges, T., and Bécares, E.: Epiphytic diatoms along environmental gradients in Western European shallow lakes, Clean-Soil Air Water, 42, 229–235, 2014.

Chen, G., Dalton, C., Leira, M., and Taylor, D.: Diatom-based total phosphorus (TP) and pH transfer functions for the Irish Ecoregion, J. Paleolimnol., 40, 143–163, doi:10.1007/s10933-007-9148-4, 2008.

Denys, L.: Calibration of littoral diatoms to water chemistry in standing fresh waters (Flanders, Lower Belgium): inference models for historical sediment assemblages, J. Paleolimnol., 35, 763–787, doi:10.1007/s10933-005-5459-5, 2006.

Denys, L.: Water-chemistry transfer functions for epiphytic diatoms in standing freshwaters and a comparison with models based on littoral sediment assemblages (Flanders, Belgium), J. Paleolimnol., 38, 97–116, doi:10.1007/s10933-006-9064-z, 2007.

Dixit, S. S. and Smol, J. P.: Diatoms as indicators in the Environmental Monitoring and Assessment Program-Surface Waters (EMAP-SW), Environ. Monit. Assess., 31, 275–307, doi:10.1007/BF00577258, 1994.

Dixit, S. S., Dixit, A. S., and Smol, J. P.: Diatom and chrysophyte functions and inferences of post-industrial acidification and recent recovery trends in Killarney lakes (Ontario, Canada), J. Paleolimnol., 27, 79–96, doi:10.1023/A:1013571821476, 2002.

Dong, X.: Using Diatoms to Understand Climate–Nutrient Interactions in Esthwaite Water, England: Evidence from Observational and Palaeolimnological Records, Ph.D. thesis, UCL (University College London), 28 December, available at: http://discovery.ucl.ac.uk/763095/ (last access: 12 February 2014), 2010.

Dong, X., Bennion, H., Maberly, S. C., Sayer, C. D., Simpson, G. L., and Battarbee, R. W.: Nutrients exert a stronger control than climate on recent diatom communities in Esthwaite Water: evidence from monitoring and palaeolimnological records, Freshwater Biol., 57, 2044–2056, doi:10.1111/j.1365-2427.2011.02670.x, 2012.

Edulnd, M. and Ramstak, J.: Diatom-Inferred TP in MCWD Lakes, St. Croix Watershed Research – Station Science Museum of Minnesota, available at: http://www.minnehahacreek.org/sites/minnehahacreek.org/files/attachments/MCWDFinalRpt06_07.pdf, 2006.

Enache, M. and Prairie, Y. T.: WA-PLS diatom-based pH, TP and DOC inference models from 42 lakes in the Abitibi clay belt area (Québec, Canada), J. Paleolimnol., 27, 151–171, doi:10.1023/A:1014281128544, 2002.

Gould Jr., R. W., Balmori, E. R., and Fryxell, G. A.: Multivariate statistics applied to phytoplankton data from two Gulf Stream warm core rings, Limnol. Oceanogr., 315, 951–968, 1986.

Harper, D. and Ryan, P.: PAST: paleontological statistics software package for education and data analysis, Palaeontol. Electron., 4, 1–9, 2001.

Hassan, G. S., Tietze, E., and Francesco, C. G. D.: Modern diatom assemblages in surface sediments from shallow lakes and streams in southern Pampas (Argentina), Aquat. Sci., 71, 487–499, doi:10.1007/s00027-009-0104-4, 2009.

Juggins, S.: Quantitative reconstructions in palaeolimnology: new paradigm or sick science?, Quaternary Sci. Rev., 64, 20–32, doi:10.1016/j.quascirev.2012.12.014, 2013.

Kirilova, E. P.: Natural and Human Induced Trophic Changes in European Lowland Lakes, LPP Contrib. Ser., 26, available at: http://dspace.library.uu.nl/handle/1874/33849 (last access: 29 January 2014), 2009.

Kirilova, E. P., Bluszcz, P., Heiri, O., Cremer, H., Ohlendorf, C., Lotter, A. F., and Zolitschka, B.: Seasonal and interannual dynamics of diatom assemblages in Sacrower See (NE Germany): a sediment trap study, Hydrobiologia, 614, 159–170, doi:10.1007/s10750-008-9504-z, 2008.

