

**DIATOM-INFERRED TOTAL PHOSPHORUS CONCENTRATIONS IN WILLIAMS
LAKE (BRITISH COLUMBIA) DURING THE PAST 150 YEARS.**

A REPORT PREPARED FOR THE MINISTRY OF THE ENVIRONMENT,
BRITISH COLUMBIA

BY

ROLAND I. HALL and Dr. JOHN P. SMOL
Paleoecological Environmental Assessment and Research Lab (PEARL)
Department of Biology,
Queen's University, Kingston
Ontario, K7L 3N6.

March 26, 1992

TABLE OF CONTENTS

SUMMARY	3
INTRODUCTION	4
MATERIALS AND METHODS	5
RESULTS	6
DISCUSSION	8
Limitations of the Diatom-Inference Model	8
Recent Total Phosphorus Trends	10
LITERATURE CITED	13
Figure 1	14
Figure 2	16
APPENDIX A -- Attached at end.	
APPENDIX B -- Attached at end.	

SUMMARY

1. We enumerated the diatoms in a sediment core from Williams Lake that includes the past 150 years. The diatom assemblage indicates that the lake was eutrophic prior to European settlement.
2. Our weighted averaging regression and calibration model to used to infer past total phosphorus concentrations in Williams Lake from the relative abundances of diatom taxa in the sediment core. Despite some changes in the composition of the diatom community, very little change in total phosphorus is inferred throughout the past 150 years. The present total phosphorus concentration (35-85 $\mu\text{g/L}$) of Williams Lake is greater than the range included in our model (5-28 $\mu\text{g/L}$) and the surface sediment diatom assemblage infers a much lower total phosphorus concentration (13.8 $\mu\text{g/L}$). We therefore believe the inferred values cannot be an accurate reflection of the true values. We will continue to refine our model so that it can provide more sensitive and accurate total phosphorus inferences in eutrophic lakes.
3. We conclude that our model may not have been able to detect past eutrophication events in Williams Lake because further enrichment may have caused the algal community to shift to other groups (e.g. blue-green algae), rather than causing shifts in the taxonomic composition of the diatom community. Dr. Peter Leavitt (University of Alberta) is presently analyzing the algal pigment record as an means of deciphering such events. The results will be forwarded to the B.C. Ministry of the Environment as soon as they become available.
4. Diatom taxonomic information contained in the Williams Lake sediment core indicates that a limnological change occurred in the early 1900's (ca. 1917). The diatom changes reflect lake eutrophication that typically occurs in response to human disturbances and agriculture.
5. In order to enhance the ability of our total phosphorus inference model to accurately hindcast extremely eutrophic lake conditions, we plan to add surface sediment diatom data and environmental data from 10 to 12 more lakes (within the range 28-85 $\mu\text{g/L}$) to our calibration lake set. This will provide better estimates of the total phosphorus optima for the diatom taxa, and improve the accuracy of total phosphorus estimates at the high end of the trophic status spectrum.

INTRODUCTION

Diatoms (a group of unicellular, siliceous algae belonging to the class Bacillariophyceae) are sensitive ecological indicators of lakewater chemistry (e.g. pH, phosphorus, salinity). Diatoms have been widely used in studies of lake eutrophication because: 1) they have restricted nutrient requirements and are sensitive indicators of lake trophic status; 2) their siliceous cell walls (or "valves") are abundant and well preserved in most lake sediments and therefore provide an "archive" of information (integrated in both space and time) of past changes in lake conditions that are otherwise unavailable or difficult to obtain; and 3) sediment diatom assemblages are composed of a large number of taxa that can be identified to the species or sub-species level, and hence contain considerable ecological information.

Previous studies have used diatoms to infer past changes in lake trophic status, but have largely been restricted to qualitative descriptions of eutrophication. We have been developing a more objective and quantitative method for inferring lake trophic status from diatoms preserved in the sediments of British Columbia lakes. Recent research (Hall and Smol 1992; presented in Appendix A) demonstrates a strong relationship between the distribution of surface sediment diatom taxa and lakewater total phosphorus concentration (TP) in a suite of 41 B.C. lakes. We have developed a model that can infer lakewater TP from the percent abundance of 131 diatom taxa in lake sediment diatom assemblages. The model is based on weighted averaging computational procedures which briefly involve two steps. First, the TP optima (weighted averages) of individual diatom taxa are calculated in a regression step based on their abundances in surface sediment samples from a number of lakes. Second, these estimated optima are used to infer past TP from the composition of fossil diatom assemblages. Our model is now available

for inferring past lakewater TP within the range 5 - 28 $\mu\text{g/L}$.

In a previous report, we inferred the recent TP history of Kamloops lake (March, 1991). Despite some changes in the diatom assemblages during the period included in the sediment core, the model inferred little or no change in the TP concentration. The diatom inferred TP estimate for the surface sediment sample (8.7 $\mu\text{g/L}$) closely agreed with average measured water chemistry data (8.8 $\mu\text{g/L}$, Colin McKean pers. comm.).

The objective of the present study is to infer the TP concentration of Williams Lake, B.C. during the past 150 years, using our weighted averaging inference model. This report also documents shifts in the most common diatom taxa in the sediment core, and discusses the limitations of our model in this lake. The TP concentration of Williams Lake (spring overturn TP = 30 - 85 $\mu\text{g/L}$ during the period 1973 - 1983; Zirnhelt 1984) is outside the range that our model can infer (5 - 28 $\mu\text{g/L}$). We therefore anticipate that our model will not be able to accurately infer the absolute TP concentration values of Williams Lake. However, we expect the model to demonstrate eutrophication events and indicate their timing.

MATERIALS AND METHODS

A 42 cm long sediment core was retrieved from Williams Lake on March 20, 1991 using a modified Kajak-Brinkhurst gravity corer. In order to minimize sediment disturbance, the core was sectioned at the lake shore immediately after retrieval. The top 15 cm of the sediment core were sectioned into 0.5 cm intervals. The remainder of the core was sectioned into 1.0 cm intervals, except for the sequence between 29 and 32 cm depth which was sectioned as follows;

29 - 30.5 cm, 30.5 - 31.5 cm, and 31.5 - 32.0 cm.

Diatom slides were prepared from 12 sediment samples throughout the length of the core using standard techniques (Smol 1983). Briefly, small aliquots (0.5 ml) of sediment were digested in strong acid (potassium dichromate and sulfuric acid, 5% w/v) to remove the organic material, and the resulting slurries were centrifuged and rinsed at least 8 times with distilled water. The cleaned diatom suspensions were pipetted onto coverslips and dried on a warming tray. The coverslips were mounted in Hyrax[®] onto glass microscope slides. A minimum of 500 diatom valves were identified and counted from each sample under oil immersion using an Olympus BHS microscope (1000x magnification, numerical aperture = 1.40) with bright field optics.

