

Assessing the potential for developing internal diatom-based transfer functions for Lake Baikal

Anson W. Mackay, Richard W. Battarbee, and Roger J. Flower

Environmental Change Research Centre, Department of Geography, University College London, 26 Bedford Way, London. WC1H 0AP UK

Nicholas G. Granin

Limnological Institute, Russian Academy of Sciences Siberian Division, 6644033, Irkutsk, P.O. Box 4199, Russia

David H. Jewson

96 Desertmartin Road, Magherafelt, Co. Derry BT45 5HE, N. Ireland UK

David B. Ryves

Environmental History Research Group, Department of Quaternary Geological, Geological Survey of Denmark & Greenland, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark

Michael Sturm

EAWAG, Überlandstrasse 133, CH-8600, Dübendorf, Switzerland

Abstract

We explore possible quantitative relationships between diatom species and environmental data for Lake Baikal using multivariate techniques. Our approach differs from published studies in other regions (on training sets and transfer functions) because (1) although only one lake is examined, we use the internal lake gradients rather than gradients among lakes and (2) the majority of the dominant diatom taxa are endemic. Canonical correspondence analysis on 93 surface sediment diatom assemblages reveals that major taxa show distinct relationships with measured environmental variables. Five significant variables were identified: snow thickness on ice (which accumulates on the frozen lake surface between January and April/May), water depth, suspended matter, annual solar radiation, and mean July temperature of surface waters. The strongest relationship is with snow thickness, which influences light levels in the water below the ice. A diatom-based inference model has been developed to predict snow thickness on the ice using weighted averaging (WA) and WA with tolerance down-weighting (WA_{tol}) models. Results suggest that Lake Baikal diatom assemblages in surface sediments can be used to develop a robust model for estimating snow thickness across frozen Lake Baikal ($r^2_{jack} = 0.607$, RMSEP = 0.138 log cm). This model has potential applications for recent paleoclimate studies in this region of continental Eurasia.

Lake Baikal in southeast Siberia is a key site for paleoclimate research, because it is positioned close to the boundaries of two important weather systems, the Siberian high pressure and the Asian monsoon zones. Lake Baikal is extremely continental: mean daily air temperatures in January fall to an average of around -25°C , whereas in July they are in the region of $\sim 19^{\circ}\text{C}$ (Kozhova and Izmet'seva 1998). Lake Baikal's midlatitude position makes it sensitive to in-

solation changes: winters are long, cold, and dry, and summers, although short, are relatively warm and wet. Climate predictions for southern Siberia suggest that temperatures in winter will increase by $2\text{--}5^{\circ}\text{C}$ in the next 50 yr (IPCC 2001).

In Lake Baikal, two ice phenology records exist for the Angara River, Baikal's outflow, and Listvyanka Bay in the south basin (Fig. 1), which extend back to 1720 and 1869 CE, respectively (Livingstone 1999; Magnuson et al. 2000). Attempts to reconstruct climate parameters prior to these dates requires proxy data, and potentially this can be provided by the paleolimnological record. However, there are relatively few high-resolution, interannual studies that have attempted to reconstruct climate parameters over the last 1000 yr in the Lake Baikal region. Notable exceptions for other Siberian regions include the northern Urals (Briffa et al. 1995) and Putoran in northern Siberia (Naurzbaev and Vaganov 2000), both of which used dendroclimatological techniques to reconstruct past summer temperatures.

Several studies have used diatoms and/or biogenic silica to reconstruct Holocene paleoclimates in Baikal (e.g., Gran-

Acknowledgments

We thank the following agencies for funding: UK NERC (GR3/10529), the Leverhulme Trust (F134 AZ), the Swiss government (EAWAG-GEOPASS 96-97.37) and The Royal Society of London, under the Baikal International Centre for Ecological Research program. We also thank members of the Limnological Institute, Irkutsk, for providing logistical support, especially during spring expeditions on the ice, and summer expeditions on the open water (crews of both the R.V. *Vereschagin* and R.V. *Titov*). We finally thank members of the Cartographic Unit in the Department of Geography, UCL, for redrawing the figures used in the study.