Korsman, T. and Birks, H. J. B.: Diatom-based water chemistry reconstructions from northern Sweden: a comparison of reconstruction techniques, J. Paleolimnol., 15, 65–77, doi:10.1007/BF00176990, 1996.

Kuylen, A. A. A. and Verhallen, T. M. M.: The use of canonical analysis, J. Econ. Psychol., 1, 217–237, doi:10.1016/0167-4870(81)90039-8, 1981.

Legendre, P. and Legendre, L.: Numerical Ecology, Elsevier, 2012.

Muma, M., Assani, A. A., Landry, R., Quessy, J.-F., and Mesfioui, M.: Effects of the change from forest to agriculture land use on the spatial variability of summer extreme daily flow characteristics in southern Quebec (Canada), J. Hydrol., 407, 153–163, doi:10.1016/j.jhydrol.2011.07.020, 2011.

Ng, S. L. and King, R. H.: Development of a diatom-based specific conductivity model for the glacio-isostatic lakes of Truelove Lowland: implications for paleoconductivity and paleoenvironmental reconstructions in Devon Island lakes, NWT, Canada, J. Paleolimnol., 22, 367–382, doi:10.1023/A:1008022501117, 1999.

Perren, B. B., Douglas, M. S. V., and Anderson, N. J.: Diatoms reveal complex spatial and temporal patterns of recent limnological change in West Greenland, J. Paleolimnol., 42, 233–247, doi:10.1007/s10933-008-9273-8, 2009.

Pienitz, R., Smol, J. P., and Birks, H. J. B.: Assessment of freshwater diatoms as quantitative indicators of past climatic change in the Yukon and Northwest Territories, Canada, J. Paleolimnol., 13, 21–49, doi:10.1007/BF00678109, 1995.

Reavie, E. D., Hall, R. I., and Smol, J. P.: An expanded weighted-averaging model for inferring past total phosphorus concentrations from diatom assemblages in eutrophic British Columbia (Canada) lakes, J. Paleolimnol., 14, 49–67, doi:10.1007/BF00682593, 1995.

Reid, M.: Diatom-based models for reconstructing past water quality and productivity in New Zealand lakes, J. Paleolimnol., 33, 13–38, doi:10.1007/s10933-004-0424-2, 2005.

Rohlf, F. J. and Corti, M.: Use of two-block partial least-squares to study covariation in shape, Syst. Biol., 49, 740–753, 2000.

Rouf, A. A., Ambak, M. A., Shamsudin, L., Phang, S.-M., and Ho, S. C.: Temporal changes in the periphytic algal communities in a drowned tropical forest reservoir in Malaysia: Lake Kenyir, Lakes Reserv. Res. Manag., 13, 271–287, doi:10.1111/j.1440-1770.2008.00382.x, 2008.

Rühland, K. M. and Smol, J. P.: Freshwater diatoms from the Canadian Arctic treeline and development of paleolimnological inference models, J. Phycol., 38, 249–264, doi:10.1046/j.1529-8817.2002.01129.x, 2002.

Ryves, D. B., Clarke, A. L., Appleby, P. G., Amsinck, S. L., Jeppesen, E., Landkildehus, F., and Anderson, N. J.: Reconstructing the salinity and environment of the Limfjord and Vejlerne Nature Reserve, Denmark, using a diatom model for brackish lakes and fjords, Can. J. Fish. Aquat. Sci., 61, 1988–2006, doi:10.1139/f04-127, 2004.

Schaninger, C. M., Parker, V., and Panton, D. B.: The complementary use of multivariate procedures to investigate nonlinear and interactive relationships between personality and product usage, J. Marketing Res., 17, 119–124, doi:10.2307/3151127, 1980.

Schmidt, R., Kamenik, C., Kaiblinger, C., and Hetzel, M.: Tracking Holocene environmental changes in an alpine lake sediment core: application of regional diatom calibration, geochemistry, and pollen, J. Paleolimnol., 32, 177–196, doi:10.1023/B:JOPL.0000029428.97961.43, 2004.

Shinneman, A. L. C., Edlund, M. B., Almendinger, J. E., and Soninkhishig, N.: Diatoms as indicators of water quality in Western Mongolian lakes: a 54-site calibration set, J. Paleolimnol., 42, 373–389, doi:10.1007/s10933-008-9282-7, 2009.