Lakewater TP inferences were generated from the diatom data using our British Columbia calibration lake set. Full details of the computational procedures are presented in our publication (Hall and Smol 1992; see Appendix A). We used the computer program WACALIB version 2.1 (Birks and Line 1990).

²¹⁰Pb was analyzed in 13 sediment samples throughout the length of the core by the Chalk River Laboratories of Atomic Energy of Canada Limited (AECL). We calculated the ²¹⁰Pb dates using the constant rate of supply model (CRS model; Binford 1990).

RESULTS

²¹⁰Pb analyses indicate background ²¹⁰Pb levels were attained at the base of the sediment core, and that the 42 cm sediment sequence encompasses the full period of European settlement (pre-1858 to present; Fig. 1,2).

According to our diatom-inference model, TP concentrations in Williams Lake have changed very little during the past 150 years (Fig. 1, Table 1). Diatom-inferred TP varied from 12.4 to 14.4 $\mu\text{g/L}$, with the lowest value corresponding to the 1842 ^{210}Pb date, and the highest value occurring in 1969. TP concentrations before 1956 are consistently lower than those since 1956 (Fig. 1). A period of TP enrichment occurs between the years 1956 and 1969, and TP concentrations have remained slightly elevated in the lake since 1956.

Table 1. Diatom-inferred total phosphorus concentrations for selected levels in the Williams lake sediment core. ^{210}Pb dates and the number of diatom taxa included in the weighted averaging calibration equations are also presented.

^{210}Pb Date	Sediment (cm)	Depth	Diatom-Inferred ($\mu\text{g/L}$)	TP	# Species included in WA model
1990		0.25		13.8	22
1988		1.25		14.2	25
1986		2.25		14.2	26
1980		4.25		14.1	26
1975		6.25		13.4	25
1969		8.25		14.4	27
1956		12.25		13.3	27
1936		16.50		13.3	24
1917		20.50		12.7	29
1884		26.50		12.5	29
1842		32.50		12.4	33
< 1842		39.00		12.9	18

The diatom assemblages at all levels of the Williams Lake sediment core are dominated by the small centric diatom Stephanodiscus minutulus (Fig. 2). S. minutulus has a relatively high optimum (or weighted average (WA)) for TP as compared to other diatom taxa in our data set (WA = 12.3 $\mu\text{g/L}$; see Appendix B), and tends to indicate that Williams Lake has always been eutrophic, both prior to and during the period of European settlement. S. minutulus decreases in relative abundance during the period 1917 - 1980, and increases again towards the surface (1980 - 1990). Two small Stephanodiscus species (S. hantschii (WA = 12.5 $\mu\text{g/L}$) and S. parvus (WA = 11.9 $\mu\text{g/L}$)) increase in relative abundance during the period of S. minutulus decline (between 1917 and 1980). Two planktonic diatom species (Asterionella formosa (WA = 11.8 $\mu\text{g/L}$) and Fragilaria crotonensis (WA = 13.9 $\mu\text{g/L}$)) increase in relative abundance after 1917, while Cyclotella michiganiana (WA = 10.1 $\mu\text{g/L}$) and Fragilaria brevistriata (WA = 10.1 $\mu\text{g/L}$), which have considerably lower optima for TP, undergo a concomitant decrease in relative abundance. The diatom assemblages between 1917 and 1980 tend to indicate a period of elevated trophic status in Williams Lake. The remaining diatom taxa show very little change in percent abundance throughout the sediment core (Fig. 2).

DISCUSSION

Limitations of the Diatom-Inference Model

In our publication (Appendix A) we state that a priori tests exist to assess if our model can be applied to a fossil diatom assemblage. All these tests tended to indicate that the Williams Lake sediment diatom assemblages were not unusual in any way and that the model may be

applied. Furthermore, the TP inferences were based on an extremely high proportion of the total sediment diatom assemblages (between 93 and 98% of the total diatom count; see Appendix B). All these considerations would tend to indicate that our model could be applied with some confidence to the Williams Lake core samples. However, because the present day TP concentration of Williams Lake (35 - 85 $\mu\text{g/L}$) is outside the range of our calibration lake set, we believe the TP inferences are unreliable.

Our TP inference model is based on the distributions of diatom taxa in 37 British Columbia lakes that range from 5 to 28 $\mu\text{g/L}$ TP. For this reason, we state in our publication (Appendix A) that the model will give poor inferences in lakes outside of this range because it cannot infer values greater than 28 $\mu\text{g/L}$, and that values should be interpreted with caution in such cases. In order to extend our model to a more useful range of lake trophic status that includes the eutrophic (and hence problematic) end of the spectrum, we must include diatom data from the surficial sediments a number of lakes with TP concentrations above 28 $\mu\text{g/L}$. At present, our calibration lake set includes only 2 lakes with TP greater than 28 $\mu\text{g/L}$ and they are so far from the mean TP of the lake set that they form outliers and presently cannot be included in the WA model. We will continue to refine our model and recommend that 10 to 12 lakes with present day TP between 28 and 85 $\mu\text{g/L}$ be added to our data set, so that it can reliably assess past lake trophic status in excess of the present 28 $\mu\text{g/L}$ limit. By adding these lakes we will be able to more accurately calculate weighted averages of the individual diatom taxa, and so will be able to provide more accurate TP estimates from eutrophic lakes.

Recent Total Phosphorus Trends

Despite some changes in the taxonomic composition of the diatom communities, inferences based on our B.C. diatom calibration set and weighted averaging regression and calibration model indicate that Williams Lake has experienced a rather consistent TP trajectory throughout the past 150 years (Table 1, Fig.1). The values of the TP inferences are likely underestimates of the true TP concentrations (for reasons outlined above), and we therefore conclude that Williams Lake was eutrophic before settlement by Europeans. The weighted averaging inference model indicates that TP concentrations increased slightly in the lake between 1956 and 1969. The timing of this shift correlates with, and may be related to, the rebuilding and paving of the Cariboo Highway along the north shore of Williams Lake in 1955 (Zirnhelt 1984). TP inferences indicate higher post-1956 TP concentrations compared to the period before 1956.