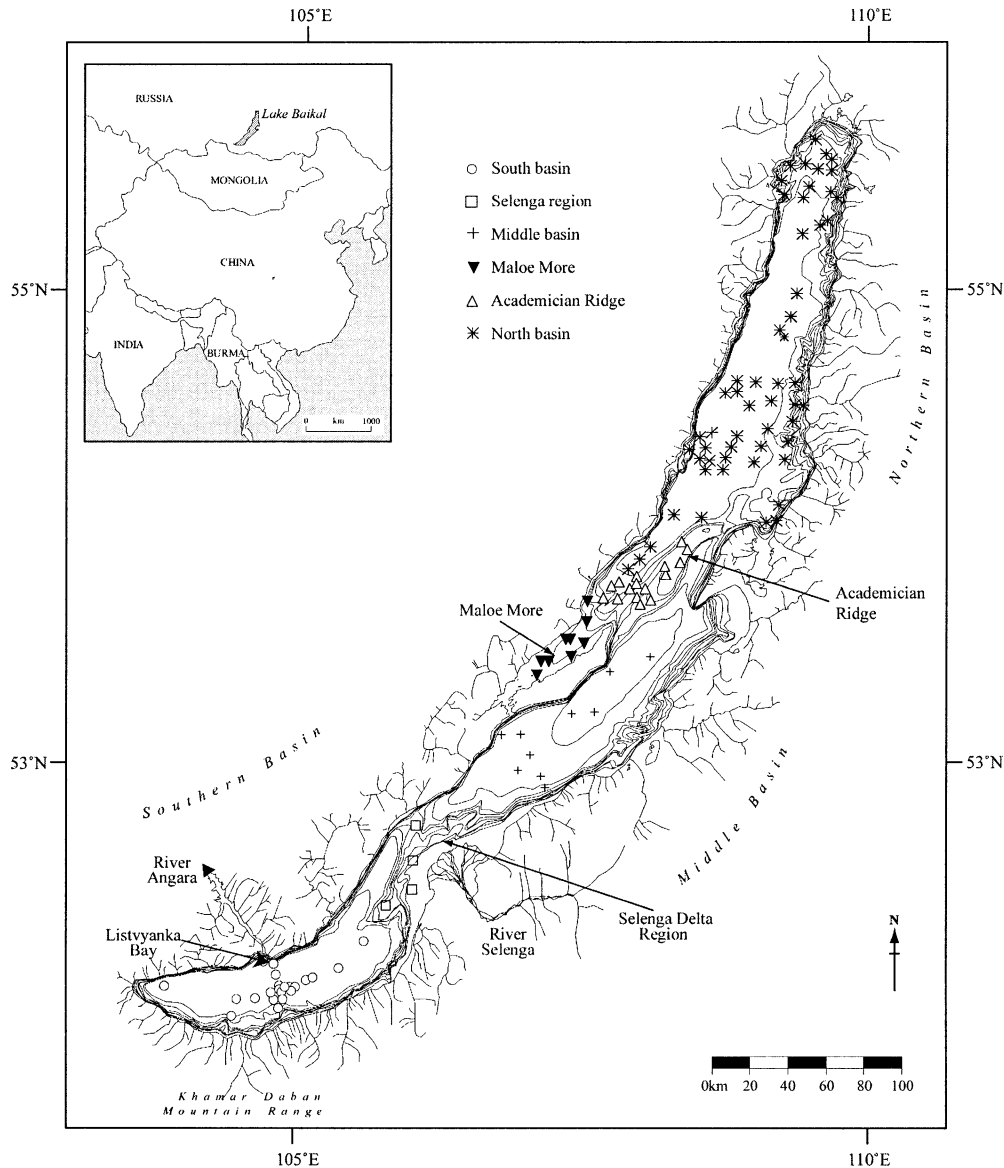


Fig. 1. Map of Lake Baikal showing location and distribution of 93 surface sediment samples. Samples taken from different regions have been coded, as shown in insert, to allow easier comparison with ordination plots. Places mentioned in the text are also highlighted, including Listvyanka Bay, shallow waters of the Selenga Delta region, the Maloe More (or "Little Sea") and the Academician Ridge.

ina et al. 1992; Qui et al. 1993; Bradbury et al. 1994; Karabanov et al. 2000), focusing mainly on millennial time-scales and on links between climate and diatom productivity. However, whereas changes in diatom productivity have been linked to orbital forcing and to influences from ice sheets and the North Atlantic, causal mechanisms have yet to be determined, although possible factors have been suggested, including surface water temperature, length of ice-free season, turbidity, and nutrient cycling (Colman et al. 1995). Karabanov et al. (2000) suggested that it is possible to use diatom abundances as a relative paleoclimate index of warmer/colder conditions, and although such studies provide useful and important information, the use of biogenic silica analysis and total diatom numbers alone ignore detail poten-

tially provided by changes in individual species and assemblages.

In the present study we use numerical techniques of regression and calibration that are used extensively in diatom-based paleolimnological studies to reconstruct environmental variables, such as pH, nutrients, salinity, conductivity, and water and air temperature. These standard approaches to transfer functions use training sets that make up modern assemblages from many lakes encompassing the environmental gradients of interest. This approach is not possible for Lake Baikal, because the majority of the key taxa are endemic. On the other hand, the lake is exceptionally large and spans a substantial climate gradient from north to south, providing an opportunity to generate an internal training set.

Lake Baikal is 636 km in length, extending across 4° of latitude from 55°46'N to 51°29'N, a distance equivalent to that used for multiple lake training sets at similar latitudes (e.g., Pienitz et al. 1995).

Certain properties of Lake Baikal, such as snow cover on the ice when the lake is frozen and subsequent timing of ice break-up, are influenced by the climate gradient across its length (Verbolov et al. 1965; Shimaraev et al. 1994) and can be used as proxy climate variables in the training set. For example, Lake Baikal freezes every winter earlier in the north basin than in the south basin, although local and regional factors (such as wind patterns) modify the precise timing (Verbolov et al. 1965). Mackay et al. (1998) have already demonstrated that many of the dominant planktonic diatoms in Lake Baikal surface sediments are distributed on a north-south gradient, whereas Granin et al. (1999) described how the ecology of many of these taxa are intricately linked with seasonal processes, such as ice formation, snow cover, and mixing processes. In general very little is known about the autecologies of endemic diatoms in Lake Baikal. We explore therefore herein the potential relationships between Lake Baikal diatoms and existing environmental data, in order to identify some key variables that can be investigated in more detail in future studies, with a view to recent climate reconstruction. This has led us to the development of a diatom-based “internal” transfer function to reconstruct climate attributes across the Lake Baikal region.

Site description

Lake Baikal occupies the world's deepest continental depression, at 1,164 m below sea level, and is made up of 3 large basins of differing ages (Fig. 1). The basins are separated by two underwater highs: the Selenga Delta Ridge in the south and the Academician Ridge in the north. There are also other notable shallow-water regions, including the Maloe More and the Barguzin Bay. Average (and maximum) water depths in the three main basins differ: 843 (1,423), 854 (1,636), and 576 (889) m in the south, middle, and north basins, respectively (Shimaraev et al. 1994). The Baikal rift zone is overlain by some 7,500 m of sediment, and, although past glaciations have influenced the catchment of Lake Baikal, leading to changes in hydrology and sedimentology, the deep-water sediments have never been directly affected.

Methods

Collation of diatom (response) and environmental (explanatory) data sets—The calibration training set is made up of 93 surface sediment samples collected from across the length of Lake Baikal between 1992 and 1997 (Fig. 1) (Mackay et al. 1998, 2000). They represent a subset of >120 cores taken, but surface samples were not used in the development of the training set from cores with either (1) visibly disturbed surface sediment-water interfaces (e.g., when the subsurface iron-manganese crust had broken through the surface layers or when the water sitting on top the surface sample was cloudy) or (2) turbidites present in the upper sections of the core (identified by visual inspection of lon-

gitudinally opened cores). Samples were prepared for diatom analysis according to the procedure detailed in Mackay et al. (1998), minimizing additional dissolution of valves by omitting chemical pretreatments during preparation. Approximately 300 valves were counted for each sample, and counts are archived in the ECRC database AMPHORA. Diatom taxa were identified according to a range of published papers and books, several of which deal specifically with the endemic diatom flora found in Lake Baikal (e.g., Meyer 1930; Skvortzow 1937; Zabelina et al. 1951; Skabichevskii 1960; Gleser et al. 1988; Khursevich 1989; Genkal and Popovskaya 1990; Genkal et al. 1992; Flower 1993; Kozhova and Kobanova 1995; Edlund et al. 1996). Taxonomic uncertainties over some of the problematic taxa were investigated in detail at two joint workshops between Russian and British diatomists. In all, 87 taxa were identified, and diatom counts were transformed to percentages prior to all analyses.