Shinneman, A. L. C., Bennett, D. M., Fritz, S. C., Schmieder, J., Engstrom, D. R., Efting, A., and Holz, J.: Inferring lake depth using diatom assemblages in the shallow, seasonally variable lakes of the Nebraska Sand Hills (USA): calibration, validation, and application of a 69-lake training set, J. Paleolimnol., 44, 443–464, doi:10.1007/s10933-010-9427-3, 2010.

StatSoft Inc.: STATISTICA, ver. 10., 2011.

Thompson, B.: Canonical Correlation Analysis: Uses and Interpretation, SAGE, 1984.

Tibby, J.: Development of a diatom-based model for inferring total phosphorus in southeastern Australian water storages, J. Paleolimnol., 31, 23–36, doi:10.1023/B:JOPL.0000013272.25122.2a, 2004.

Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D. A., Mann, D. G., Juggins, S., Van de Vijver, B., Jones, V., Flower, R., Roberts, D., Chepurnov, V. A., Kilroy, C., Vanormelingen, P., and De Wever, A.: Historical processes constrain patterns in global diatom diversity, Ecology, 88, 1924–1931, 2007.

Weckström, J. and Korhola, A.: Patterns in the distribution, composition and diversity of diatom assemblages in relation to ecoclimatic factors in Arctic Lapland, J. Biogeogr., 28, 31–45, doi:10.1046/j.1365-2699.2001.00537.x, 2001.

Yang, X., Wang, S., Xia, W., and Li, W.: Application of CCA for study on modern lake diatoms and environment in the Tibetan Plateau, Sci. China Ser. D, 44, 343–350, doi:10.1007/BF02912005, 2001.

Yang, X., Kamenik, C., Schmidt, R., and Wang, S.: Diatom-based conductivity and water-level inference models from eastern Tibetan (Qinghai-Xizang) Plateau lakes, J. Paleolimnol., 30, 1–19, doi:10.1023/A:1024703012475, 2003.

Yang, X.-D., Dong, X.-H., Gao, G., Pan, H.-X., and Wu, J.-L.: Relationship between surface sediment diatoms and summer water quality in shallow lakes of the middle and lower reaches of the Yangtze River, J. Integr. Plant Biol., 47, 153–164, doi:10.1111/j.1744-7909.2005.00035.x, 2005.

Table 1. Descriptive statistics of the analyzed dataset ($n = 40$). Variable medians and ranks, and percentage of variance explained in the diatom assemblages.

	Median [rank]	% of variance explained (median [rank])	[a]	[b]
Latitude (absolute)	51 [5–76]			
Longitude (E)	8 [(−159)–173]			
Elevation (m a.s.l.)	258 [0–4290]	5.25 [1.00–13.8]	19.51	2.44
Mean depth (m)	7.10 [0.39–51.05]	5.1 [1.20–31.00]	43.90	7.32
Area (km^2)	0.23 [0.01–369.00]	8.10 [1.50–25.00]	12.20	2.44
pH	7.75 [5.80–9.11]	7.17 [0.00–40.00]	78.05	29.27
Alkalinity (mEq L^{-1})	7.40 [0.08–3471.78]	0.00 [0.00–17.00]	17.07	9.76
DOC[c] (mg L^{-1})	7.57 [0.50–22.80]	5.10 [0.00–16.00]	14.63	0.00
Conductivity (µS cm^{-1})	234.73 [10.12–88 068.51]	3.48 [0.00–37.00]	39.02	24.39
Turbidity (NTU)	7.08 [1.40–121.60]	1.21 [0.00–3.58]	12.20	0.00
Water T (°C)	13.39 [5.90–29.64]	0.61 [0.00–18.00]	26.83	0.00
O_2 (mg L^{-1})	10.84 [7.66–90.70]	1.90 [0.00–32.00]	17.07	2.44
TN[d] (mg L^{-1})	0.89 [0.22–3.44]	2.97 [0.00–13.20]	36.59	4.88
Ammonia (mg L^{-1})	0.03 [0.00–38.80]	0.00 [0.00–7.20]	7.32	0.00
Nitrates (mg L^{-1})	0.05 [0.00–30.00]	0.00 [0.00–10.10]	9.76	0.00
TP[e] (mg L^{-1})	0.06 [0.00–0.36]	4.70 [0.00–23.00]	43.90	12.20
SRP[f] (mg L^{-1})	0.03 [0.00–2.60]	3.40 [0.00–21.90]	19.51	4.88
Silica (mg L^{-1})	1.93 [0.23–46.72]	2.20 [0.00–22.00]	17.07	0.00

[a] Percentage of papers where the variable was forward-selected (CCA-Monte Carlo test).