Taxonomic shifts in the diatom assemblages suggest that limnological conditions in Williams Lake began to change as early as 1917 (Fig. 2). At this time Cyclotella michiganiana and Fragilaria brevistriata decline in relative abundance. These two taxa are generally considered to be indicative of less productive water bodies, and have relatively low optima for TP in our data set. Their replacement by three planktonic species (Stephanodiscus hantzschii, Fragilaria crotonensis and Asterionella formosa) with higher TP optima, indicates that Williams Lake may have become more eutrophic between 1917 and 1980. All three taxa are generally considered to be good indicators of eutrophication, and commonly increase in percent abundance following human activity. Stephanodiscus hantzschii and Asterionella formosa, in particular, have been associated with increases in nutrient loading that result from agricultural practices.

Our inference model, however, does not display a marked increase in TP concentration during this period. The year 1917 saw the building of the PGE railway along the south shore of Williams Lake and led to increases in cattle ranching (Zirnhelt 1984). We believe these events had an affect on the diatom communities of Williams Lake, but our inference models are as yet too coarse to quantitatively document these changes.

Prior to 1850, the diatom community of Williams lake was predominantly composed of planktonic taxa that are typical of fairly productive waters (Fig. 2). We believe that the diatom assemblage in Williams Lake prior to 1850 is indicative of eutrophic conditions. It is possible that further eutrophication of the lake was only coupled with small changes in the diatom flora, because there are very few planktonic diatom taxa with higher TP optima that could replace the "background", or natural diatom assemblage. For this reason, we believe the diatom community may have been rather insensitive to further eutrophication in Williams Lake. It is probable that further lake eutrophication led to dramatic shifts in the types of algae growing in Williams Lake, rather than just causing shifts in the relative abundance of diatom species. Changes in other algal groups (e.g. blue-green (Cyanophyta) and green (Chlorophyta) algae) may provide a better signal of eutrophication events in this lake because they tend to be more abundant and dominant than diatoms in highly eutrophic lake environments. These algal groups leave biogeochemical fossils in the form of pigments and pigment derivatives that can be measured in lake sediments. Dr. Peter Leavitt (University of Alberta) is presently analyzing the remains of the algal pigments in the Williams Lake sediment core in order to test this hypothesis and to provide an independent indicator of lake eutrophication. We believe this research will provide some useful insights into the recent history of Williams Lake. Preliminary, visual (qualitative) scans of wet lake sediment

using a light microscope indicate an abrupt increase and abundance in blue-green algal remains above the 4 cm level in the sediment core, and suggest recent (post-1980) eutrophication. Blue-green algal remains are digested in the process of making diatom slides, and so we require the analyses by Dr. Leavitt to confirm this observation.

LITERATURE CITED

- Binford, M.W. 1990. Calculation and uncertainty analysis of ^{210}Pb dates for PIRLA project lake sediment cores. *Journal of Paleolimnology* 3: 253-267.
- Hall, R.I. and Smol, J.P. 1992. A weighted-averaging regression and calibration model for inferring total phosphorus concentration from diatoms in British Columbia (Canada) lakes. *Freshwater Biology*, in press.
- Smol, J.P. 1983. Paleophycology of a high arctic lake near Cape Herschel, Ellesmere Island. *Canadian Journal of Botany* 61: 2195-2204.
- Zirnhelt, N. 1984. Preliminary paleolimnology of Williams Lake, B.C. Undergraduate Research Report, Simon Fraser University.

Figure 1. Diatom inferred total phosphorus stratigraphy of a sediment core spanning the past 150 years of Williams Lake, British Columbia.

Williams Lake, British Columbia

Diatom Inferred Total Phosphorus

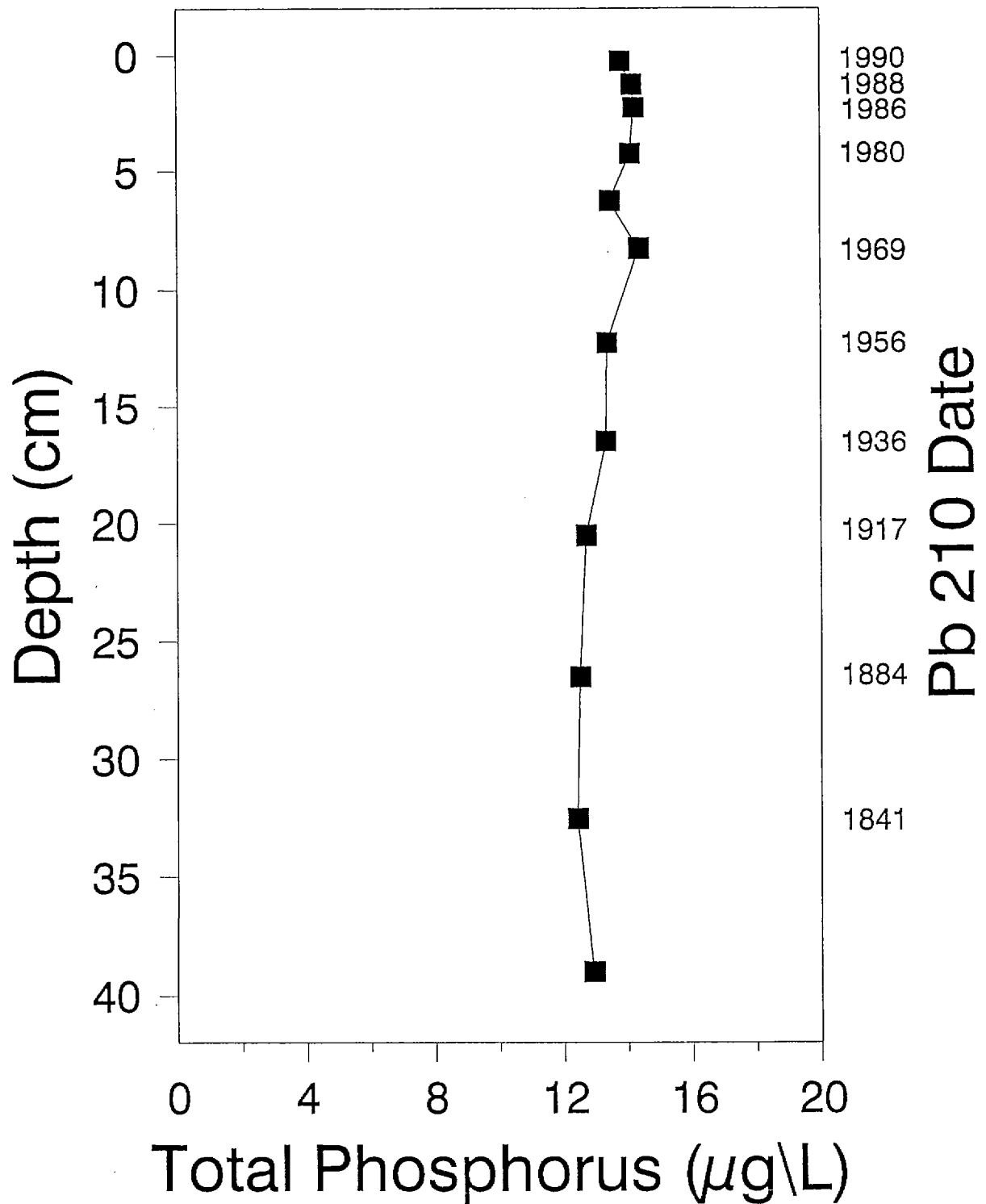


Figure 2. A stratigraphy of the most common diatoms in a surface sediment core from Williams Lake, British Columbia.