With respect to collation of the diatom training set, account needs to be taken of taphonomy (e.g., diatom dissolution), bioturbation, sediment resuspension, and sediment redeposition processes. Diatom dissolution in Lake Baikal has a major impact on the composition of species assemblages being incorporated into the sedimentary record (Mackay et al. 2000; Ryves et al. in press). Nevertheless, transfer functions from dissolved assemblages are still effective tools (e.g., Pichon et al. 1992), and work is currently under way to develop “correction factors” to provide more realistic quantitative estimates of Lake Baikal diatoms in sedimentary assemblages. Bioturbation may potentially be a problem because of the lake's oxygenated hypolimnion, which allows for diverse (and almost wholly endemic) faunal communities to inhabit the surface sediments. Using thin-sectioning techniques, Francus and Karabanov (2000) provided evidence, however, for the presence of laminae of *Aulacoseira* species in the bottom sediments. They suggested that these laminae are directly related to blooms of *Aulacoseira baicalensis* (Meyer) (Simonsen) that occur every 3–5 yr in the lake (so-called *Melosira* years) and concluded that the presence of these laminae are an unusual occurrence in an environment prone to bioturbation. The actual extent of bioturbation of the surface sediments of Lake Baikal is therefore unknown, and, until more research is carried out, it is difficult to fully quantify its effects. Sediment resuspension has been suggested to occur in Lake Baikal (e.g., Kempe and Schaumburg 1995), although our studies suggest that there is little sediment resuspension at depth (Ryves et al. in press), and the process is not considered further herein. Finally, account needs to be taken of sediment redeposition processes, especially in the form of turbidites. These processes are complex, but in Lake Baikal sediment profiles affected by turbidites can be recognized through a combination of techniques, such as radiometric dating and dry weight analysis (e.g., Mackay et al. 1998), mineral magnetic analysis (e.g., Lees et al. 1998), and visual inspection of cores opened longitudinally. In the present study, we have further screened out cores where the uppermost layers have been affected by secondary deposition.

Regular chemical and biological monitoring of Lake Baikal has been carried out for many decades, whereas, for some variables, such as surface water temperatures, measurements have been made for >100 yr. The most compre-

Table 1. Limnological, climatic, chemical, and biological data used as potential explanatory variables for the distribution of diatoms in surface sediments. Original source references include: * data from the present study; † Atlas Baikala (1993); ‡ Verbov et al. (1965); § Shimaraev et al. (1994); ¶ Votinsev 1975; || Tarasova and Mesheryakova (1992); ** Bondarenko et al. (1996); †† Popovskaya (1991); n/a, not available; ‡‡ n/a, long term observations over the last century.

Explanatory variable	Code	Units	Dates measured	Min	Max	Mean	STDS
Limnology							
Water depth*	Depth	m	1992–1997	20.0	1678.0	809.6	480.3
Transparency in July†	Transjul	m	1961	9.0	25.0	21.4	4.6
Transparency in September†	Transsep	m	1961	2.0	11.0	6.0	2.0
Albedo in July†	Albedo	%	n/a	5.0	7.0	5.1	0.4
Temperature of July water surface‡	Tempws	°C	1896–1959	1.0	13.0	5.9	2.3
Climate							
Annual solar radiation†	Solar	mJ m ⁻²	n/a	4100.0	4700.0	4478.5	187.6
Absorbed radiation in July†	Absorb	Kcal cm ⁻²	n/a	13.0	15.0	14.2	1.0
Mean July air temperature†	Julair	°C	n/a‡‡	9.0	15.0	12.2	1.7
Annual precipitation†	Precip	mm	n/a‡‡	175.0	550.0	257.3	67.7
Snow thickness on the lake in March†	Snow	cm	1972–1973	2.5	12.5	6.2	3.2
Ice thickness in March§	Icethick	cm	1972–1985	65.0	100.0	82.9	7.2
Length of ice cover§	Icelength	days	n/a‡‡	126.0	168.0	145.3	10.7
Chemistry							
N-NO ₃ ^{5¶}	Nitrate	mg m ⁻³	July 1978	15.0	70	58.17	16.4
P-PO ₄ ^{5¶}	Phosphate	mg m ⁻³	July 1978	4.0	12.00	7.73	2.7
Suspended matter	Suspmat	mg L ⁻¹	July 1978	0.8	4.00	2.32	0.6
Suspended organic C	SusporgC	mg L ⁻¹	July 1978	0.3	1.50	0.38	0.2
Organic C	Org C	mg L ⁻¹	July 1978	1.1	1.70	1.33	0.2
Organic N	Org N	µg L ⁻¹	July 1978	70.0	175.00	118.38	24.2
Organic P	Org P	µg L ⁻¹	July 1978	7.0	16.50	9.96	2.1
Biology							
Average length of 100 <i>A. baicalensis</i> valves in surface sediment samples*	Length	µm	1992–1997	20.3	29.4	25.2	1.7
Diurnal primary production**	Primary	g O ₂ m ⁻²	June 1991	0.3	2.8	1.0	0.6
Phytoplankton biomass††	Phytopl	mg m ⁻³	Spring 1985	50.0	1100.0	150.0	172.0
Zooplankton biomass in September†	Zoobio	G m ⁻²	1985	5.0	30.0	10.0	8.6

hensive source of accessible data exists in the *Atlas Baikala* (1993), which reviews (among others) limnological, chemical, biological, and climatic data and currently represents one of the best available published data sources for sites across the whole lake. We know from published studies that many climate and nutrient variables have a significant impact on diatom distributions, and although >40 variables are mapped out in the *Atlas*, we only selected those that exhibited distinct gradients and that are known or suspected to have an effect on diatom ecology. Selection prior to analyses was carried out because standard diatom transfer function methodology relates changes in relative species abundances (the response data set) to measured environmental variables that directly affect species distributions or are correlated with variables that do. Thus, variables that exhibit no gradients cannot, by definition, explain variance in the diatom (response) data set. Although we do not assume that nonsignificant variables are unimportant for individual species populations, the transfer function approach seeks to find the dominant environmental gradient(s) that best explain changes in community structure within the response data set.

A total of 23 variables from several sources (including the *Baikal Atlas*) were selected as explanatory variables for diatom species distribution across the length of Lake Baikal (see Table 1 for full details). The data in the *Atlas* are most

commonly presented as contour maps of monthly mean values (most often between June and December when the lake is free of ice), and only occasionally are annual mean values given. Maps for July measurements were most common between all the variables, and, as a consequence these are the ones used in the present study, but exceptions include snow and ice measurements made in March and zooplankton biomass measurements made in September. Where possible, original sources of the data were consulted to determine when measurements were made, although, for some variables, dates are only given as long-term observations over the past century (see Table 1). Environmental data derived from the *Atlas* and used in the present study were linearly interpolated from nearest mapped values for each core location. Other variables in the *Atlas* that were not used in the present study include bacterial concentrations, wave height, fog cover, cloudiness, and catchment characteristics—for example, vegetation cover.

An important limitation to our methodology is the mismatch between surface sediment sampling (1992–1997), with periods when environmental data were originally collected (see Table 1). However, Mackay et al. (1998) have demonstrated that the sedimentary composition of diatoms (in a suite of cores taken across the length of Baikal) has not changed over at least the past 50 yr, and so the com-

position of dominant taxa in surface sediments is similar on collection to when the environmental data were measured. Although no diatom changes are suggested in the sedimentary record, monitoring studies do indicate that a few cosmopolitan species (e.g., *Nitzschia acicularis* W. Smith) are increasing in number, perhaps because of anthropogenic impacts on the lake (Popovskaya 1991). Selective species dissolution is suggested to be the important reason for this discrepancy (Ryves pers. comm.).