[b] Percentage of papers where the variable accounted for the most variance explained.

[c] Dissolved organic carbon.

[d] Total nitrogen.

[e] Total phosphorus.

[f] Soluble reactive phosphorus.

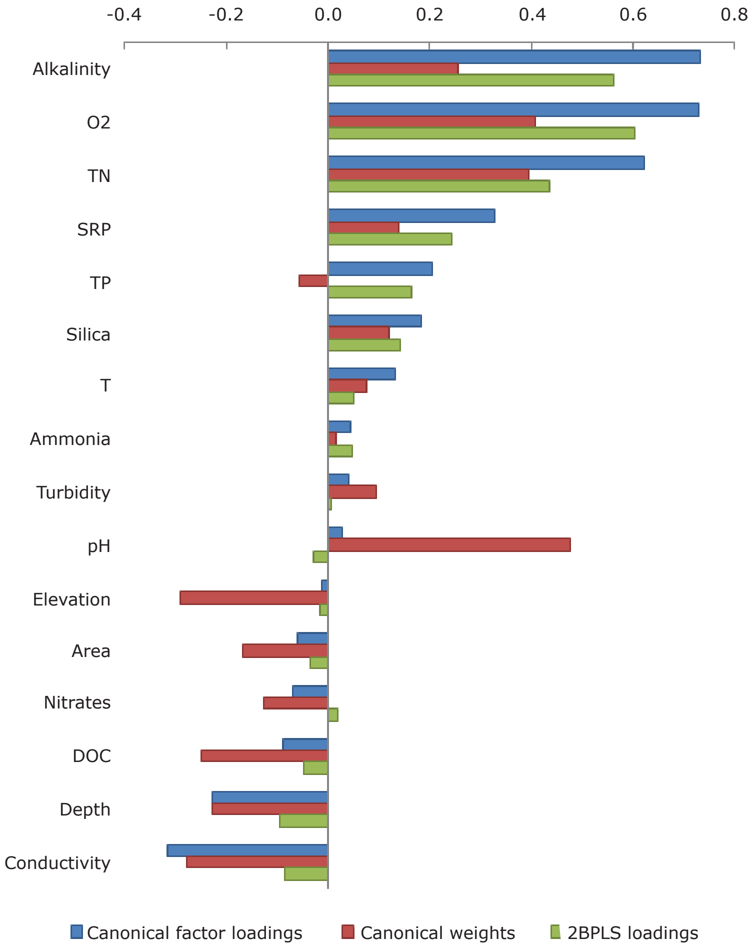

Figure 1. Factor loadings and weights resulting from the CCorA and 2BPLS analyses for variables in the criterion matrix.