- Ennis G.L., Northcote T.G. & Stockner J.G. (1983) Recent trophic changes in Kootenay Lake, British Columbia, as recorded by fossil diatoms. Canadian Journal of Botany, **61**, 1983-1992.
- Farley A.L. (1979) Atlas of British Columbia. The University of British Columbia Press, Vancouver, Canada. pp.136.
- Fritz S.C. (1990) Twentieth-century salinity and water-level fluctuations in Devils Lake, North Dakota: Test of a diatom-based transfer function. Limnology and Oceanography, **35**, 1771-1781.
- Gauch H.G. Jr. (1982) Noise reduction by eigenvector ordinations. Ecology, **63**, 1643-1649.
- Gause G.F. (1930) Studies on the ecology of the Orthoptera. Ecology, **11**, 307-325.
- Gensemer R.W. (1991) The effects of pH and aluminum on the growth of the acidophilic diatom Asterionella ralfsii var. americana. Limnology and Oceanography, **36**, 123-131.
- Germain H. (1981) Flore des diatomées: Diatomophycées. Société Nouvelle des Éditions Boubée, Paris. pp. 444.
- Håkansson H. (1986) A taxonomic reappraisal of some Stephanodiscus species (Bacillariophyta). British Phycological Journal, **21**, 25-37.
- Håkansson H. & Stoermer E.F. (1984a) An investigation of the morphology of Stephanodiscus alpinus Hust. Bacillaria, **7**, 159-172.
- Håkansson H. & Stoermer E.F. (1984b) Observations on the type material of Stephanodiscus hantzschii Grunow in Cleve & Grunow. Nova Hedwigia, **39**, 477-495.
- Hill M.O. (1979) DECORANA. -A FORTRAN program for detrended correspondence analysis and reciprocal averaging. Ecological Systematics. Cornell University, Ithaca, N.Y.
- Hustedt F. (1930) Bacillariophyta (Diatomeae). Die Süßwasserflora Mitteleuropas, **10**, 1-466.
- Hustedt F. (1937-1949) Systematische und ökologische Untersuchung über die Diatomeen Flora von Java, Bali, und Sumatra. Archiv für Hydrobiologie Supplementband, **15**, 131-177, 187-295, 393-506, 790-836, **16**, 1-155, 274-394.
- Kilham S.S. (1975) Kinetics of silicon-limited growth in the freshwater diatom Asterionella formosa. Journal of Phycology, **11**, 396-399.
- Kingston J.C., Birks H.J.B., Uutala A.J., Cumming B.F. & Smol J.P. (1992) Assessing trends in fishery resources and lake water aluminum from paleolimnological analyses of siliceous algae. Canadian Journal of Fisheries and Aquatic Science, **49**, 127-138.
- Koppen J.D. (1975) A morphological and taxonomic consideration of Tabellaria (Bacillariophyceae) from northcentral United States. Journal of Phycology, **11**, 236-244.
- Krammer K. & Lange-Bertalot H. (1986) Süßwasserflora von Mitteleuropa. Bacillariophyceae 1. Teil: Naviculaceae. Gustav Fischer Verlag, Stuttgart. pp.

References

- Agbeti M. (1992) The relationship between diatom assemblages and trophic variables: a comparison of old and new approaches. Canadian Journal of Fisheries and Aquatic Sciences (in press).
- Agbeti M. & Dickman M. (1989) Use of lake fossil diatom assemblages to determine historical changes in trophic status. Canadian Journal of Fisheries and Aquatic Sciences, **46**, 1013-1021.
- Beil C.E., Taylor R.L. & Guppy G.A. (1976) The biogeoclimatic zones of British Columbia. Davidsonia, **7**, 45-55.
- Birks H.J.B., Line J.M., Juggins S., Stevenson A.C. & ter Braak C.J.F. (1990) Diatoms and pH reconstructions. Philosophical Transactions of the Royal Society of London, series B, **327**, 263-278.
- Birks H.J.B., Juggins S. & Line J.M. (1990) Lake surface-water chemistry reconstructions from palaeoecological data. In Mason B.J. (Ed.) The Surface Waters Acidification Programme, Cambridge University Press, Cambridge. pp. 301-311.
- Bothwell M.L. (1985) Phosphorus limitation of lotic periphyton growth rates: An intersite comparison using continuous-flow troughs (Thompson River system, British Columbia). Limnology and Oceanography, **30**, 527-542.
- Bothwell, M.L. (1992) Nutrient and grazer control of algal biomass in the Thompson River, British Columbia: a case history of water quality management. Proceedings of the Rawson Conference on Aquatic Ecosystems in Semi-Arid Regions, Saskatoon, Canada, August 1990 (in press).
- Boucherle M.M, Smol J.P., Oliver T.C., Brown S.R. & McNeely R. (1986) Limnological consequences of the decline in hemlock 4800 years ago in three Southern Ontario lakes. Hydrobiologia, **143**, 217-225.
- Brown S.R. (1969) Paleolimnological evidence from fossil pigments. Mitteilungen Internationale Vereinigung für Limnologie, **17**, 95-103.
- Brugam R.B. (1978) Human disturbance and historical development of Linsley Pond. Ecology, **59**, 19-36.
- Camburn K.E., Kingston J.C. & Charles D.F. (Eds.) (1986) PIRLA Diatom Iconograph. PIRLA unpublished report series #3, Dept. of Biology, Indiana University, Bloomington, U.S.A.
- Christie C.E. (1988) Surficial diatom assemblages as indicators of lake trophic status. M.Sc. thesis, Queen's University, Kingston, Ontario, Canada. pp.124.
- Cumming B.F., Smol J.P. & Birks H.J.B. (1992) Scaled chrysophytes (Chrysophyceae and Synurophyceae) from Adirondack drainage lakes and their relationship to environmental variables. Journal of Phycology (in press).
- Dixit S., Dixit A.S., & Smol J.P. (1991) Multivariable environmental inferences based on diatom assemblages from Sudbury (Canada) lakes. Freshwater Biology, **26**, 251-266.