Data analyses—Unconstrained (detrended correspondence analysis [DCA] and principal components analysis [PCA]) and constrained (canonical correspondence analysis [CCA]) ordinations were carried out using the software package CANOCO v. 4 (ter Braak and Šmilauer 1998). Environmental variables positively skewed around the mean were transformed using $\log(x + 1)$ except for total annual absorbed radiation, for which transformation was not required. Major trends in the full explanatory data set were initially explored using PCA on a correlation biplot. The principal gradient of floristic variation within the calibration training set was assessed prior to ordination analyses using DCA, with detrending by segments, nonlinear rescaling of axes, and down-weighting of rare species. A first axis gradient of 3.06 SDs was obtained, and further numerical analyses involving species data were therefore based on unimodal species-response models.

Relationships between diatom distributions and the full environmental data set were initially examined using the direct gradient technique of CCA, which constrains ordination axes as linear combinations of environmental variables (ter Braak and Šmilauer 1998). Forward selection (a form of stepwise regression) was used to remove redundant variables, leaving only those that explained significant amounts of variation in the species data (identified using unrestricted Monte Carlo permutation tests; ter Braak and Šmilauer 1998). Variables were considered to be significant when $p < 0.05$; 199 permutations, but, because multiple tests were performed on the data, Bonferroni corrections were applied to p (Legendre and Fortin 1989), to produce a more parsimonious model. Environmental variables that were not significant were omitted from further analyses. The significance of both first and second canonical axes was also tested for all the data sets outlined above, again using unrestricted Monte Carlo permutation tests (199 permutations).

Additionally, samples and environmental variables were screened within CANOCO to determine samples with extreme environmental values and to exclude those environmental variables that exhibited high collinearity—that is, variance inflation factors >20 , which have the effect of destabilizing the data set (ter Braak and Šmilauer 1998). Further tests to detect poorly fitted samples (outliers) within the model were also made (see below).

Development of inference models—Diatom-based inference models, or transfer functions, were constructed using the techniques of weighted averaging (WA) and tolerance down-weighted WA, both with inverse and classical de-shrinking, using the program CALIBRATE 0.81 (Juggins and ter Braak 1997). The predictive ability of the various

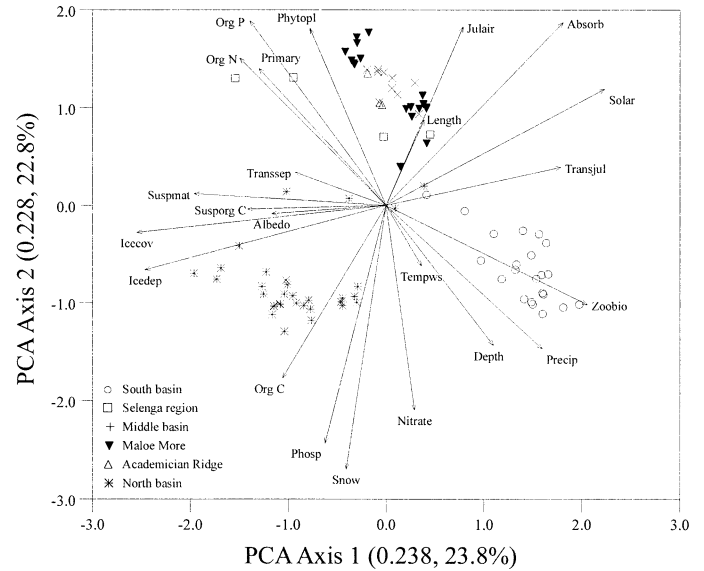


Fig. 2. PCA biplot showing 23 environmental variables (vectors) and 93 sampling locations.

models are assessed in terms of the correlation coefficients (r^2) between the observed and diatom-inferred values, in terms of the apparent root mean square error (RMSE) and in terms of the cross-validated error of prediction using the technique of leave-one-out jackknifing.

Results

A PCA biplot of all the environmental variables is shown in Fig 2. Eigenvalues for axes 1 and 2 are 0.238 and 0.228, whereas those for axes 3 and 4 are only 0.150 and 0.059, respectively, and are not considered further in the present study. The first two axes therefore capture just $>46\%$ of the variation in the environmental data set. Axis 1 is dominated by ice cover and ice thickness gradients, which all decrease with positive values along axis 1. Axis 2 is related to a heterogeneous set of variables, principally snow depth, phosphate, and nitrate concentrations. The positions of surface samples are also shown on the biplot, which demonstrates that sites located in the south basin and sites located in the north basin are separated along the first axis. South basin sites have the highest annual precipitation levels and zooplankton biomass, whereas sites in the northern basin have the highest values for length of ice cover, thickness of ice, thickness of snow cover on the ice, and suspended matter (see Table 1 for details on sampling dates). Sites in the middle basin, the Maloe More, and especially the shallow waters of the Selenga Delta region have greatest values of organic nitrates and organic phosphates (measured in July 1978), which are inversely related to depth. These shallow sites are also associated with high spring phytoplankton biomass (Fig. 2).

Initially, the direct gradient technique of CCA was performed on the full data set. However, herein we present results for a reduced set of environmental variables determined using forward selection and Monte Carlo permutation tests.

Table 2. CCA of diatom and reduced environmental data set after forward selection.

	CCA Axes				Total inertia
	1	2	3	4	
Eigenvalue	0.264	0.121	0.069	0.053	2.055
p	0.01	0.01			
Cum. % variance of response data	12.8	18.7	22.1	24.7	26.7
Sum of all unconstrained canonical eigenvalues					2.055
Sum of all canonical eigenvalues					0.549

This procedure identified five environmental variables that are independently significant in explaining variation in the diatom data: snow thickness in March, water depth, concentration of suspended matter in July, annual solar radiation (predominantly in the spectral range of 400–700 nm; Shimaraev et al. 1994), and temperature of surface waters in July averaged over the period 1896–1959: see Table 1 for further details of sampling frequency and original data sources.

The results of CCA using only the five environmental variables are given in Table 2, which explain 26.7% of the variation in the whole data set, whereas Monte Carlo permutation tests performed on axes 1 and 2 demonstrate that both axes are significant. Relationships between the (1) 5 significant variables, (2) 93 site locations, (3) 5 dominant planktonic taxa, and (4) the remaining 82 taxa in the training set are shown in a CCA triplot (Fig. 3). It is convenient here to outline some of the major features of a CCA triplot, because this will aid interpretation in the discussion below: (1) surface sediment samples with similar species composition are located close together and tend to be dominated by species located near to them in the ordination diagram; (2) the order of samples along the environmental variable vector approximates the order of values of that environmental variable among samples, with the origin representing the mean value of that environmental variable in the data set (similarly, one can extend the arrow back through the origin to get decreasing values); (3) small angles between environmental variable vectors indicate that those variables are closely correlated, whereas if the angle between variables is close to 90°, then there is little correlation, and angles close to 180° suggest a negative correlation; (4) a species position on the triplot approximates its WA optimum relative to other taxa and approximates its relationship to the environmental variables (Pienitz et al. 1995).

