BOREAL ENVIRONMENT RESEARCH 8: 215–228 ISSN 1239-6095
Helsinki 29 September 2003

A diatom-total phosphorus transfer function for freshwater lakes in southeastern Finland, including cross-validation with independent test lakes

Juha O. Miettinen

Karelian Institute, University of Joensuu, P.O. Box 111, FIN-80101 Joensuu, Finland

Miettinen, J. O. 2003: A diatom-total phosphorus transfer function for freshwater lakes in southeastern Finland, including cross-validation with independent test lakes. *Boreal Env. Res.* 8: 215–228. ISSN 1239-6095

A diatom total phosphorus (TP) training set and weighted averaging (WA) calibration model for lakes in eastern and southeastern Finland are presented. The data set comprises 78 lakes, with autumnal TP-range of 3–125 μg l^{-1}, pH-range of 5.9–7.7, and lake area of 0.13–283 km^2. The correlation between the observed and WA-reconstructed (jackknifed) log TP values in the training set is 0.86 (R^2 = 0.73), and root mean squared error of prediction (RMSEp) 0.193 μg TP l^{-1}. The transfer function was cross-validated using an independent test set of 30 lakes. The correlation between the observed and predicted log TP values for the 30 lakes is 0.89 (R^2 = 0.79), and RMSEp 0.161 μg TP l^{-1}. These performance statistics (R^2, RMSEp), based on the independent cross-validation with the test set of lakes, are within the range of values based on jackknifing estimates, suggesting that diatoms provide reliable estimates of lakewater TP concentration in Finnish lakes.

Introduction

Many Finnish lakes impacted by industrial and municipal effluents received elevated nutrient loads until the 1970s. Since the beginning of control on the point-sources of nutrients in the 1970s, these polluted lakes have generally been recovering (Wahlström *et al.* 1992). At the same time, however, diffuse nutrient loads have slowly degraded water quality of lakes in general (National Board of Waters and the Environment 1990). Agriculture has become the largest single source of total phosphorus, total nitrogen and biologically available fraction of phosphorus to surface waters (Rekolainen 1993). Nearly 3000 lakes (9%) in Finland had total phosphorus concentrations exceeding 35 μg l^{-1} in 1995, indicating eutrophic status (Mannio *et al.* 2000).

Several restoration projects for eutrophied lakes are being carried out in Finland. The determination of realistic target conditions for lake restoration requires knowledge of the degree of change over time. Development of multivariate statistical methods makes it possible to quantify

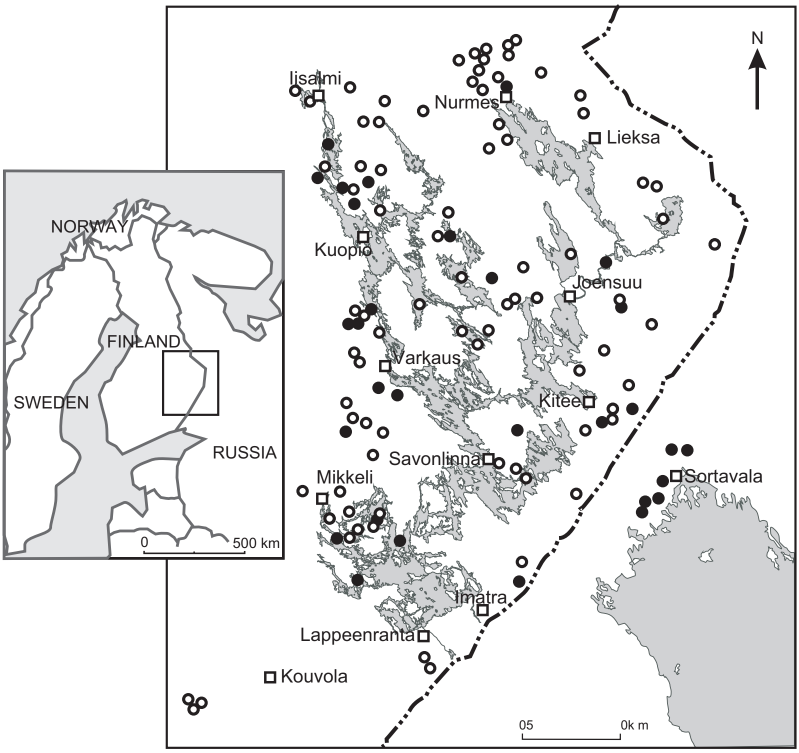


Fig. 1. Study area, with calibration lakes presented as open circles (N = 78) and test lakes as closed circles (N = 30).

the relationship between biological communities and environmental parameters in different ecosystems, and reconstructions of past environmental conditions on the basis of fossil species assemblages (Birks 1995).

Diatom frustules are generally well preserved in lake sediments, and they have been widely used as indicators of lake water quality in several studies (*see* Hall and Smol 1999). In this study, surface sediment diatom assemblages (0–1 cm interval) were used to assess the relationship between diatoms and total phosphorus concentration in the calibration set of lakes. The top or surface sediment typically contain diatoms that have been deposited over the last few years, thus their composition is not strongly influenced by the sampling time/season.

Several diatom-TP models are now available in Europe, e.g. Bennion *et al*. (1996), Lotter *et al*. (1998), Bradshaw and Anderson (2001) and Kauppila *et al*. (2002). The present model was constructed to fulfil the need of an appropriate model for the central-eastern part of Finland (Finnish lake district, described in Mannio *et al*. 2000) and Russian Karelia (Fig. 1). In particular, the model was developed for paleolimnological studies of lakes in the Sortavala region, SW Karelian Republic, Russia (Simola and Miettinen 1997, Simola *et al*. 2000). This new diatom-TP model extends the range of use for diatom based TP reconstructions in Finland, by containing large and europhic, and also concurrently dystrophic lakes in the calibration set. Of the preexisting diatom-TP models in Finland, the model by Turkia *et al*. (1998) is based on small, mostly oligotrophic forest lakes, and on the other hand, highly coloured lakes are excluded from the calibration set of Kauppila *et al*. (2002).

Independent cross-validation on a 30-lake test set was used to assess the performance of the diatom-TP transfer function, in addition to the statistical resampling approach (jackknife) on the calibration set. Cross-validation using only statistical resampling techniques for the calibration set is dependent on the statistical method used. Independent cross-validation is considered a more objective validation method, avoiding the pitfalls (Hämäläinen 2000) in error estimation on the basis of the training set alone.

Material and methods

Lake selection procedure

The samples were taken during two separate sampling periods, March–April 1999 and January–March 2001. The area of four Regional Environment Centres in eastern Finland (South Savo, North Savo, Southeast Finland and North Karelia), were used as the sampling area, with the exception of the three southernmost lakes, Pyhäjärvi (Artjärvi), Säyhtee and Viljakkalanjärvi (Fig. 1).

In 1999, 23 lakes chosen subjectively from eastern and southern Finland, were sampled for the calibration set, emphasizing eutrophic lakes to cover the TP range. In 2001, all the lakes with comparable water chemistry data (at least one TP value from 1 meter water depth, sampled in September–November, during the years 1998–2000) available in the Finnish Environment Institute database were sampled (81 lakes). In addition, six lakes in Russian Karelia, analyzed earlier by diatom stratigraphies, are used, summing the sampled lakes up to 110. All the lakes are situated between 60° and 64° northern latitude and between 26° and 32° eastern longitude, mainly within the River Vuoksi drainage area.