Our model has been applied to diatom assemblages preserved in a sediment core from Kamloops Lake, British Columbia (a lake not included in the calibration set). The diatom inferred TP ($8.7\mu\text{g l}^{-1}$) from the surficial sediment sample closely agrees with the average TP ($8.8\mu\text{g l}^{-1}$) measured by the British Columbia MOE (Colin McKean, pers. comm.). Although this preliminary result indicates the model may perform well, we realize that this is a single sample with no statistical degrees of freedom and must be interpreted with caution. The predictive ability of our model will be tested further by applying it to lakes that have long term water chemistry data, and comparing TP inferences to directly measured values.

Acknowledgements

We thank Colin McKean at the Ministry of the Environment, Victoria, British Columbia for contributing the surface sediment samples and environmental data. Without his help, this project would never have been possible. Hedy Kling helped with the taxonomy of small centric diatoms. We thank the scientists in our lab for many helpful comments and discussions. This project was funded by the Natural Sciences and Engineering Research Council of Canada via an operating grant to John P. Smol, and a postgraduate scholarship to Roland Hall.

Weighted-Averaging Models for Total Phosphorus Reconstruction

The ratio of λ_1/λ_2 tends to be higher in studies of lake acidification than the ratio presented here. For example, a diatom calibration set from the Sudbury region of Canada has a λ_1/λ_2 ratio for pH of 0.84 (Dixit, Dixit, and Smol, 1991). The relationship between diatoms and pH is strong because pH exerts a direct physiological stress on diatoms (Gensemer, 1991), and also strongly influences other lakewater chemistry variables (Stumm and Morgan, 1981). The relationship between diatoms and TP is not quite so direct because other factors (such as intracellular phosphorus storage, physical mixing, and silica availability) are also involved (Reynolds, 1980). The diatom distributions in our data set do not appear to be strongly influenced by pH, largely because the lakes include a relatively narrow pH range (7.5 - 8.5).

Simple WA models provide a lower apparent root mean squared error of prediction and a higher correlation between observed and inferred TP than WA_(tol) models (Fig.3, Table 6). WA calibration with inverse deshrinking has a significant trend in the residuals, and tends to underestimate TP at the high end of the TP gradient and over-estimate values at the low end (Fig. 3). We therefore conclude that simple WA calibration with classical deshrinking provides the most reliable inferences of TP, and is the best model for inferring TP in our data-set.

Our WA model can be used to infer past lakewater TP from diatom assemblages preserved in the sediments of British Columbia lakes within the TP range of 5 to 28 $\mu\text{g l}^{-1}$. The TP inferences are based on the weighted averages (or, optima) of 131 diatom taxa that are common in the British Columbia lakes. TP hindcasts are obtained by using the following classical regression equation; $y = 1.98 + 0.23x$, where y and x are, respectively, the corrected and uncorrected TP inferences (Birks *et al.*, 1990).

Quantitative estimates of past TP generated by the model can provide useful information concerning the onset, rate and magnitude of changes in lake trophic status that result from natural and anthropogenic causes. Our model can also provide a useful approach for testing hypotheses and validating models concerning land use and lake-watershed interactions that require long-term water chemistry data. Because our inference model is based on WA regression and calibration, we believe it possesses at least two advantages over existing trophic status models that are based on multiple linear regression (Agbeti and Dickman, 1989; Whitmore, 1989). First, reconstructions are based on the responses of individual taxa to TP, and do not involve grouping the taxa into a small number of ecological categories (ter Braak and van Dam, 1989). Second, the model assumes that diatoms respond to TP in a unimodal, rather than a linear, fashion. Furthermore, *a priori* tests are available to assess if the model can be applied to fossil diatom assemblages. Sediment diatom assemblages with: (i) a sample score outside a 95% confidence limit about the mean in a DCA of the diatom data in the 37 lake calibration set; or (ii) a residual fit outside the 95% confidence limit when run as a passive sample in a CCA of the 37 lake-set constrained to TP, would have a poor fit to the model. In either of these cases, the inferences are likely to be unreliable.

confident that our calibration lake-set and WA inference model are developed from a broad range of lake conditions, and that they include a useful range of lakewater TP.

We removed a further four samples (Echo L. (#12), Champion#2 L. (#20), Tahla L. (#32), and Comfort L. (#46); Table 1) from our WA regression and calibration analyses because diatom assemblages found in these lakes have a poor fit to the TP axis in a constrained CCA. The WA models therefore include a TP range of 5-28 $\mu\text{g l}^{-1}$. It is difficult to ascertain the exact reason for the poor fit to the TP axis in these samples. Possibly, the single "snapshot" TP measurements were unusually high or low only during the sampling period and not during the majority of the diatom growing season, and therefore the diatom assemblages show poor agreement to the measured TP values in these lakes. Also, diatoms in these outlier lakes could be responding to some other major perturbation(s) or limnological variable(s) that could mask or alter the TP signal. Three of the four lakes (Echo, Tahla, and Comfort L.) have very reduced species diversity in the surficial sediment samples, and are dominated by between one and four diatom species. Data that could help us determine why these lakes are outliers are unavailable. Fortunately, *a priori* tests are available to check that the WA model for TP inferences can be applied to fossil diatom assemblages from unknown samples (see below).

Ordination

The discrepancy between eigenvalues of the first 2 axes in a CCA and a DCA of the 41 lake-set indicates that a large proportion of the variance in the diatom data is accounted for by the environmental variables included in the data, but other unmeasured variables also influence the distribution of diatoms. For example, it is well known that physical mixing (or turbulence) plays an important role in maintaining planktonic diatoms within the photic zone of lakes (Reynolds, 1980), and that silica availability controls diatom assemblages (Kilham, 1975). Turbulence and silica availability could play a role in the seasonality of diatoms, and influence the distributions of certain diatom taxa. We believe information concerning these and other variables would likely improve the ordination, but are not available for this lake-set.