These results confirm those of Meyer (1930) that regions within Baikal are floristically distinct. For example, diatom assemblages in the Maloe More are dominated by *Stephanodiscus meyerii* Genkal & Popovskaya (a species characteristic of shallower, more nutrient-rich waters), whereas species such as *Synedra acus* v. *radians* Hust., *Aulacoseira skvortzowii* Edlund, Stoermer & Taylor and *S. meyerii* are not common in the north basin (Fig. 3). Values for suspended matter are greatest in northern basin sites and are highly correlated with increasing values for snow thickness. Snow is also correlated with water depth, which indicates that snow lies more thickly across the deep-water basins than shallow water regions. Several species exhibit interesting relationships with specific environmental variables, most notably *A. baicalensis* and *Cyclotella minuta* with thickness of snow on the ice, and *S. acus* v. *radians* with mean annual solar radiation. *S. meyerii* is not characterized well by any of the five significant environmental variables.

Inference models should only realistically be developed for variables that have first and second eigenvalue ratios (λ_1/λ_2) > 0.50 (Dixit et al. 1991). In the present study, the most robust variable is snow thickness, with an eigenvalue ratio of 0.55, which alone explains 9.6% of species variation in the diatom data set ($p = 0.002$; $n = 499$ permutations). This ratio suggests that there is an independent diatom signal for snow thickness on the ice and that reliable inferences can be obtained. Using these criteria, WA_{tol} down-weighting performs better than simple WA as a predictive model of snow depth. Significant relationships can therefore be established between diatom assemblages in Lake Baikal and snow depth ($r_{\text{jack}}^2 = 0.607$, RMSEP = 0.138 log cm; Table 3). The plot of observed and predicted snow depths on Lake Baikal shows a rather complex relationship (Fig. 4) and is clearly not appropriate for low values of snow thickness, where the range of predicted values is very large. Potential sample out-

Table 3. Summary statistics for log₁₀ snow depth WA and tolerance down-weighted (Tol d/w) WA regression with cross-validation by jack-knifing of the data set with 89 samples (and 83 taxa) used for the derivation of a transfer-function. Results for inverse deshrinking only are shown.

	Inverse deshrinking			
	RMSE(P) Log snow depth, cm	r_{jack}^2	Avg bias	Max bias
Apparent Simple WA	0.136	0.617	0	0.135
Apparent Tol d/w WA	0.124	0.681	0	0.126
Jack-knifed WA	0.148	0.547	0.0021	0.152
Jack-knifed Tol d/w WA	0.138	0.607	0.0036	0.126

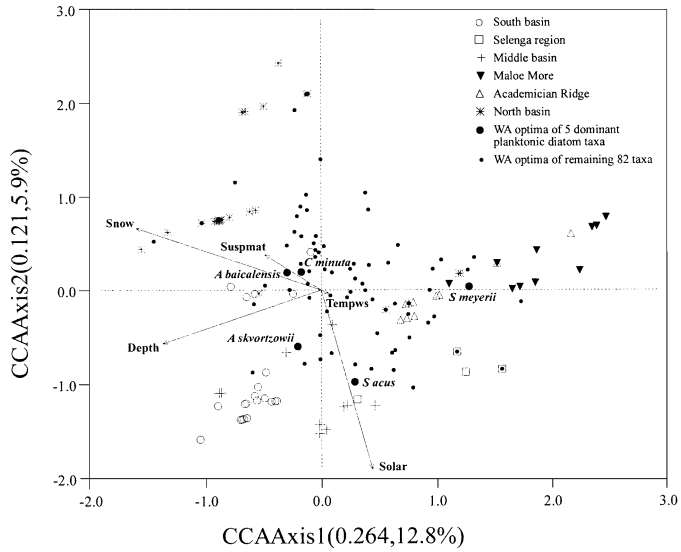


Fig. 3. CCA triplot, showing (1) five forward-selected environmental variables (vectors); (2) surface sediment samples (symbols as in Fig. 1); (3) WA optima of five dominant planktonic diatom taxa (large black circles); (4) WA optima of remaining 82 taxa (small black circles).

liers in the training set were identified as those with a jackknifed residual (inferred-observed) greater than the SD of the environmental parameter. The influence of these samples on model parameters was then assessed using Cook's *D* (Cook and Weisberg 1982) and compared with the critical value of $4/n$, with n being number of sites in the model (Jones and Juggins 1995). Poorly fitted samples with a low Cook's *D* can be deleted safely, whereas those with a high Cook's *D* were carefully inspected for their effect on species' optima before deletion. Using these criteria, four samples were deleted from the model only: BAIK 17, 65, 70, and 126. As a final check on the data, the original data for these sites were examined to ascertain why they were outliers. For example, BAIK 17, taken from near the Selenga Delta region, is characterized by fluvial diatoms from the Selenga River rather than by open-water or even littoral species characteristic of the other sites. On deleting these four samples, four taxa not present in any of the other samples were also deleted, leaving 83 taxa used in the model.

Discussion

Relationships of diatoms to environmental variables—Diatom phytoplankton in Lake Baikal, which are dominated by a small number of taxa whose populations vary significantly from year to year, are subject mainly to physical controls, such as light, temperature (Richardson et al. 2000; Jewson et al. unpubl. data), ice formation, and water-column mixing (Granin et al. 2000). Their growth is only limited by nutrient depletion during summer stratification, which lasts for ~19–22 weeks, taking into account summer heating and autumn cooling (Shimaraev et al. 1994). During this period of stratification, picoplankton are dominant (especially *Synechocystis limnetica* Popovskaya), although taxa from other groups such as Chrysophyta and Cryptophyta are also abundant.