The lakes were divided into a calibration set and a test set. All the lakes sampled in 1999, were included in the 80-lake training set (23 lakes), and all the lakes sampled with piston corer for long cores were included in the 30-lake test set (8 lakes). The six lakes in Russian Karelia were included in the test set. The amount of lakes in the test set was decided to be 30, for statistical reasons, and so 16 additional lakes were picked randomly from the volume of 96 lakes remaining. The resulting test set of 30 lakes is considered to be independent from the calibration set. Two basins were later excluded from the calibration set, as they were considered not truly separate lakes from the Annilanselkä (no. 3 in the calibration set), so the final calibration set comprises 78 lakes.

Sampling

The sediments were cored with a Kajak gravity corer, with the exception of eight lakes cored with a piston corer (modification of Livingstone 1955), at or near the deepest point of the lake. Out of the 23 lakes cored in 1999, water chemistry data from the Finnish Environment Institute was obtained for 14 lakes, and water chemistry was sampled on 25 and 26 October 1999 for the remaining nine lakes (water analysis in Karelian Institute, University of Joensuu).

Only one set of representative measurements of chemical variables was available for most of the sites, and so one representative data point was selected for each site. The data points (dates) were selected according to the month of sampling preferring months in the order: October > September > November > December. Chemistry for all the sites was sampled from 1 m water depth in central parts of the lakes, using Ruttner-type sampler (length 0.5 m, volume 2 l). The water sampling points may be several kilometres apart from the sediment sampling points in some of the large lakes.

The water chemistry of the calibration set and test set includes both isothermal and stratified conditions. Of the calibration lakes, 55 were sampled during the autumn overturn of the water column (October–December, water temperature < 10 °C), and 22 were sampled in September–October before the autumn overturn. One lake, Tohmajärvi, was sampled in July 1999 (Table 1). Of the test lakes, 22 were sampled during the autumn overturn, seven during summer stratification, and one during winter stagnation.

Diatom identification

The sediment samples were treated using standard methods (Battarbee 1986), and at least 700 diatom valves were counted from each of the calibration samples, using a Leitz Dialux 20 EB microscope (phase contrast, 1000× magnification).

The taxonomy of Krammer and Lange-Bertalot (1986–1991) was followed, except for high-celled forms of *Aulacoseira* (*A. granulata, A. italica, A. subarctica*), which were harmonized with the calibration set of Kauppila *et al*. (2002), and for *Aulacoseira alpigena* and *Aulacoseira distans*, which were identified in accordance with Haworth (1988). In addition to the typical

Table 1. Location, chemical data, water depth at the sampling sites, and lake surface area for the 78 site data of Finnish lakes. Marking n.a. used, where data is not available.

Site no.	Site/statistic	Lat. N	Long. E	Date	TP (μg l^{-1})	Colour (mg Pt l^{-1})	pH	Cond. (mS m^{-1})	Depth (m)	Area (km^2)
1	Aittolampi	62°37´	29°57´	27.11.2000	6	5	6.9	5.4	9.1	0.48
2	Ala-Lyly	62°23´	27°27´	06.10.1999	19	30	7	6.8	4	0.55
3	Annilanselkä	61°39´	27°18´	25.10.1999	30	50	7.3	14.3	19.5	2
4	Etelä-Virmas	62°04´	27°36´	25.10.1999	18	80	6.8	4.3	5.2	11.8
5	Haapajärvi	63°36´	26°56´	09.10.2000	70	140	6.8	5.1	2.1	25.9
6	Haapajärvi	62°29´	29°54´	18.10.1999	5	n.a.	7.4	n.a.	18	1.37
7	Haapajärvi	63°39´	28°50´	29.09.1999	57	180	6.4	3	10	6.01
8	Haapajärvi	62°05´	27°22´	12.10.1999	32	140	6.5	4	6	5.24
9	Hanhijärvi	60°59´	28°11´	28.09.2000	38	35	7	11.6	2.5	4.45
10	Hyypiijärvi	62°04´	30°12´	25.10.1999	23	40	7	9.4	6.4	2.17
11	Höytiäinen	62°44´	29°46´	03.10.2000	5	30	6.7	5.2	45.6	283
12	Iso-Hietajärvi	63°10´	30°43´	05.10.2000	5	20	6.5	1.6	6	0.82
13	Iso-Lyly	63°00´	28°32´	13.10.1999	13	30	6.8	3.3	7.3	0.84
14	Jamalinjärvi	63°24´	29°57´	06.10.1998	14	100	6.3	2.4	12.5	1.22
15	Juojärvi	62°44´	28°37´	19.10.2000	4	n.a.	7.2	4.3	44	220
16	Juurusvesi	63°06´	27°46´	20.10.1999	17	n.a.	7.2	7.7	20.3	159
17	Karhunpää	63°36´	28°55´	25.10.1999	56	150	6.4	3.3	7.5	0.49
18	Kermajärvi	62°27´	28°40´	14.11.2000	6	20	6.7	4.5	53	85.6
19	Kesk. Sulkama	62°36´	29°11´	13.09.2000	60	n.a.	6.9	15.8	7.2	0.28
20	Keyritty	63°28´	28°16´	01.09.1999	41	130	6.3	2.1	9.6	18.3
21	Kinnasjärvi	62°30´	30°45´	08.11.2000	19	140	6.3	2.5	20.1	1.4
22	Kiteenjärvi	62°06´	30°11´	01.10.1999	28	50	7.3	10.4	5.4	15.1
23	Koirusjärvi	62°36´	27°32´	12.10.2000	18	50	7.2	5.2	22	42
24	Koitere	62°57´	30°39´	03.10.2000	8	70	6.6	1.8	29.5	164
25	Korpijärvi	61°14´	27°07´	20.09.1999	4	20	7.1	5.6	28.5	31.4
26	Kumpunen	63°04´	27°39´	20.10.1999	14	100	6.4	2.8	7.3	0.13
27	Kuohattijärvi	63°37´	29°29´	12.10.2000	11	60	6.4	1.8	13.2	10.8
28	Kuonanjärvi	61°58´	29°15´	07.10.1999	33	60	6.6	4.8	3	5.74
29	Kuorinka	62°37´	29°24´	10.11.2000	3	5	6.8	4.9	24.4	12.9
30	Kyyhkylänselkä	61°38´	27°18´	25.10.1999	30	50	7.3	13.6	12.5	1.75
31	Kärinki	61°17´	28°55´	09.10.2000	8	15	6.6	7.1	25	5.04
32	Lahnajärvi	60°47´	28°03´	03.09.1998	19	50	7.2	9.8	11.5	2.02
33	Leppäselkä	61°36´	27°21´	25.10.1999	15	25	7.3	10.5	16.1	2.14
34	Louhivesi	61°31´	27°22´	25.10.1999	8	20	7.3	10.1	31.5	16
35	Mensunlampi	62°16´	29°38´	01.11.2000	34	260	6.5	5.3	2.2	0.26
36	Mujejärvi	63°47´	29°33´	09.10.2000	20	n.a.	5.9	n.a.	14	3.51
37	Muntsurinjärvi	63°32´	29°59´	06.10.1998	23	210	6.4	2.3	4.6	0.83
38	Nerkoonjärvi	63°25´	28°18´	12.10.1999	67	80	7	5.1	10.7	15.6
39	Nuorajärvi	62°41´	31°09´	18.10.2000	18	140	6.3	1.8	5	40.2
40	Osmajärvi	62°26´	27°41´	18.10.1999	17	15	6.9	3.8	1.9	9.28
41	Pieni-Hietajärvi	63°10´	30°42´	05.10.2000	12	140	6	1.6	6	0.44
42	Pieni Rautjärvi	61°42´	29°45´	20.11.1996	28	40	7.2	14.5	5.5	4.8
43	Pieni Valtimojärvi	63°38´	28°52´	25.10.1999	52	120	6.3	2.7	4	0.74
44	Polvijärvi, Juuka	63°11´	29°00´	28.09.1999	11	80	6.3	1.9	27.8	1.84
45	Porovesi	63°32´	27°12´	16.11.1999	62	100	7.2	4.2	8.4	21.7
46	Pyhäjärvi, Artjärvi	60°43´	26°00´	08.11.2000	76	100	6.9	11.2	67	7.85
47	Pyhäjärvi, Kitee	62°02´	29°55´	17.10.2000	5	10	7.1	5.4	18	248
48	Reittiönjärvi	63°18´	27°55´	20.10.1999	27	100	6.8	6.3	11.9	0.14
49	Ruokojärvi	61°38´	28°23´	20.10.1999	10	40	6.5	6.4	15.5	1.62
50	Ruokonen	62°49´	29°31´	11.09.2000	19	n.a.	6.7	4.2	9.1	0.14
51	Saarilampi	62°35´	30°16´	01.11.2000	11	60	6.6	3.4	5.8	0.43
52	Sappulanselkä	61°34´	27°42´	09.09.1999	5	15	7.2	6.6	15.6	3.9