The CCA of the 41 lake data-set demonstrates a strong relationship between diatom distributions and the 8 environmental variables (Fig. 2). Maximum depth, conductivity, TP, and [CA] each exert significant and independent influences on the diatom distributions (Fig. 2). Furthermore, TP accounts for a statistically significant direction of variance in the diatom data, as indicated by a Monte Carlo permutation test of the eigenvalue of the first axis in a CCA constrained to TP. Based on this, and a $\lambda_1/\lambda_2 = 0.40$ in a CCA constrained to TP, we conclude that the relationship between diatoms and TP is sufficiently robust for reconstructions based on WA models, but they should also be interpreted with the consideration that other variables may strongly influence the distributions of diatom taxa.

as the former produces TP reconstructions with a lower apparent root mean squared error of prediction and has a stronger correlation between observed and inferred TP (Table 6, Fig. 3). Deshrinking by inverse regression produces a lower root mean squared error of prediction than classical deshrinking (Table 6). However, inverse deshrinking displays a significant trend in the residuals ($r^2 = 0.38$, $p < 0.001$; d.f. = 35, $n = 37$) (Fig. 3). Inverse deshrinking tends to over-estimate values at the low end of the TP range and under-estimate the high values. Classical deshrinking does not produce a significant trend in the residuals ($r^2 = 0.001$, $p > 0.05$; d.f. = 35, $n = 37$) (Fig. 3), and therefore, provides the most reliable inferences of TP.

Discussion

Data Screening

The full 46 lake calibration set includes a TP range of 5-85 $\mu\text{g l}^{-1}$. The final CCA ordination was performed on 41 of these lakes and includes a TP range of 5-28 $\mu\text{g l}^{-1}$ because lakes with the highest TP were outliers. Five lakes were identified as outliers (based on the criteria discussed earlier) and were removed from the ordination analyses.

Gardom L. (#6) is unusual because the surface sediment diatom assemblage is dominated by a single taxon; *Cyclotella michiganiana* comprises 89% of the diatom count. Furthermore, Gardom L. has the highest specific conductivity (455 $\mu\text{S cm}^{-1}$; calibration lake-set mean = 203 $\mu\text{S cm}^{-1}$, standard deviation = 100 $\mu\text{S cm}^{-1}$) of the full 46 calibration lake-set and is 2.5 times greater than the mean conductivity.

The other four lakes were removed from the final CCA analysis because they possess extreme value(s) for the environmental variables (i.e. they had a value greater than three standard deviations from the mean; ter Braak 1990b). Kootenay L. (#5) has extreme lake depth (maximum depth = 154 m; calibration lake-set mean = 23.6 m, standard deviation = 24.9 m) and extreme lake surface area (38,900 ha; calibration lake-set mean = 1,100 ha, standard deviation = 5,655 ha). Boss L. (#11) has extreme lakewater Fe concentration (8760 $\mu\text{g l}^{-1}$; calibration lake-set mean = 265 $\mu\text{g l}^{-1}$, standard deviation = 1268 $\mu\text{g l}^{-1}$). Holden L. (#26) and Little Bobtail L. (#38) have extreme TP concentrations (45 and 85 $\mu\text{g l}^{-1}$ respectively; calibration lake-set mean = 15 $\mu\text{g l}^{-1}$, standard deviation = 13 $\mu\text{g l}^{-1}$).

Limnologically, there is perhaps nothing particularly unusual about any of these four lakes; they simply possess environmental variables greater than three standard deviation units from the sample mean, and therefore exert extreme leverage in CCA (ter Braak 1990b). The lakes form outlier samples because the calibration lake-set does not include a sufficient distribution of lakes along the full ranges of all the important environmental gradients. For example, if our calibration lake-set included several more lakes with TP between 28 and 85 $\mu\text{g l}^{-1}$, then Holden and Little Bobtail L. would no longer form outlier samples. However, British Columbia lakes with lakewater TP greater than 30 $\mu\text{g l}^{-1}$ are much more rare than lakes with TP less than 30 $\mu\text{g l}^{-1}$, and they are usually perturbed systems. We are

contribute significantly to axis 1; and Secchi and [Al] contribute significantly to axis 2.

CCA axes 1 and 2 (Fig. 2) roughly separate deep, low conductivity lakes with well developed planktonic diatom floras (e.g. Echo L. (#12), Lac de Roche (#21), Keno L. (#39)) from shallow, high conductivity lakes dominated by benthic diatoms (e.g. Champion#2 L. (#20), Windermere L. (#28), Comfort L. (#46)). Benthic taxa characteristic of shallow, unproductive, dilute lakes, including *Fragilaria construens* var. *venter* (#82), *F. lapponica* (#78), *F. brevistriata* (#36), *Achnanthes biasolettiana* (#47), and *Cymbella diluviana* (#24, not shown in Fig. 2) occur in the lower right hand quadrant of the CCA biplot (Fig. 2, Table 4). Planktonic taxa characteristic of deeper and more productive lakes, including *Stephanodiscus parvus* (#66), *S. minutulus* (#90), *Tabellaria flocculosa* str. IIIp (#114, not shown in Fig. 2), and *Aulacoseira granulata* var. *angustissima* (#131, not shown in Fig. 2), occur in the upper left hand quadrant (Fig. 2, Table 4).

TP, maximum depth and surface area also make significant contributions to axis 3, based on canonical coefficients and an approximate t-test ($p \leq 0.05$; data not shown). Intra-set correlations indicate that the influences of maximum lake depth (0.38) and TP (-0.85) on the diatoms are split along axis 3. CCA is therefore able to separate diatom assemblages that occur in deeper, more unproductive lakes from those occurring in shallower, more nutrient rich lakes.

Weighted-Averaging Models for Total Phosphorus Reconstruction

Reliable reconstruction of lakewater chemistry from the relative abundance of diatom taxa in a sample requires a good statistical relationship between the diatoms and the variable to be reconstructed (ter Braak, 1987a). The eigenvalue of the first axis ($\lambda_1 = 0.23$) in a constrained CCA with TP as the only environmental variable is significant ($p \leq 0.05$), based on a Monte Carlo permutation test (99 unrestricted permutations; ter Braak, 1988b, 1990b). This indicates that TP accounts for a statistically significant direction of variance in the diatom data. The ratio of λ_1/λ_2 in the constrained CCA is 0.40, indicating that TP is an important variable controlling diatom distributions, but other factors strongly influence the composition of the diatom assemblages.

The WA regression and calibration equations for inferring TP are derived from 37 lakes in British Columbia that include a wide variety of limnological conditions. The range of important environmental variables included in the calibration lake-set are presented in Table 5. The model should be applied only to British Columbia lakes that fall within these ranges. Our 37 lake-set includes a TP range of 5-28 $\mu\text{g l}^{-1}$. TP inferences are based on the optima of 131 diatom taxa common in the British Columbia lakes, and the following classical deshrinking regression equation; $y = 1.98 + 0.23x$, where y and x are, respectively, the corrected and uncorrected TP inferences (Birks *et al.*, 1990). We provide the estimated optima (or weighted averages) of all the diatom taxa, and the number of occurrences upon which these estimates are based, in Table 4.