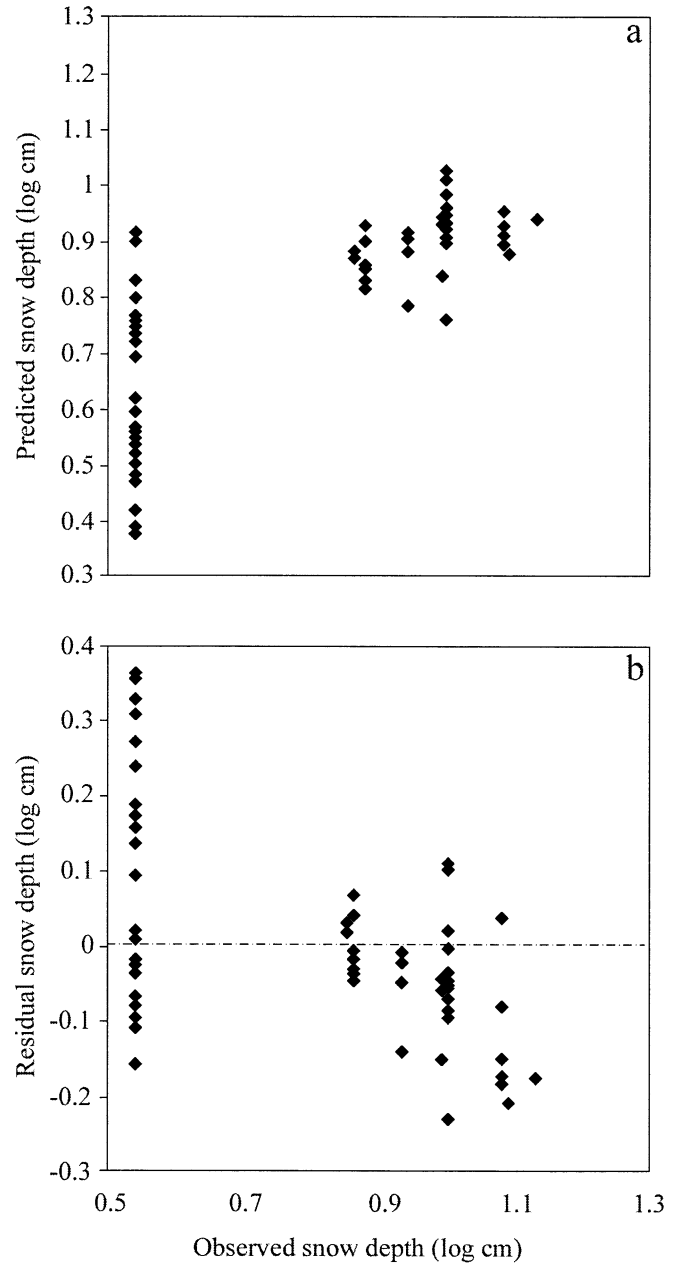


Fig. 4. (a) Observed versus predicted snow thickness values (log cm) on frozen Lake Baikal. Predictive values were determined using jackknifing. (b) Prediction residuals (measured-predicted) versus measured snow thickness values using a WA_{tol} model.

Previous authors working on Lake Baikal have discussed compositional changes in diatom assemblages over the Holocene period and have related these to changes in climate. For example, Bradbury et al. (1994) discussed the possible role that climate has on the productivity of shallow water regions of Lake Baikal, concomitant with effects on common species found in these areas such as *S. meyerii* and *S. acus*. More recent diatom assemblage changes have been associated with climate changes since the middle of the 19th century (e.g., Edlund et al. 1995; Mackay et al. 1998). These assemblages are dominated by *C. minuta*, *A. baicalensis*, *A.*

skvortowii, *S. meyerii*, and *S. acus* v. *radians* and are still important in the lake today. In the present study, the discussion will focus on these five taxa, taking into account their relationships with significant environmental variables, as revealed by CCA.

S. acus v. *radians* is strongly correlated with annual solar irradiation (Fig. 3) (the main source of heat into the water column of the lake) (Shimaraev et al. 1994), in agreement with culturing experiments which suggest that *S. acus* v. *radians* grows well at high irradiances (e.g., between 30 and 70 $\mu\text{mol Q m}^{-2} \text{s}^{-1}$) and high temperatures, in comparison to other Baikal diatoms—for example, *A. baicalensis* and *C. minuta* (Jewson et al. unpubl. data), which have the lowest WA optima of the five species for solar irradiation. *S. meyerii* populations are highest in shallow-water environments such as those in the region of the Selenga Delta (Fig. 3), which are more nutrient-rich than deep water environments. Our study suggests that, of the variables selected through forward selection, no measured significant environmental variable is closely related to *S. meyerii*, although it is associated with very shallow water and low values of snow thickness. In our larger study, *S. meyerii* populations were greatest in early summer during May and June, after ice break-up (Ryves et al. in press). Results herein serve to highlight that further work needs to be done in assessing the role of environmental variables on the ecology of *S. meyerii*. Population growth of *A. skvortowii* is complex, because this species blooms both in the spring under the ice (Popovskaya 1977) and during the autumn in near-shore regions where surface temperatures can reach $>26^{\circ}\text{C}$ (Kozhova and Izmet'eva 1998). In Fig. 3, *A. skvortowii* has one of the highest WA optima of all species for water depth, which suggests that populations here are represented by those blooming under the ice.

Populations of *C. minuta* increase early in the autumnal cooling period (Jewson et al. unpubl. data), when water transparency (determined by Secchi disc) is at a minimum because of high concentrations of picoplankton leading to increased organic matter in the surface waters of the lake (Kozhova and Izmet'eva 1998). This is reflected in Fig. 3, where *C. minuta* has high WA optima values for suspended matter (Fig. 3). It is interesting to note that *C. minuta* has similar WA optima for both suspended matter and snow cover close to those for *A. baicalensis*. This apparent similarity is unlikely to suggest that *C. minuta* has similar strategies in response to snow cover but more likely suggests that the two species are present in similar proportions in the surface sediments across much of the lake.

In terms of ranked species optima for snow thickness, *A. baicalensis* narrowly has the highest WA optima of the five species under consideration, whereas *S. meyerii* has the lowest. Granin et al. (1999) outline how the penetration of light through the frozen surface of Lake Baikal provides enough energy for a dynamic community to survive under the ice in the spring, which is especially important for *A. baicalensis* populations. Light penetration, however, is also dependent on the thickness of snow on top of the ice (Granin et al. 2000). Kelly (1997) estimates that between 4% and 11% of solar radiation is able to penetrate through clear ice, providing enough energy for algal growth in spring, while as little

as 5 cm of snow cover on the ice can reduce solar transmission by a factor of 50; thus, when depths of surface snow are >10 cm, reduced light penetration is limiting for Baikal diatom growth (Granin et al. 1999, 2000; Jewson and Granin 2000). Moreover, Semovski et al. (2000) outlined how, when surface snow layers melt, “pipkrake,” or needle ice, is produced, which allows more effective transmission of solar radiation, from 20% to 80% (Sherstyankin 1975 in Semovski et al. 2000). A combination of persistent snow cover and the reduced production of wet ice in specific regions of Baikal (related to increased snow cover) act to reduce levels of light transmission through the ice, thereby preventing large population growths of *A. baicalensis* as monitored in other regions. For example, monitoring studies demonstrate variability in crop sizes of *A. baicalensis* in the south basin: snow thickness is usually greater toward the east of the lake than to the west, and, when this occurs, the concentration of *A. baicalensis* valves is lower on the east side of the south basin (Jewson and Granin 2000; Mackay et al. 2000). Work is currently under way to explore relationships between diatoms and snow cover using remote sensing and to monitor at regular intervals responses of diatom phytoplankton to changing hydrological conditions across the length of the lake.