Continues

Table 1. Continued.

Site no.	Site/statistic	Lat. N	Long. E	Date	TP (μg l^{-1})	Colour (mg Pt l^{-1})	pH	Cond. (mS m^{-1})	Depth (m)	Area (km^2)
53	Siilinjärvi	63°04´	27°43´	18.10.2000	20	n.a.	7.3	10	27	2.83
54	Sonkajärvi	63°41´	27°31´	03.11.1999	47	n.a.	7	7	5.2	5.34
55	Sulkavanjärvi	63°07´	28°03´	18.10.2000	12	n.a.	7.6	9.6	17.3	8.2
56	Suomujärvi	63°08´	30°45´	11.10.1999	5	50	6.8	1.9	19	6.63
57	Suuri-Hautanen	62°23´	28°23´	14.10.1999	12	50	6.5	2.8	3.5	0.41
58	Suurijärvi	60°41´	27°54´	20.09.1999	6	30	6.8	2.4	8.8	16.2
59	Suuri Palojärvi	63°42´	29°18´	11.10.2000	25	n.a.	6.1	n.a.	6	1.1
60	Suvasvesi	62°35´	28°14´	23.10.2000	5	30	7.2	4.9	17.5	234
61	Lahnankuttava	62°27´	27°25´	16.10.2000	23	120	7	6.1	2	14.8
62	Säyhtee	60°45´	26°05´	19.10.1999	105	70	7.4	9.4	11.9	2.15
63	Tervajärvi	63°21´	29°11´	25.10.1999	38	100	6.8	4.6	10	0.18
64	Tohmajärvi	62°11´	30°23´	29.07.1999	19	80	7.2	8.8	13	12.1
65	Tuusjärvi	61°58´	28°02´	26.10.1999	12	60	7	6.6	17	15.6
66	Ukonvesi	61°36´	27°17´	25.10.1999	17	30	7.1	13.1	29	24.2
67	Unnukka	62°28´	27°51´	12.10.2000	17	30	7.2	5.3	30	80.5
68	Valtimojärvi	63°41´	28°50´	30.11.1993	54	160	5.9	3	6.3	4.01
69	Vestimo	63°33´	28°54´	25.10.1999	82	180	6.8	6.1	9	0.31
70	Viemenenjärvi	63°29´	28°58´	25.10.1999	43	150	6.2	6.2	9.6	1.83
71	Viinijärvi	62°45´	29°18´	02.10.2000	12	n.a.	7.1	6.1	28	135
72	Viljakkalanjärvi	60°46´	26°02´	25.10.1999	125	140	7.4	10.5	8.4	6.75
73	Virtasalmi	62°08´	27°27´	27.09.1999	25	100	6.6	4.8	2.3	5.9
74	Vuokko	63°24´	29°07´	25.10.1999	43	120	6.6	4	14.5	2.59
75	Ylikylänjärvi	63°38´	29°03´	25.10.1999	28	140	6.4	2.9	4.5	1.03
76	Ylimm. Sulkama	62°36´	29°10´	20.11.2000	37	100	6.7	11.2	9	0.31
77	Ylä-Siikajärvi	63°16´	28°20´	20.10.1999	24	80	6.8	4.2	15.5	2.21
78	Ylä-Valtimojärvi	63°42´	28°55´	30.11.1993	34	240	5.9	3	6.5	1.87
	Minimum				3	5	5.9	1.6	1.9	0.13
	Maximum				125	260	7.6	15.8	67	283
	Mean				27.0	81.6	6.79	5.95	14.2	26.7
	Median				19.0	70.0	6.80	5.10	10.0	4.23
	St. Deviation				23.6	58.4	0.40	3.51	12.1	60.5

Aulacoseira italica cells, thin and narrow valves with distinctive pseudoseptum (Ross *et al*. 1979), were counted separately as 'narrow' *A. italica*. Only taxa appearing in at least three samples, and exceeding 1% in relative abundance in at least one sample, were used for the ordinations and model development.

Ordinations

The chemistry values, except pH, were $\log_{10}$-transformed prior to all analyses as their distributions were found to be skewed (Table 1). Correlations between the environmental variables were checked, and no need of removing any variables was indicated.

Ordination methods assuming unimodality (correspondense analysis; CA, and canonical correpondense analysis; CCA) were applied to reveal the major patterns of variation in the data (ter Braak 1995). TP was used as the sole environmental variable in CCA, with Monte Carlo tests to evaluate the significance of the variable in explaining the variance in diatom data.

Eigenvalues of CA were compared with CCA for the evaluation of how much the measured environmental variables can explain of the variance in the species data. CANOCO 4.0 (ter Braak and Šmilauer 1998) was used for all ordinations, with square-root transformation of the species data. The calibration set lakes were screened for possible outliers using CCA. A lake was considered an outlier if it had an extreme score on any of the four axes of the CCA and a large (> 5%) residual distance to the TP-axis in CCA with TP as the sole variable (ter Braak & van Dam 1989).

Diatom-TP model development

Weighted averaging (WA; ter Braak and Van Dam 1989) with inverse and classical deshrinking, weighted averaging partial least squares (WA-PLS; ter Braak and Juggins 1993) and linear-based partial least squares (PLS) methods were compared for the regression and calibration of the TP-values for the calibration samples (Table 2). Program CALIBRATE version 0.81 (Juggins and ter Braak 1997) was used for the model development, with leave-one-out jackknifing as a cross-validation method for the models. To assess global and local systematic errors, mean and maximum bias, (ter Braak and Juggins 1993) were estimated by jackknifing.

The species data were square-root transformed for the model runs, as this was found to give lower prediction errors and higher correlations for the models, than non-transformed data. The tolerances are adjusted for the effective number of species (Hill's N^2) in the samples.

Test reconstructions

A minimum of 500 diatom valves were counted for the test samples. Taxonomy is consistent with the training set. The mean error statistics for the test set were calculated in the same way, as for the training set. The RMSEp (root mean squared error of prediction) is derived from the equation:

$$\mathrm{RMSE} = (s_1^{\,2} + s_2^{\,2}) \tag{1}$$

where s_1 is the mean and s_2 is the standard deviation of the residuals (estimated-observed TP values). Program WACALIB version 3.3 (Line *et al*. 1994) was used for the test reconstructions to obtain sample spesific error estimates (with 1000 bootstrap cycles).