Simple WA calibration of the 37 lake data-set performed better than $WA_{(tol)}$,

variables, including Secchi. TDS was omitted from the final CCA ordination because it had a variance inflation factor >20 ; this indicates that its canonical coefficient is unstable and does not merit further interpretation (ter Braak, 1986).

Unusual samples

Five outlier samples were omitted from the final CCA ordination because they fulfilled one of the criteria of outliers (as outlined in the methods). The outliers comprise: Kootenay L. (#5), Gardom L. (#6), Boss L. (#11), Holden L. (#26) and Little Bobtail L. (#38) (Table 1). Gardom L. fulfilled criterion #1, whereas Kootenay, Boss, Holden, and Little Bobtail lakes fulfilled criterion #2. The final CCA, therefore, includes 41 lakes (Fig. 2) that span a TP range of 5-28 $\mu\text{g l}^{-1}$.

A further four outlier samples were omitted from WA regression and calibration analysis because they fulfilled criterion #3 for outlier samples. They comprise: Echo L. (#12), Champion#2 L. (#20), Tahla L. (#32), and Comfort L. (#46) (Table 1). WA regression and calibration of TP are therefore based on 37 of our original 46 calibration lakes, and include a TP range of 5-28 $\mu\text{g l}^{-1}$.

Ordination

The eigenvalues ($\lambda_1=0.58$, $\lambda_2=0.35$) for the first two DCA axes (run with detrending by segments and nonlinear rescaling of axes) account for 17.5% of the cumulative variance in the weighted averages of diatom taxa in our 41 lake data-set. The low percentage of explained variance is typical of noisy data-sets containing a large number of taxa and many zero values (Stevenson *et al.*, 1991).

A CCA of the same 41 lake data-set constrained to the 8 environmental variables explains a smaller proportion of the taxon variance; the eigenvalues for CCA axis 1 (0.38) and axis 2 (0.25) explain 11.8% of the cumulative variance in the weighted averages of the diatom taxa (Fig. 2). However, the CCA includes a large proportion of the variance in the diatom-environmental relationship; the first two axes account for 44% of the variance in the diatom-environment relationship. The diatom-environment correlations for CCA axis 1 (0.90) and axis 2 (0.87) are high, and indicate a strong relationship between diatom taxa and the 8 environmental variables. Monte Carlo unrestricted permutation tests of the axes indicate that axis 1 (99 permutations) and axis 2 (99 permutations of axis 2 with axis 1 as a covariable) are statistically significant ($p \leq 0.05$).

The forward selection option of CCA with Monte Carlo permutation tests (99 permutations, $p \leq 0.05$) indicates that maximum depth, conductivity, TP, and [Ca] make significant and independent contributions to explaining the underlying variation in the diatom assemblages (Fig. 2). The remaining variables do not individually explain any significant additional proportion of the variance of the taxon weighted averages. Canonical coefficients of the environmental variables are used to construct the CCA axes, and the significance of their contribution to each axis is best judged by an approximate t-test (ter Braak, 1988a). We present the canonical coefficients and t-values of the 8 environmental variables in Table 3, and the approximate t-test indicates that maximum depth and TP contribute significantly ($p \leq 0.05$) to axis 1 and 2; conductivity, [Ca] and surface area

of a diatom taxon (i.e. the average of all TP values of lakes in which the taxon occurs, weighted by the taxon's relative abundance in each lake) provides a computationally simple and reliable estimate of its TP optimum (Birks *et al.*, 1990). This computational process is referred to as WA regression. WA calibration performs the reverse function; it uses the optima and relative abundances of diatoms in a sediment sample to estimate TP. A more complete discussion of WA theory (ter Braak and Barendregt, 1986; ter Braak and Looman, 1986; ter Braak and Prentice, 1988) and methodology (ter Braak and van Dam, 1989; Birks *et al.*, 1990) are available elsewhere.

Variations of the simple WA regression and calibration model can be derived. Species tolerances can be incorporated into WA models (WA_{tol}) by downweighting each species by its variance for TP (ter Braak and van Dam, 1989; Line and Birks, 1990). Also, shrinkage in the range of inferred values occurs because averages are taken twice; this can be corrected by either classical or inverse deshrinking (or regression) methods (ter Braak and van Dam 1989). The predictive abilities of the various models were assessed in terms of the correlation between observed and inferred TP (Birks *et al.*, 1990), the apparent root mean squared error (RMSE) of prediction (Wallach and Goffinet, 1989), and the distribution of residual values (observed-inferred TP) (Birks *et al.*, 1990).

Results

Data Screening

Redundant environmental variables

Eleven of the original 19 environmental variables were eliminated from the analysis prior to forward selection. Three variables (NH_4^+ , NO_3^- , organic N) were eliminated because values were below the limit of detection or were not available for many of the lakes. Eight other variables (watershed area, mean depth, pH, TDS, hardness, Fe, Mg, Mn) were excluded on the basis of our data screening criteria (i.e. the excluded variables did not independently influence the distributions of diatom taxa, and were significantly correlated with other environmental variables already selected for inclusion in the analysis).

Prior to forward selection, the 8 environmental variables remaining in the analysis were: specific conductivity, maximum depth, Secchi, altitude, surface area, and concentrations of TP, Ca, and Al. Altitude and [Al] were not significantly correlated with any other environmental variables; the remaining variables were significantly correlated with one or more environmental variables (Table 2). Each of the latter variables, therefore, represents a group of correlated variables. For example, conductivity was included in the analysis, but was significantly correlated with pH, hardness, and [Mg]. Likewise, Secchi represents a group of correlated variables, including TP, [Fe], and [Mn]; [Ca] represents the group of variables including pH and TDS; maximum lake depth represents a pair of correlated variables (maximum depth and mean depth); and surface area represents a pair of correlated variables (surface area and watershed area). The influence of TP on diatom distributions was found to be independent of all other environmental

diatom distributions. In each case, the environmental variable identified in step (ii) would be used as the sole variable. Each correlated environmental variable was tested for independent influence on the diatom distributions by entering it as the sole covariable in a partial CCA and testing the significance of the first canonical axis (Monte Carlo test with 99 unrestricted permutations, $p \leq 0.05$) (ter Braak, 1988b). If the covariable did not exert an independent influence, it would be excluded in subsequent analyses.

This routine reduced the amount of redundancy in the environmental data by selecting environmental variables which were either: (i) not significantly correlated to other environmental variables, or (ii) were correlated, but exerted an independent influence on diatom distributions.