The correspondence between the distribution of planktonic diatoms in surface sediments and the ecology of these species established from monitoring and culture studies underlines the potential of using an internal transfer function for climate reconstruction from sediment cores. However, interpretation of sediment data also needs to allow for problems related to diatom dissolution: Ryves et al. (in press) have found that dissolution accounts for almost 20% of diatom variation in surface sediment samples. For example, in certain years, the finely silicified taxon *N. acicularis* is abundant in the water column (e.g., in 1995), especially in the south basin, but it is rarely found in the sediment record. As yet we do not fully understand the complexity of diatom dissolution in the lake, but sequencing trap data show that, for most species, dissolution in the water column is not extensive, whereas analyses of dissolution in the sediments suggests that valves once incorporated into the sediment record no longer undergo further dissolution (Ryves et al. in press). In surface sediments, ~50% of diatom frustules are affected to some degree by dissolution, and provisional results suggest that only a small proportion of the valves or cells produced in the water column are ultimately preserved in the sediments (Mackay et al. 2000). Experimental work on diatom dissolution from saline lakes has suggested that dissolution can be quantified and linked to biogenic silica and diatom valve loss (Ryves et al. 2001). Work is currently in progress to establish correction factors (linking concentrations of specific species in the water column with concentrations found in surface sediments) for the major planktonic species that should enable more confident reconstructions to be made in the future.

Our study is the first to calibrate modern Lake Baikal diatom assemblages to environmental variables, but because the data set is based on assemblages dominated by endemic diatoms, the model cannot be applied to other sites. The “internal” methodology rests on the assumption (supported

by, e.g., Shimaraev et al. 1994) that significant environmental gradients exist within the lake. The explanatory data set has largely been derived from an amalgamation of previously published work, and, although these data have specific limitations, the resultant model is robust and promising. Research is now being undertaken to incorporate additional data from remotely sensed parameters to improve the model further (e.g., recent measurements of ice cover, snow cover and chlorophyll *a*), to provide a more spatially coherent, internally consistent, and high-resolution data source across the length of the lake.

References

- ATLAS BAIKALA. 1993. Baikal atlas (in Russian). Federal Service of Surveying and Cartography.
- BONDERENKO, N. A., N. E. GUSELNIKOVA, N. F. LOGACHEVA, AND G. V. POMAZKINA. 1996. Spatial distribution of phytoplankton in Lake Baikal, spring 1991. *Freshw. Biol.* **35**: 517–523.
- BRADBURY, J. P., YE. V. BEZRUKOVA, G. P. CHERNYAeva, S. M. COLMAN, G. KHURSEVICH, J. W. KING, AND YE. V. LIKOSHWAY. 1994. A synthesis of post-glacial diatom records from Lake Baikal. *J. Paleolim.* **10**: 213–252.
- BRIFFA, K. R., P. D. JONES, F. H. SCHWEINGRUBER, S. G. SHIYATOV, AND E. R. COOK. 1995. Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia. *Nature* **376**: 156–159.
- COLMAN, S. M., J. A. PECK, E. B. KARABANOV, S. J. CARTER, J. P. BRADBURY, J. W. KING, AND D. F. WILLIAMS. 1995. Continental climate response to orbital forcing from biogenic silica records in Lake Baikal. *Nature* **378**: 769–771.
- COOK, D. R., AND S. WEISBERG. 1982. Residuals and influence in regression. Chapman & Hall.
- DIXIT, S. S., A. S. DIXIT, AND J. P. SMOL. 1991. Multivariable environmental inferences based on diatom assemblages from Sudbury (Canada) lakes. *Freshw. Biol.* **26**: 251–266.
- EDLUND, M. B., E. F. STOERMER, AND C. H. PILSKALN. 1995. Siliceous microfossil succession in the recent history of two basins in Lake Baikal, Siberia. *J. Paleolim.* **14**: 165–184.
- EDLUND, M. B., E. F. STOERMER, AND C. M. TAYLOR. 1996. *Aulacoseira skvortzowii* sp. nov. (Bacillariophyta), a poorly understood diatom from Lake Baikal, Russia. *J. Phycol.* **32**: 165–175.
- FLOWER, R. J. 1993. A taxonomic re-evaluation of endemic *Cyclotella* taxa in Lake Baikal, Siberia. *Nova Hedwigia* **106**: 203–220.
- FRANCUS, P., AND E. KARABANOV. 2000. A computer-assisted thin-section study of Lake Baikal sediments: A tool for understanding sedimentary processes and deciphering their climatic signal. *Int. J. Earth Sci.* **89**: 260–267.
- GENKAL, S. I., AND G. I. POPOVSKAYA. 1990. On the fine structure of *Cyclotella baicalensis* (K. Meyer) Skv. and *Cyclotella minuta* (Skv.) (in Russian). *Antipova Biol. Inland Waters Inf. Bull.* **87**: 21–26.
- GENKAL, S. I., AND OTHERS. 1992. Diatom algae of the USSR: Relict and modern (in Russian). V. II, *Stephanodiscus*, p. 7–48. In I. V. Makarova [ed.] Nauka.
- GLESER, S. I., I. V. MAKAROVA, A. I. MOISSEVA, AND V. A. NIKOLAIEV [EDS.]. 1988. The diatoms of the U.S.S.R., fossil and recent (in Russian). V. 2. Nauka.
- GRANIN, N. G., AND OTHERS. 1999. Turbulent mixing in the water layer just below the ice and its role in development of diatomic algae in Lake Baikal. *Dokl. Akad. Nauk.* **366**: 835–839.
- . 2000. Turbulent mixing under ice and the growth of diatoms in Lake Baikal. *Verh. Int. Ver. Limnol.* **27**: 1–3.
- GRANINA, L. Z., E. B. KARABANOV, M. K. SHIMARAeva, D. F. WILLIAMS, AND V. M. KUPTSOV. 1992. Biogenic silica of Baikal bottom sediments used for palaeo reconstruction. *In International Project on Palaeolimnology and Lake Cenozoic Climate Newslett.* **6**: 53–59.
- IPCC 2001. Climate change 2001: The scientific basis. Working Group I contribution to the IPCC Third Assessment Report (Summary for Policy Makers) (<http://www.ipcc.ch>).
- JEWSON, D. H., AND N. G. GRANIN. 2000. How can present day studies of diatoms help in understanding past climatic change in Baikal? *Terra Nostra* **9**: 29–33.
- JONES, V. J., AND S. JUGGINS. 1995. The construction of a diatom-based chlorophyll *a* transfer function and its application at three lakes on Signy Island (maritime Antarctic) subject to differing degrees of nutrient enrichment. *Freshw. Biol.* **34**: 433–445.
- JUGGINS, S., AND C. J. F. TER BRAAK. 1997. CALIBRATE—a computer program for species environmental calibration by [weighted averaging] partial least squares regression. 0.81. Univ. of Newcastle.
- KARABANOV, E. B., A. A. PROKOPENKO, D. F. WILLIAMS, AND G. K. KHURSEVICH. 2000. A new record of Holocene climate change from the bottom sediments of Lake Baikal. *Palaeogeog. Palaeoclim. Palaeoecol.* **156**: 211–224.
- KELLY, D. E. 1997. Convection in ice-covered lakes. Effects on algal suspension. *J. Plankton Res.* **19**: 1859–1880.
- KEMPE, S., AND M. SCHAUMBURG. 1995. Vertical particle flux in Lake Baikal, p. 323–355. In V. Ittekkot, S. Honjo, and P. J. Depetris [eds.], *Particle flux in the ocean*. Wiley.
- KHURSEVICH, G. K. 1989. Atlas of *Stephanodiscus* and *Cyclotella*, Bacillariophyta, from the Upper Cenozoic deposits of the USSR (in Russian). In F. Yu. Velichkevich [ed.], *Nauka & Teknika*.
- KOZHOVA, O. M., AND G. I. KOBANOVA. 1995. On the morphology and systematics of centric diatom algae of Lake Baikal. In O. M. Kozhova [ed.], *Ecological problems V. 2* (in Russian). Nauka.
- KOZHOVA, O. M., AND L. R. IZMEST'EVA. 1998. Lake Baikal: Evolution and biodiversity. Backhuys Publishers.
- LEES J. A., R. J. FLOWER, D. B. RYVES, E. VOLOGINA, AND M. STURM. 1998. Identifying sedimentation patterns in Lake Baikal using whole core and surface scanning magnetic susceptibility. *J. Paleolimnol.* **20**: 187–202.
- LEGENDRE, P., AND M.-J. FORTIN. 1989. Spatial patterns and ecological analysis. *Vegetatio* **80**: 107–138.
- LIVINGSTONE, D. M. 1999. Ice break-up on southern Lake Baikal and its relationship to local and regional air temperatures in Siberia and to the North Atlantic Oscillation. *Limnol. Oceanogr.* **44**: 1486–1497.
- MACKAY, A. W., AND OTHERS. 1998. Diatom succession trends in recent sediments from Lake Baikal and their relationship to atmospheric pollution and to climate change. *Phil. Trans. R. Soc. Lond. B* **353**: 1011–1055.
- MACKAY, A. W., R. W. BATTARBEE, R. J. FLOWER, D. JEWSON, J. A. LEES, D. B. RYVES, AND M. STURM. 2000. The deposition and accumulation of endemic planktonic diatoms in the sediments of Lake Baikal and an evaluation of their potential role in climate reconstruction during the Holocene. *Terra Nostra* **9**: 34–48.
- MAGNUSON, J. J., AND OTHERS. 2000. Historical trends in lake and river ice cover in the northern hemisphere. *Science* **289**: 1743–1747.
- MEYER, K. I. 1930. Introduction into the algae flora of Lake Baikal (in Russian). *Bull. Moscow Soc. Nat.* **39**: 179–396.