The suitability of the diatom-TP model for calibration using test samples was assessed by Modern Analog Technique (MAT; Birks *et al*. 1990). The presence or absence of modern analogs for the test samples in the calibration set was examined using the MAT software of Juggins (1995). A test sample was considered to lack a good modern analog if the squared chord distance to its closest analog in the calibration set was longer than the distance at the fifth percentile of the distribution of dissimilarities between the calibration set samples.

Results

The diatom communities

The diatom data set was dominated by planktonic taxa: the mean abundance of planktonic (incl. tychoplankton) taxa was 75.3% and range 41.6%–94.9% in the calibration samples. Genera *Aulacoseira, Cyclotella, Tabellaria, Asterionella, Stephanodiscus, Cyclostephanos* and species *Fragilaria capucina, F. crotonensis, F. tenera* and

Table 2. Performance of the different log TP models based on the 78-site dataset. The performances are evaluated by apparent and cross-validated (jackknife) R^2 and root mean squared error (RMSE), both for the calibration set and for the test reconstructions. Mean and maximum bias statistics are based on the differences between inferred and observed TP in cross-validation for the calibration set. Inverse and classical deshrinking is compared for weighted averaging (WA).

Model	R^2 app.	RMSE app. (μg TP l^{-1})	R^2 jackknifed	RMSEp (μg TP l^{-1})	Mean bias (μg TP l^{-1})	Max. bias (μg TP l^{-1})	R^2 test, N = 30	RMSE test, N = 30 (μg TP l^{-1})
$WA_{inverse}$	0.81	0.164	0.73	0.193	0.001	0.222	0.79	0.161
$WA_{classical}$	0.81	0.182	0.74	0.202	0.002	0.161	0.79	0.194
WA-PLS (2 comp.)	0.87	0.133	0.74	0.192	0.001	0.232	0.71	0.201
PLS (3 comp.)	0.86	0.139	0.75	0.185	–0.001	0.179	0.78	0.168

F. ulna were classified as planktonic. Altogether 90 taxa with abundance exceeding 1% in at least one of the samples, and appearing in at least three lakes, were encountered in the calibration set.

Detrended correspondence analysis (DCA) resulted in gradient length of 2.04 for the first ordination axis (square-root transformation of species data, no downweighting of rare taxa), suggesting that both linear- and unimodal-based ordination methods would be appropriate (ter Braak & Prentice 1988). The configuration of the sites was very similar in the indirect CA and in CCA with all the six environmental variables. CA axis 1 eigenvalue was 0.205, and the ratio of the axis 1 eigenvalues CCA/CA 0.86. This small difference in the eigenvalues of CA and CCA suggests that the measured environmental variables offer a reasonable reflection of the main floristic gradients in the diatom data (ter Braak 1995).

The first two axes in the CCA had eigenvalues of 0.18 and 0.13. They explained 10.5% and 7.9% of the variance in the species data, respectively. Based on Monte Carlo tests (499 unrestricted permutations under full model), all the measured variables yielded a significant CCA axis, when other variables were entered as covariables ($p < 0.001$ for TP, lake area and depth, and $p < 0.01$ for conductivity, colour and pH). TP as the sole environmental variable explained 7.8% of the variance in the species data, and 4.6% when the other variables were entered as covariables.

No outliers were detected among the sites in the calibration set, according to the CCA. In CCA with all the environmental variables (Fig. 2), oligotrophic, clearwater lakes like Höytiäinen (no. 11), Kuorinka (29), Juojärvi (15), Kärinki (31) and Suvasvesi (60) are grouped at the low end of the phosphorus gradient (top in the Fig. 3b). The diatom taxa characterizing these lakes include *Cyclotella iris, C. bodanica, C. rossii,* and *Stephanodiscus rotula*. A number of oligotrophic lakes with relatively high humic content (colour), like Pieni-Hietajärvi (41), Suomujärvi (56), Nuorajärvi (39) and Suuri Palojärvi (59), are grouped at the low end of the conductivity and pH gradient (right in Fig. 3b). These lakes are characterized by the genera *Anomoeoneis* and *Eunotia,* as well as *Frustulia rhomboides* and *Aulacoseira tenella*.

The lakes Karhunpää (17), Keyritty (20), Valtimojärvi (68) and Vuokko (74) are examples of lakes more towards the upper end of the phosphorus gradient, and also characterized by high colour. The taxa most clearly associated with the combination of high colour and phophorus are *Stauroneis anceps, Eunotia zasuminensis, Pinnularia viridis* and genus *Tabellaria*. Lakes such as Pyhäjärvi (Artjärvi, 46), Sulkavanjärvi (55), Annilanselkä (3) and Siilinjärvi (53) are characterized by high phosphorus and high conductivity (left in Fig. 3b). These lakes are situated on fertile clayey or silty areas favourable for agriculture, and feature species including *Aulacoseira granulata*, *Cyclostephanos dubius, Stephanodiscus hantzschii* and *S. minutulus*.

The diatom-TP calibration model

The diatom inferred TP is strongly correlated with the observed TP in the training set (Fig. 3). The PLS model (three components) has the lowest RMSE and the highest correlation coefficient for the observed vs. predicted TP values (Table 2). WA with inverse deshrinking performs best with the independent test set.

Models without tolerance downweighting worked best with the WA. Inverse deshrinking was found superior to the classical deshrinking, according to the test reconstructions (Table 2). The equation for the inverse deshrinking is:

$$\text{Final value} = -6.650 + 6.201 \times \text{initial} \quad (2)$$

This WA model for log TP with inverse deshrinking applied to the calibration set resulted in apparent RMSE 0.164 μg TP l^{-1} and RMSEp (jackknife) 0.193 μg TP l^{-1}. The species optimum and tolerance values for the WA model are presented in the Appendix.

Test results

The WA model with inverse deshrinking resulted in RMSEp of 0.161 μg TP l^{-1}. The correlation between the estimated and observed log-transformed TP concentrations for the test lakes is