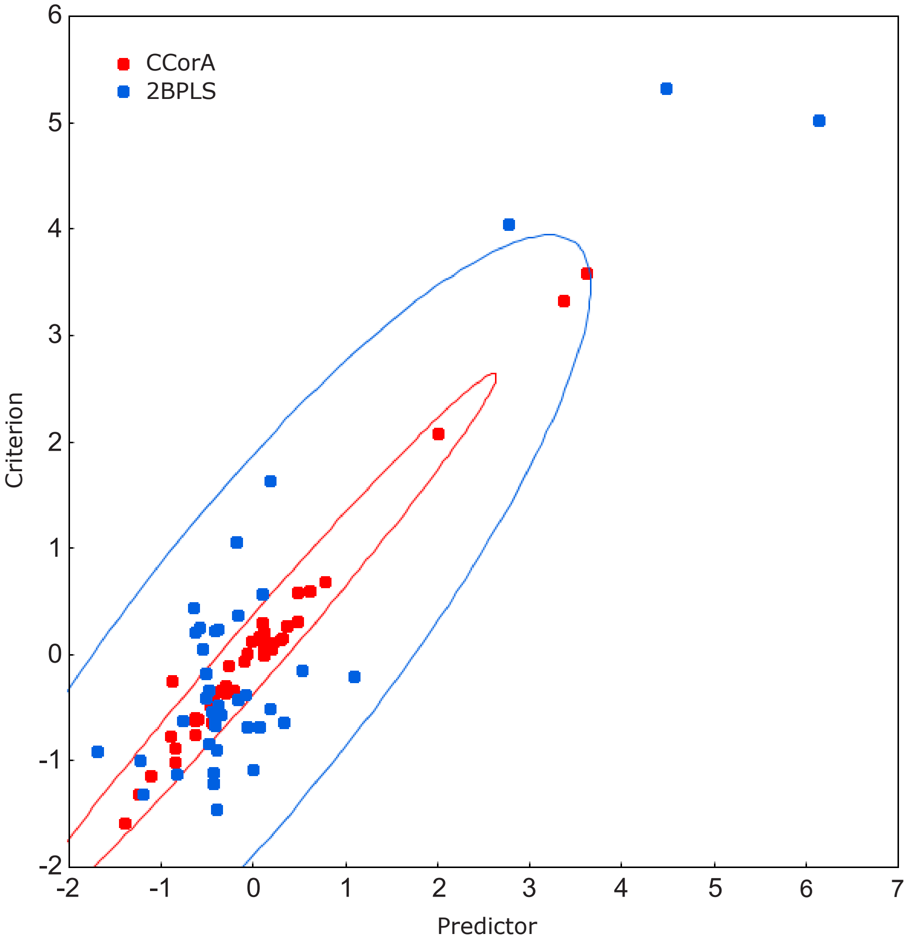

Figure 2. Canonical and 2BPLS scores for variables in the predictor and criterion sets (first canonical root/dimension). Data fitted to 95 % confidence ellipses.

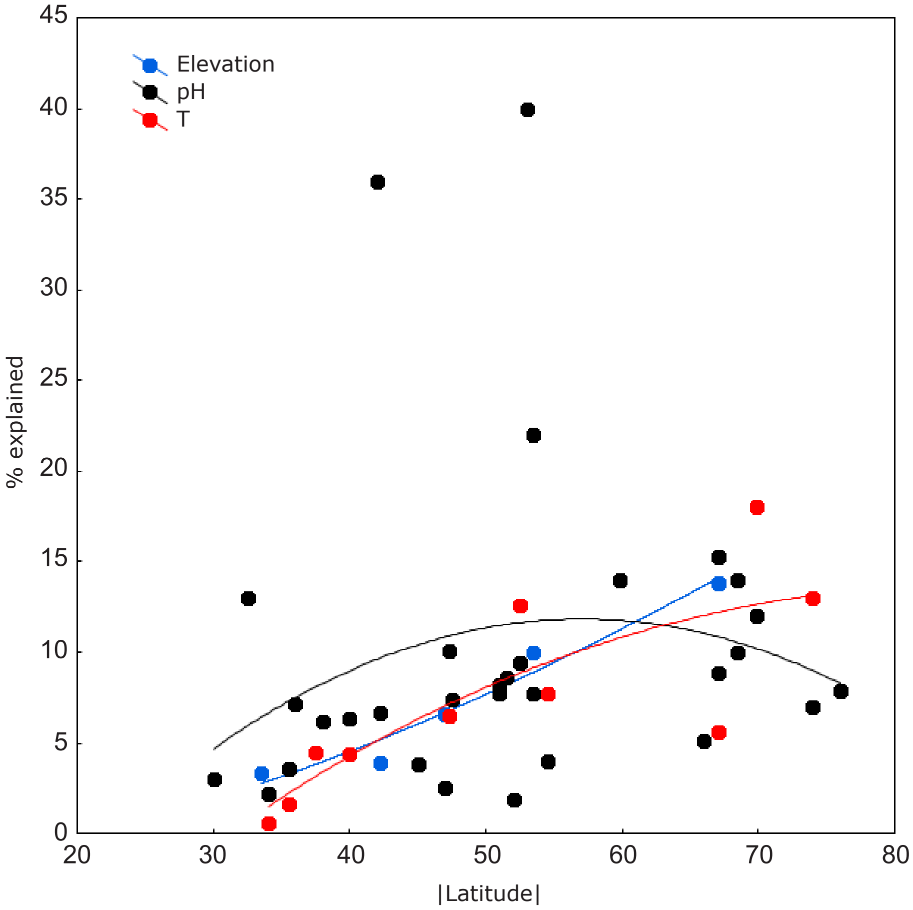

Figure 3. Percentage of variance captured by elevation, water temperature and pH along the latitudinal gradient. Data fitted to distance-weighted least squares curves.