- NAURZBAEV, M. M., AND E. A. VAGANOV. 2000. Variation of early summer and annual temperature in east Taymir and Putoran Siberia over the last two millennia inferred from tree rings. *J. Geophys. Res.* **105**: 7317–7326.
- PICHON, J.-J., L. D. LABEYRIE, G. BAREILLE, M. LABRACHERIE, J. DUPRAT, AND J. JOUZEL. 1992. Surface water temperature changes in the high latitudes of the southern hemisphere over the last glacial-interglacial cycle. *Paleoceanography* **7**: 289–318.
- PIENITZ, R., J. P. SMOL, AND H. J. B. BIRKS. 1995. Assessment of freshwater diatoms as quantitative indicators of past climate change in the Yukon and Northwest Territories, Canada. *J. Paleolimnol.* **13**: 21–49.
- POPOVSKAYA, G. I. 1991. Baikal phytoplankton and its long-term observations (1958–1990) (in Russian). Professorship thesis, Biological Sciences, Novosibirsk.
- POPOVSKAYA, G. I. 1977. Dynamics of pelagic plankton 1964–1974, pp. 5–39 (in Russian). *In* M. Yu. Bekman [ed.], *Biological productivity of Baikal pelagic zone and its changeability*. Nauka.
- QUI, L., D. WILLIAMS, A. GVORZDKOV, E. KARABANOV, AND M. SHIMARAEVA, M. 1993. Biogenic silica accumulation and paleoproductivity in the northern basin of Lake Baikal during the Holocene. *Geology* **21**: 25–28.
- RICHARDSON, T. L., C. E. GIBSON, AND S. I. HEANEY. 2000. Temperature growth and seasonal succession of phytoplankton in Lake Baikal, Siberia. *Freshw. Biol.* **44**: 431–440.
- RYVES, D. B., S. JUGGINS, S. C. FRITZ, AND R. W. BATTARBEE. 2001. Experimental diatom dissolution and the quantification of microfossil preservation in sediments. *Palaeogeog. Palaeoclim. Palaeoecol.* **172**: 99–113.
- , D. H. JEWSON, M. STURM, R. W. BATTARBEE, R. J. FLOWER, A. W. MACKAY, AND N. G. GRANIN. In press. Quantitative and qualitative relationships between planktonic diatom communities and diatom assemblages in sedimenting material and surface sediments in Lake Baikal, Siberia. *Limnol. Oceanogr.*
- SEMOVSKI, S. V., N. YU. MOGILEV, AND P. P. SHERSTYANKIN. 2000. Lake Baikal ice. analysis of AVHRR imagery and simulation of under-ice phytoplankton bloom. *J. Mar. Syst.* **27**: 117–130.
- SHIMARAEV, M. N., V. I. VERBOLOV, N. G. GRANIN, AND P. P. SHERSTYANKIN. 1994. *In* M. N. Shimaraev and S. Okuda [eds.], *Physical limnology of Lake Baikal*. A review. BICER.
- SKABICHEVSKII, A. P. 1960. Planktonic diatoms of the freshwaters of the USSR: Systematics, ecology and distribution (in Russian). Moscow State University.
- SKVORTZOW, B. W. 1937. Diatoms from Lake Baikal. *Philipp. J. Sci.* **62**: 296–395.
- TARASOVA, E. N., AND A. I. MESCHERYAKOVA. 1992. Modern state of the hydrochemical regime of Lake Baikal (in Russian). Nauka.
- TER BRAAK, C. J. F., AND P. ŠMILAUER. 1998. CANOCO reference manual and user's guide to Canoco for windows: Software for canonical community ordination (v. 4). Microcomputer Power.
- VERBOLOV, V. I., V. N. SOKOL'NIKOV, AND M. N. SHIMARAEV. 1965. Hydrometeorological conditions and heat balance of Lake Baikal. Nauka.
- VOTINSEV, K. K., A. I. MESCHERYAKOVA, AND G. I. POPOVSKAYA. 1975. Organic matter turnover in Lake Baikal (in Russian). Novosibirsk.
- ZABELINA, M. M., I. A. KISELEV, A. I. PROSHKINA-LAVRENKO, AND V. S. SHEPUKOVA. 1951. *In* A.I. Proshkina-Lavrenko [ed.], *Diatom algae* (in Russian). Sovetsya Nauka.

Received: 21 January 2002

Accepted: 3 October 2002

Amended: 9 December 2002