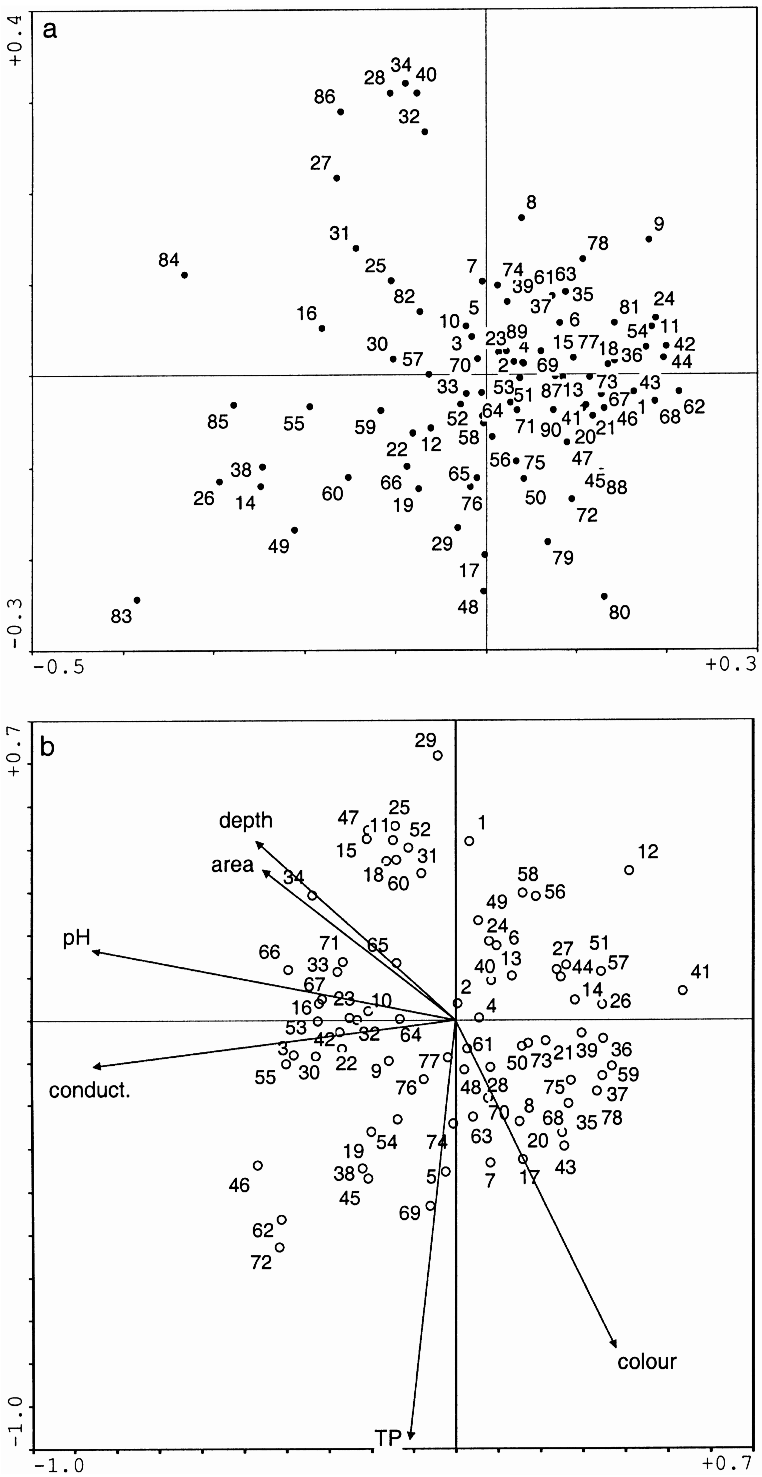
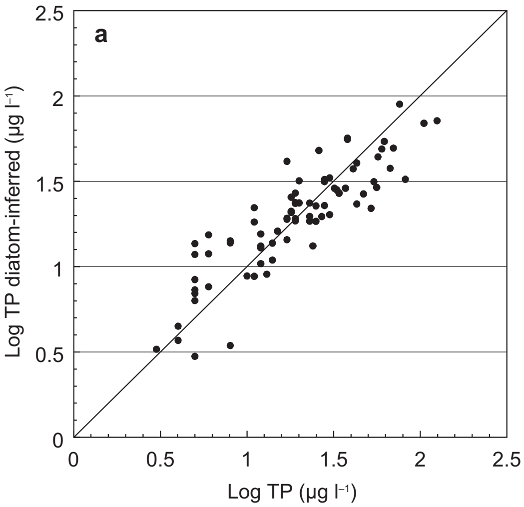


Fig. 2. CCA ordination plot, showing the ordination scores of (**a**) species (N = 90) and (**b**) sites (N = 78) and environmental variables on the ordination axes 1 and 2. Species are numbered according to the Appendix and sites according to Table 1.



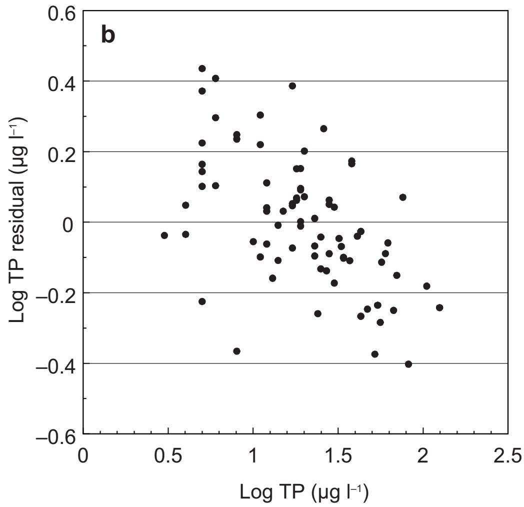


Fig. 3. Plot of observed log TP against (**a**) diatom-inferred TP and (**b**) predicted residuals, based on WA inverse regression of the 78 site dataset.

Table 3. Location, chemical data, diatom-inferred WA estimates for TP (backtransformed), water depth at the sampling sites, and lake surface area for the 30 site test set. Lakes are numbered in ascending order according to observed TP concentrations. Marking n.a. used, where data is not available.

	Lake	Lat. N	Long. E	Date	TP (μg l^{-1})	DI-TP (μg l^{-1})	Colour (mg Pt l^{-1})	pH	Cond. (mS m^{-1})	Depth (m)	Area (km^2)
1	Luonteri	61°38´	27°47´	23.11.2000	4	5.9	10	6.9	6.1	68.2	n.a.
2	Sorsavesi	62°31´	27°32´	12.10.2000	4	6.4	20	6.9	3.6	40	55.0
3	Nurmijärvi	61°23´	29°10´	19.10.2000	5	4.1	10	6.7	4.6	28	9.8
4	Iso-Heinäjärvi	62°05´	30°21´	03.11.2000	6	5.1	10	6.9	5.5	22	3.5
5	Koirusjärvi	62°36´	27°32´	13.09.2000	6	6.4	20	6.9	3.7	25	4.3
6	Kuolimo	61°21´	27°24´	10.10.2000	6	7.6	10	6.8	5.3	9.9	79.0
7	Lietvesi	61°30´	28°00´	08.09.1999	6	8.7	30	7.2	5	52	n.a.
8	Löytöjärvi	62°34´	30°19´	23.11.2000	8	6.8	40	6.7	2.9	10	1.3
9	Oravilahti	62°35´	27°38´	28.10.1999	10	13.3	15	7.3	13	20	n.a.
10	Ristijärvi	61°48´	30°46´	25.10.2000	10	8.0	50	6.7	4.9	24	1.5
11	Kuivinjärvi	62°51´	28°39´	13.10.1999	11	5.0	35	6.6	2.6	11.7	0.3
12	Ryttyjärvi	61°50´	30°42´	25.10.2000	13	6.1	35	6.7	4.5	10	1.0
13	Päähkeenselkä	61°34´	27°18´	25.10.1999	16	20.0	30	7.4	12.3	23.5	n.a.
14	Kannantakainen	61°53´	29°09´	25.10.1999	17	15.1	50	6.3	3.7	13.2	n.a.
15	Kuokkajärvi	61°39´	30°25´	26.10.2000	18	18.4	25	6.9	6.4	17.7	2.6
16	Lavijärvi	61°38´	30°30´	26.10.2000	19	16.7	20	7.1	6.5	21	2.7
17	Laukunlampi	62°40´	29°09´	10.01.1996	21	11.6	5	7	12.5	25	0.1
18	Hympölänjärvi	61°41´	30°39´	04.11.1999	22	19.8	70	7.1	7.4	16	5.4
19	Räimäjärvi	63°02´	27°37´	20.10.1999	22	23.0	30	7	10.6	25	1.3
20	Pitkäjärvi	61°37´	30°24´	26.10.2000	24	13.9	20	6.9	4.7	11	1.7
21	Ätäskö	62°01´	30°01´	17.10.2000	24	27.2	25	6.8	8.4	4	13.9
22	Lautiainen	63°35´	29°06´	05.09.2000	25	32.8	120	6.5	2.2	6	n.a.
23	Pöljänjärvi	63°08´	27°36´	25.10.2000	28	23.9	n.a.	7.1	n.a.	13.2	3.0
24	Kolmisoppi	63°09´	27°42´	18.10.2000	29	27.5	n.a.	7.7	2.1	9	0.5
25	Sääksjärvi	62°04´	27°56´	26.10.1999	31	21.4	80	6.5	5.2	2.5	2.7
26	Kuusjärvi	62°41´	28°56´	16.09.1999	37	24.4	160	6.87	6.8	3.6	1.0
27	Valvatus	62°13´	27°50´	19.09.2000	38	51.8	n.a.	7.1	14.3	6.5	3.0
28	Kangasjärvi	62°00´	35°20´	12.10.1999	41	18.8	150	6.3	3.6	5.7	22.0
29	Maaninkajärvi	63°11´	27°13´	17.11.1999	47	45.2	60	6.8	4.9	35	24.4
30	Onkivesi	63°15´	27°22´	21.10.1999	60	42.4	80	7	4.9	14.6	114