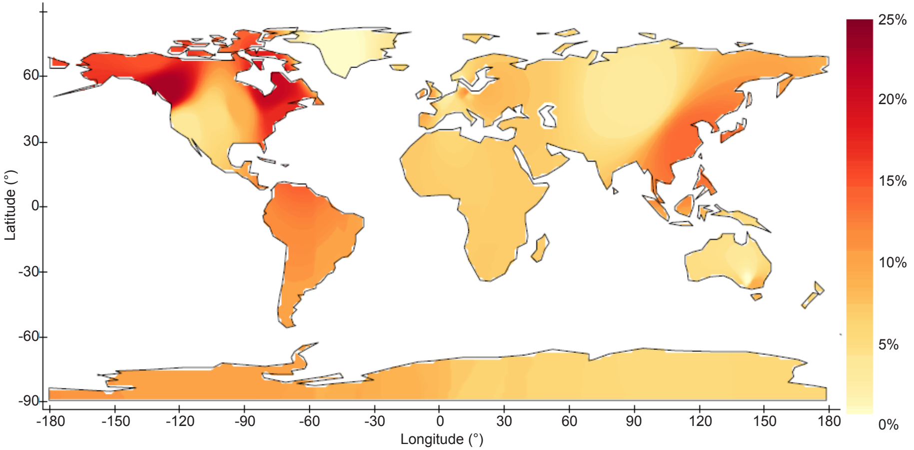

Figure 4. Amount of variance explained by the concentration of total phosphorus in lake diatom assemblages worldwide. Data were interpolated using the inverse distance weighting (IDW) method.

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀▌
▶▌
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

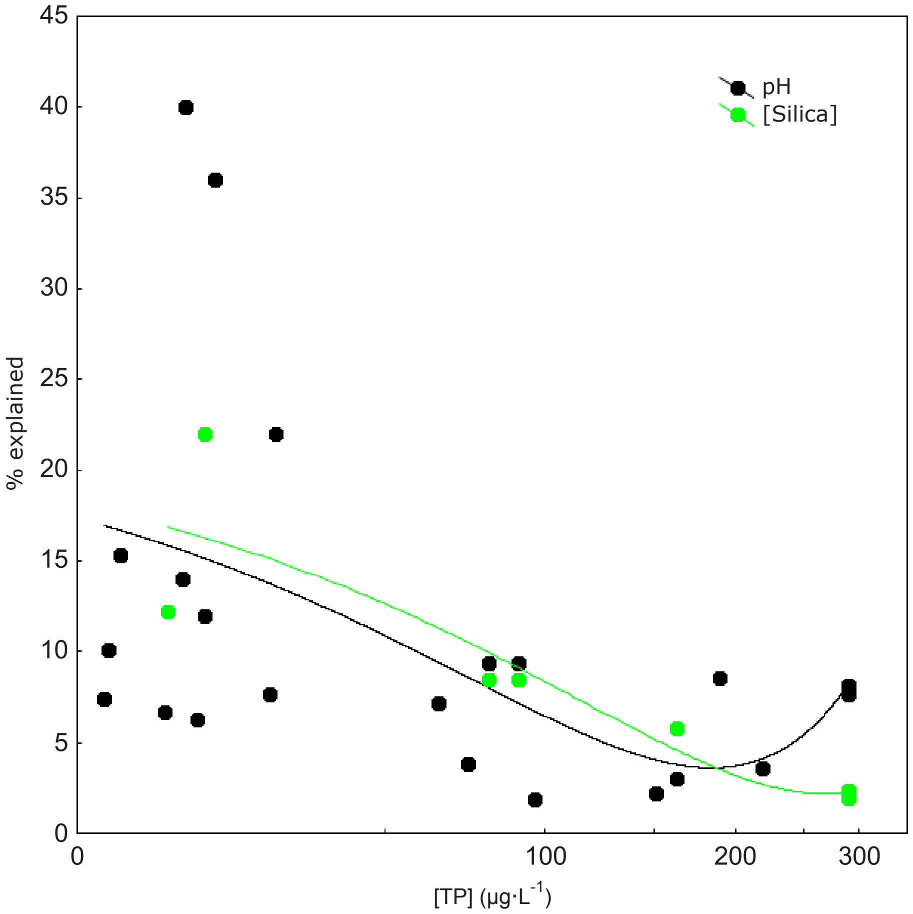

Figure 5. Percentage of variance captured by pH and silica concentration along the trophic gradient. Data fitted to distance-weighted least squares curves.