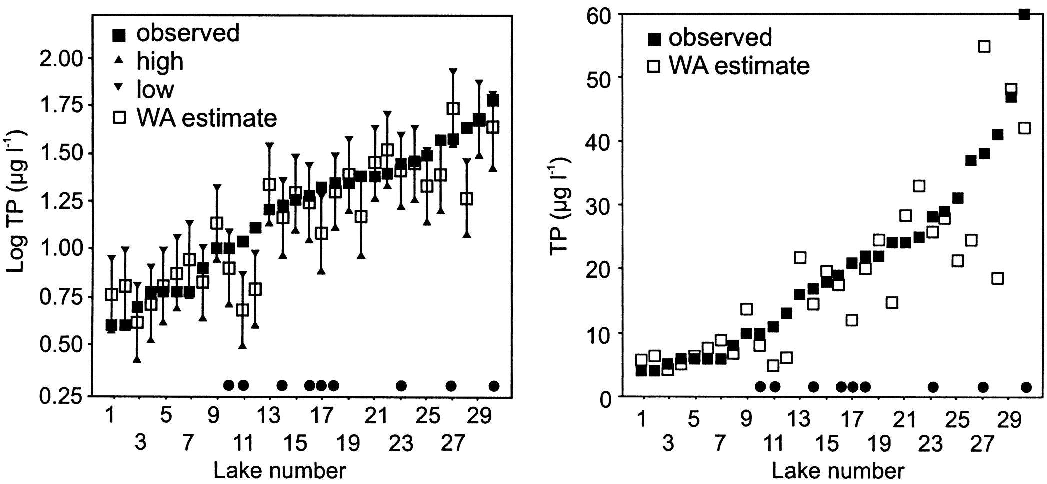


Fig. 4. Observed TP concentrations for the test lakes in ascending order, on log-transformed scale, with WA estimate, ± estimated standard error of prediction as confidence limits on the left, and on linear scale, with backtransformed WA estimates on the right. The WA- and SE-estimates are obtained by WACALIB v. 3.3, using bootstrapping as the cross-validation method for the predictions. Lakes lacking a good modern analog are indicated with dots.

0.89 ($R^2 = 0.79$). The lakes are presented in Table 3 and the deviation of the predicted TP from the observed concentrations is shown in Fig. 4.

Sample spesific error estimates (by WACALIB with bootstrapping as cross-validation method) were used for the test reconstructions in Fig 2. These sample spesific error estimates are larger (0.199–0.210 μg TP l^{-1}) than the RMSEp (jackknife) for the full test set (0.162 μg TP l^{-1}). Nine out of the 30 test samples were considered to lack a good modern analog. These test samples are marked with a dot at the bottom of Fig. 4. Four lakes have observed TP outside the range defined by reconstructed TP and error estimates (RMSEp). Two of these lakes (11, Kuivinjärvi; and 17, Laukunlampi) lack a good modern analog in the calibration set.

Combining of the training set with the test set did not substantially improve the error statistics: R^2 is 0.73 and RMSEp is 0.188 μg TP l^{-1} for the 108 lake calibration set, when using the best model in this case (WA-PLS). The PLS model (3 components) has the lowest RMSE and the highest correlation coefficient for the observed vs. predicted TP values among the calibration set lakes (Table 2). However, the difference between PLS and WA (with inverse deshrinking) is small for the jackknifed predictions (RMSEp 0.185 vs. 0.193 μg TP l^{-1}). WA is simple, and performs best with the independent test set. Both the PLS and WA underestimate the TP measured at high levels (Fig. 4). Based on these results, WA with inverse deshrinking is considered the best method in this case.

Discussion

The constructed diatom-TP model is robust, as judged by the stability of the error estimates (Table 2), which are comparable to those obtained by other diatom-TP models in Europe and in Canada (e.g. Wunsam and Schmidt 1995, Hall and Smol 1996, Bennion *et al.* 1996, Lotter *et al.* 1998, Bradshaw and Anderson 2001, Reavie and Smol 2001, Enache and Prairie 2002, Kauppila *et al.* 2002). Cross-validated (jackknife or bootstrap) R^2 values range from 0.41 to 0.91 for these models. Five of the eight models use log-transformed TP values, with RMSEp ranging from 0.16 to 0.34 μg TP l^{-1}.

The diatom-TP transfer function failed to reconstruct present TP concentrations in four lakes of the 30-lake test set, based on the measured TP concentrations in the lakes. Out of the four lakes, two are lacking modern analog in the calibration set (according to the MAT with fifth percentile as the exclusion limit), and the transfer function may be considered unsuitable for these lakes. Presence of *Fragilaria*-species, with inferred optimum TP values of 19.3 μg TP l^{-1} for *F. capucina* and 21.4 μg TP l^{-1} for *F. construens*, cause over-prediction of TP for these two small, deep, and oligotrophic clearwater lakes (11, Kuivinjärvi; and 17, Laukunlampi).

The use of a single chemical measurement as the reference for environmental reconstruction is problematic, as is the case in this study. The failing of the transfer function for two of the test lakes (12, Ryttyjärvi; and 28, Kangasjärvi) with good modern analogs in the calibration set does not necessarily mean unfitness of the calibration set for these lakes, but may partly result from unrepresentative single TP measurement.

The constructed inference model seems to result in larger errors at high TP values. In the test set, Pearson correlation is 0.37 for observed TP vs. error ($p < 0.05$). This may result from increasing variation in lake TP with increasing mean TP concentrations. The deviation between the observed and inferred TP values increase also for the calibration set lakes with TP on the range 3–60 μg TP l^{-1} (Pearson correlation 0.50, $p < 0.01$, $N = 71$).

The model tends to underestimate TP concentrations at TP values over 60 μg TP l^{-1}. Due to the random selection of the lakes, the model is validated only for the TP range up to 60 μg l^{-1} with the independent test set. When lakes with TP > 60 μg l^{-1} (seven lakes) were deleted from the calibration set, WA model for log TP with inverse deshrinking resulted in a slightly smaller RMSEp for the calibration lakes (0.175 μg TP l^{-1}), and for the test set (0.147 μg TP l^{-1}), than with the full calibration set of 78 lakes. However, the full 78-lake set is presented here, because it is more widely applicable across the ecological and chemical range of lakes in the eastern and southeastern Finland.

Jackknifing can give over-optimistic results, if there are similar, interrelated, samples in the calibration set (Hämäläinen 2000). In this case the test set is used to validate the error estimates derived by jackknifing. The RMSEp is lower for the independent test set than the jackknifed RMSEp for the calibration set, suggesting that the jackknifed error estimates are realistic for the calibration set, and that the diatoms provide reliable estimates of the past and present TP concentrations in the lakes of the Finnish lake district.

These methods of diatom-TP model development and validation have now been well tested. TP reconstructions from sediment cores have been validated by comparison with monitoring data (e.g. Bennion *et al*. 1995, Bradshaw and Anderson 2001, Kauppila *et al*. 2002). As a novel aspect, this study presents an evaluation of the diatom-TP transfer function by cross-validation using an independent 30-lake test set. Bigler and Hall (2002) used a 15-lake test set to validate pH and temperature transfer functions in Sweden, but the test lakes belonged to the calibration set, and were used as an intra-set cross-validation exercise.

The developed transfer function is suitable for oligo-, eu- and dystrophic lakes, but not for hypertrophic lakes. Many oligotrophic and slightly mesotrophic (< 15 μg TP l^{-1}) lakes in the calibration set (Table 1) allow reconstruction of subtle shifts in nutrient concentrations of currently oligotrophic lakes. Palaeolimnological techniques may have a special role in such waters, where changes in the diatom assemblages from dated sediment cores may indicate nutrient enrichment before any increase in TP concentrations can be detected using standard water chemistry techniques (Bennion *et al*. 1996). In management of eutrophied lakes, the transfer function approach provides an estimate of baseline conditions, helping to define whether a pollution problem exists and to set realistic targets for restoration.

Acknowledgements: I am thankful to Heikki Simola for initiating this project, supervising and corrections to the paper. Thanks to Heikki Hämäläinen for examining the model and corrections. Hannu Tarnanen did much of the field work and Kirsti Kyyrönen finished the figures. Maj and Tor Nessling Foundation funded the project. This paper has benefited from thorough reviews of three anonymous referees.

References

Battarbee R.W. 1986. Diatom analysis. In: Berglund B.E. (ed.), *Handbook of Holocene palaeoecology and palaeohydrology*. Wiley, Chichester, pp. 527–570

Bennion H., Wunsam S. & Schmidt R. 1995. The validation of diatom-phosphorus transfer functions: an example from Mondsee, Austria. *Freshwater biology* 34: 271–283.

Bennion H., Juggins S. & Anderson N.J. 1996. Predicting epilimnetic phosphorus concentrations using an improved diatom-based transfer function and its applica-

tion to lake eutrophication management. *Environmental Science and Technology* 30: 2004–2007.
Bigler C. & Hall R.I. 2002. Diatoms as indicators of climatic and limnological change in Swedish Lapland: a 100-lake calibration set and its validation for paleoecological reconstructions. *J. Paleolimnol*. 27: 97–115.
Birks H.J.B. 1995. Quantitative palaeoenvironmental reconstructions. In: Maddy D. & Brew S. (eds.), *Statistical modelling of Quaternary science data*. Quaternary Research Association, Cambridge, Technical Guide 5, pp. 161–254.
Birks H.J.B., Line J.M., Juggins S., Stevenson A.C. & ter Braak C.J.F. 1990. Diatoms and pH reconstruction. *Phil. Trans. R. Soc. Lond. B* 327: 263–278.
Bradshaw E.G. & Anderson N.J. 2001. Validation of a diatom-phosphorus calibration set for Sweden. *Freshwater Biology* 46: 1035–1048.
Enache M. & Prairie Y.T. 2002. WA-PLS diatom-based pH, TP and DOC inference models from 42 lakes in the Abitibi clay belt area (Quebec, Canada). *J. Paleolimnol*. 27: 151–171.
Hall R.I. & Smol J.P. 1996. Paleolimnological assessment of long-term water-quality changes in south-central Ontario lakes affected by cottage development and acidification. *Can. J. Fish. Aquat. Sci*. 53: 1–17.
Hall R.I. & Smol J.P. (eds.) 1999. *The diatoms: Applications for the environmental and earth sciences*. Cambridge University Press, 469 pp.
Haworth E.Y. 1988. Distribution of diatom taxa of the old genus *Melosira* (now mainly *Aulacoseira*) in Cumbrian waters. In: Round F.E. (ed.), *Algae and the aquatic environment*. Biopress Ltd., Bristol, pp. 138–167.
Hämäläinen H. 2000. *Weighted averaging models in contemporary freshwater monitoring with benthic invertebrate assemblages*. University of Joensuu, Ph.D. Dissertations in Biology, No. 2, 102 pp.
Juggins S. 1995. *Program MAT — Modern Analog Technique. Version 1.1*.
Juggins S. & ter Braak C.J.F. 1997. *CALIBRATE — a computer program for species–environmental calibration by [Weighted-Averaging] Partial Least Squares Regression (0.81)*. University of Newcastle, Newcastle.
Kauppila T., Moisio T. & Salonen V.-P. 2002. A diatom-based inference model for autumn epilimnetic total phosphorus concentration and its application to a presently eutrophic boreal lake. *Journal of Paleolimnology* 27: 261–273.
Krammer K. & Lange-Bertalot H. 1986–1991. Bacillariophyceae. Teil 1–4. *Süsswasserflora von Mitteleuropa*, Band 4/1-4. G. Fischer Verlag, Stuttgart.
Line J.M., ter Braak C.J.F. & Birks H.J.B. 1994. WACALIB, Version 3.3 — a computer program to reconstruct environmental variables from fossil assemblages by weighted averaging and to derive sample-spesific errors of prediction. *J. Paleolimnol*. 10: 147–152.
Livingstone D.A. 1955. A lightweight piston sampler for lake deposits. *Ecology* 36: 137–139.
Lotter A.F., Birks H., 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.
Mannio J., Räike A. & Vuorenmaa J. 2000. Finnish lake survey 1995: regional characteristics of lake chemistry. *Verh. Internat. Verein. Limnol*. 27: 362–367.
National Board of Waters and the Environment 1990. *Vähittäinen rehevöityminen ja kemikalisoituminen uhkaavat pintavesien tilaa*. Tiedote 29.6.1990.
Reavie E.D. & Smol J.P. 2001. Diatom-environmental relationships in 64 alkaline southeastern Ontario (Canada) lakes: a diatom-based model for water quality reconstructions. *J. Paleolimnol*. 25: 25–42.
Rekolainen S. 1993. *Assessment and mitigation of agricultural water pollution*. Publications of the Water and Environment Research Institute. National Board of Waters and the Environment, Finland. No. 12. 34 pp.
Ross R., Cox E.J. & Karayeva N.I. 1979. An amended terminology for the siliceous components of the diatom cell. *Proceedings of the fifth symposium on recent and fossil diatoms, Antwerp 1978*. Cramer, Vaduz.
Simola H. & Miettinen J. 1997. Mitä tapahtui kun väestö evakuoitiin? — Luonnontilan palautumisprosessi on tallentunut luovutetun Karjalan järvien sedimentteihin. *Geologi* 49: 86–89.
Simola H., Miettinen J., Grönlund E. & Saksa A. 2000. Palaeoecological reflections of rapid land-use change in the northern Lake Ladoga region. In: Peltola A., Grönlund E. & Viljanen M. (eds.), *Proceedings of the Third International Lake Ladoga Symposium, Petrozavodsk, 1999*. University of Joensuu, Publications of Karelian Institute 129, pp. 504–507.
ter Braak C.J.F. 1995. Ordination. In: Jongman R.H.G., ter Braak C.J.F. & Van Tongeren O.F.R. (eds.), *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, UK. pp. 91–173.
ter Braak C.J.F. & Prentice I.C. 1988. A theory of gradient analysis. *Adv. Ecol. Res*. 18: 271–317.
ter Braak C.J.F. & van Dam H. 1989. Inferring pH from diatoms: a comparison of old and new calibration methods. *Hydrobiologia* 178: 209–223.
ter Braak C.J.F. & Juggins S. 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia* 269/270: 485–502.
ter Braak C.J.F. & Šmilauer P. 1998. *Canoco for Windows (4.0)*. Centre for Biometry, Wageningen.
Turkia J., Sandman O. & Huttunen P. 1998. Palaeolimnological evidence of forestry practices disturbing small lakes in Finland. *Boreal Env. Res*. 3: 45–61.
Wahlström E., Reinikainen T. & Hallanaro E.-L. (eds.) 1992. *Ympäristön tila Suomessa*. Ympäristötietokeskus Gaudeamus, Helsinki, 364 pp.
Wunsam S. & Schmidt R. 1995. A diatom-phosphorous transfer function for Alpine and pre-alpine lakes. *Memorie dell'Istituto Italiano di Idrobiologia* 53: 85–99.

Received 4 January 2002, accepted 1 April 2003

Appendix. Jackknifed species optima and tolerances for simple WA model, optima backtransformed to linear scale, and number of occurrences for the taxa in the calibration set.

No.	Taxon	Log TP optimum (μg l^{-1})	Log TP tolerance (μg l^{-1})	TP optimum (μg l^{-1})	Number occur.
1	*Achnanthes impexa*	1.285	0.307	19.3	26
2	*Achnanthes laevis*	1.244	0.264	17.5	33
3	*Achnanthes levanderi*	1.257	0.347	18.1	71
4	*Achnanthes linearis*	1.258	0.345	18.1	54
5	*Achnanthes minutissima*	1.198	0.398	15.8	76
6	*Achnanthes subatomoides*	1.167	0.347	14.7	28
7	*Amphora copulata*	1.150	0.378	14.1	23
8	*Anomoeoneis brachysira*	1.021	0.337	10.5	38
9	*Anomoeoneis serians*	0.990	0.237	9.8	5
10	*Asterionella formosa*	1.216	0.357	16.4	75
11	*Aulacoseira alpigena*	1.128	0.225	13.4	3
12	*Aulacoseira ambigua*	1.369	0.322	23.4	76
13	*Aulacoseira distans*	1.266	0.352	18.5	64
14	*Aulacoseira granulata*	1.486	0.295	30.6	21
15	*Aulacoseira humilis*	1.222	0.357	16.7	12
16	*Aulacoseira islandica*	1.205	0.314	16.0	33
17	*Aulacoseira italica*	1.583	0.373	38.3	14
18	*Aulacoseira valida*	1.248	0.408	17.7	35
19	*Aulacoseira italica ‘narrow’*	1.469	0.369	29.4	59
20	*Aulacoseira lacustris*	1.337	0.27	21.7	30
21	*Aulacoseira lirata*	1.333	0.315	21.6	50
22	*Aulacoseira subarctica tall*	1.377	0.402	23.8	65
23	*Aulacoseira subarctica short*	1.234	0.385	17.1	73
24	*Aulacoseira distans* var. *tenella*	1.140	0.280	13.8	28
25	*Cocconeis placentula*	1.155	0.445	14.3	7
26	*Cyclostephanos dubius*	1.464	0.279	29.1	19
27	*Cyclotella bodanica*	1.009	0.367	10.2	16
28	*Cyclotella iris*	0.874	0.263	7.5	19
29	*Cyclotella glomerata*	1.506	0.228	32.1	12
30	*Cyclotella pseudostelligera*	1.266	0.456	18.5	31
31	*Cyclotella radiosa*	1.096	0.310	12.5	44
32	*Cyclotella rossii*	0.893	0.289	7.9	45
33	*Cyclotella stelligera*	1.311	0.333	20.7	64
34	*Cymbella descripta*	0.816	0.279	6.5	12
35	*Cymbella naviculiformis*	1.148	0.403	14.1	28
36	*Cymbella silesiaca*	1.235	0.372	17.2	43
37	*Denticula tenuis*	1.255	0.427	18.0	18
38	*Diatoma tenuis*	1.461	0.413	29.0	18
39	*Diploneis elliptica*	1.155	0.521	14.3	5
40	*Epithemia adnata*	0.856	0.195	7.2	9
41	*Eunotia bilunaris*	1.313	0.349	20.6	40
42	*Eunotia faba*	1.193	0.295	15.6	15
43	*Eunotia incisa*	1.281	0.296	19.2	45
44	*Eunotia meisteri*	1.217	0.335	16.5	25
45	*Eunotia paludosa*	1.397	0.353	24.9	4
46	Eunotia pectinalis	1.310	0.358	20.4	46
47	*Eunotia praerupta*	1.379	0.326	23.9	29
48	*Eunotia zasuminensis*	1.613	0.098	41.0	4
49	*Fragilaria berolinensis*	1.615	0.308	41.2	5
50	*Fragilaria brevistriata*	1.446	0.289	28.0	41

Continues

Appendix. Continued.

No.	Taxon	Log TP optimum ($\mu g\ l^{-1}$)	Log TP tolerance ($\mu g\ l^{-1}$)	TP optimum ($\mu g\ l^{-1}$)	Number occur.
51	*Fragilaria capucina*	1.285	0.343	19.3	71
52	*Fragilaria capucina* var. *gracilis*	1.328	0.391	21.3	57
53	*Fragilaria construens*	1.331	0.324	21.4	72
54	*Fragilaria constricta*	1.199	0.369	15.8	17
55	*Fragilaria crotonensis*	1.366	0.301	23.2	51
56	*Fragilaria parasitica* var. *subconstricta*	1.406	0.317	25.5	31
57	*Fragilaria pinnata*	1.282	0.332	19.1	31
58	*Fragilaria rumpens*	1.375	0.189	23.7	5
59	*Fragilaria tenera*	1.355	0.407	22.6	32
60	*Fragilaria ulna* var. *acus*	1.451	0.360	28.4	26
61	*Fragilaria virescens*	1.226	0.300	16.8	44
62	*Frustulia rhomboides*	1.271	0.338	18.7	37
63	*Gomphonema acuminatum*	1.148	0.365	14.1	19
64	*Gomphonema parvulum*	1.354	0.39	22.6	55
65	*Hantzschia amphioxys*	1.442	0.301	27.7	19
66	*Navicula elginensis*	1.435	0.221	27.2	13
67	*Navicula impexa*	1.293	0.383	19.6	21
68	*Navicula mediocris*	1.302	0.315	20.0	24
69	*Navicula pupula*	1.266	0.334	18.5	47
70	*Navicula radiosa*	1.298	0.380	20.0	42
71	*Navicula radiosa* var. *tenella*	1.333	0.404	21.5	48
72	*Navicula schoenfeldii*	1.441	0.356	27.6	7
73	*Neidium ampliatum*	1.283	0.348	19.2	22
74	*Nitzschia fonticola*	1.130	0.395	13.5	41
75	*Nitzschia palea*	1.417	0.363	26.3	25
76	*Opephora olsenii*	1.500	0.316	31.6	13
77	*Pinnularia gibba*	1.240	0.357	17.5	66
78	*Pinnularia maior*	1.128	0.418	13.4	10
79	*Pinnularia viridis*	1.542	0.291	34.8	7
80	*Stauroneis anceps*	1.649	0.152	44.6	6
81	*Stauroneis phoenicenteron*	1.173	0.313	14.9	39
82	*Stephanodiscus alpinus*	1.051	0.409	11.2	5
83	*Stephanodiscus hantzschii*	1.712	0.464	51.5	7
84	*Stephanodiscus minutulus*	1.154	0.256	14.3	4
85	*Stephanodiscus neoastraea*	1.380	0.418	24.1	19
86	*Stephanodiscus rotula*	0.886	0.312	7.7	13
87	*Surirella ovalis*	1.285	0.335	19.3	9
88	*Tabellaria fenestrata*	1.421	0.280	26.4	37
89	*Tabellaria flocculosa*	1.232	0.342	17.5	76
90	*Tabellaria flocculosa* ('type III')	1.329	0.343	21.5	50