Unusual samples

Ordination was also used to screen the data to detect any unusual or "outlier" samples. Outlier samples (Birks *et al.*, 1990) may have an unusual diatom assemblage, an unusual combination of environmental variables, or a diatom assemblage with a poor relationship to the environmental variable of interest (i.e. TP). An outlier was detected and deleted from the final CCA if: 1) the sample score fell outside the 95% confidence limits about the sample score means (only on axes reflecting data structure; Gauch, 1982) in both a DCA of the species data and a principal components analysis (PCA) of the environmental data in the full calibration set; or 2) the sample had an environmental variable with extreme ($> 5x$) influence, or "leverage", as detected by CCA (ter Braak, 1990a; Birks, Juggins and Line, 1990). Samples were also omitted from WA calibration and regression analysis of TP if: 3) their residual length to the TP axis fell outside a 95% confidence limit in a constrained CCA with TP as the only environmental variable (Birks *et al.*, 1990).

Environmental variables unrelated to diatom distributions

Following the data screening process described above, a final CCA was performed with forward selection. Forward selection and Monte Carlo permutation tests of the statistical significance of each variable identified a subset of environmental variables such that each environmental variable explained a significant or near significant amount of variance in the species data (Birks, Juggins and Line, 1990). In our analysis, several environmental variables were identified as explaining a significant proportion of the variance, but all environmental variables (significant and not significant) were included in the final CCA ordination.

(c) Regression and calibration

WA regression and calibration were performed using the computer program WACALIB 2.1 (Line and Birks, 1990) to infer TP based on diatom assemblages in the calibration lake-set. WA techniques assume that diatom taxa are most abundant in lakes that have a TP at, or near, their optima. The weighted average

(a) Ordination

Ordinations were performed using the computer program CANOCO (ter Braak, 1988a) version 3.10 (ter Braak, 1990a,b), and rare species were downweighted in all cases. Principal patterns in the distribution of diatom taxa within the calibration lakes were explored using detrended correspondence analysis (DCA) and CCA. The maximum amount of variation in the species data was determined using DCA (Hill, 1979; ter Braak, 1985). CCA was used to explore the relationships between diatom distributions and measured environmental variables.

DCA is an indirect-gradient analytical technique that effectively ordines samples and diatom taxa in low dimensional space for data that have a large number of taxa and many zero values (i.e. for samples in which taxa are missing; ter Braak, 1987b). DCA assumes a unimodal distribution of taxa along an environmental gradient.

CCA is a direct-gradient technique that simultaneously represents diatom taxa, samples and environmental variables in low dimensional space (ter Braak, 1987b). Most of the variation is usually accounted for by the first 2 axes. The axes are constrained to be linear combinations of the environmental variables. CCA assumes that species respond to environmental variables in a unimodal manner. Forward selection (ter Braak, 1990b) was used to identify a subset of environmental variables that explain statistically significant and independent directions of variation in the diatom distributions.

(b) Data Screening

Large ecological data-sets contain: (i) redundancies in environmental information, (ii) unusual samples, and (iii) environmental variables which do not determine diatom distributions (Birks *et al.*, 1990). Prior to analysis, the species and environmental data files were screened to identify and eliminate these problems.

Redundancies in environmental information

The full British Columbia data-set contained superfluous environmental variables that were significantly correlated among lakes and did not influence diatom distributions in an independent fashion. To decide on the environmental variables to be used in CCA, we performed the following routine:

(i) groups of significantly ($p \leq 0.05$) correlated environmental variables were identified from a Pearson correlation matrix with Bonferroni adjusted probabilities (Wilkinson, 1988).

(ii) we ran a preliminary CCA with forward selection (outlier samples omitted; see below) to determine which environmental variables in each group explain the greatest amount of variance in the diatom data. This environmental variable would be used in subsequent analyses as a representative of that particular group of correlated variables.

(iii) we ran a series of partial constrained CCA's to determine if other members of a particular variable group exerted an independent influence on the

Håkansson (1984), Camburn, Kingston and Charles (1986), and Krammer and Lange-Bertalot (1986).

Sediment samples are stored at the Department of Biology, Queen's University, Kingston, Ontario, K7L 3N6, Canada. A set of microscope slides and cleaned diatom slurries are archived in the Algal Collection, Canadian Museum of Nature, Ottawa, Canada (Accession numbers 35721 - 35767).

Data analysis

Quantitative predictions of lakewater chemistry typically depend on establishing relationships between present-day diatom species and environmental variables using a suite of study lakes (a "training set" or "calibration set"). Species data usually consist of relative abundances of diatom taxa in surficial sediment samples. These samples provide an integrated sample (in space and time) of the diatom taxa that have accumulated over the previous few years. The environmental data consist of present-day physical, chemical and biological variables that have been measured, and are thought to be ecologically important. If the distributions of diatom taxa are strongly related to a particular environmental variable, realistic reconstructions of that variable should be possible from the diatom species composition of sediment samples (ter Braak, 1987a).

Our calibration set consists of diatom counts (expressed as relative abundance) from the surficial sediments of 46 British Columbia lakes with accompanying water chemistry and limnological information. The diatom data include all taxa that were present in a minimum of 2 samples and had a relative abundance of $\geq 1\%$ of the diatom count in at least 1 sample. A total of 134 diatom taxa met these criteria.

Chemistry and limnological data were available for the following variables: altitude, watershed area, surface area, maximum depth, mean depth, Secchi, pH, specific conductivity, total dissolved solids (TDS), hardness, and concentrations of NH_4^+ , NO_3^- , organic nitrogen, TP, Ca, Fe, Mg, Mn, and Al. Median values were substituted in the few cases where data for some variables were missing for individual lakes (Table 1). The distributions of 9 environmental variables (watershed area, surface area, mean depth, maximum depth, and concentrations of TP, Fe, Mg, Mn, and Al) were found to be positively skewed and were therefore $\text{Ln}(x+1)$ transformed (Zar, 1984). Nitrogen data (NH_4^+ , NO_3^- , and organic nitrogen) were not included in the numerical analyses, because NH_4^+ and NO_3^- concentrations were below the detection limit in nearly all of the calibration lakes, and organic nitrogen data were lacking for 11 of our lakes. The calibration data set, therefore, consists of 46 samples x 134 diatom taxa x 16 environmental variables. The data are stored on computer spreadsheets and maintained by R. Hall.

Numerical procedures involved ordination, data screening, and WA regression and calibration, as discussed below.

Study Area

The topography, geology, climate, vegetation and soils of British Columbia are extremely heterogeneous (Farley, 1979), and the lakes display a similarly broad range of physical, chemical, and biological characteristics. This study includes 46 freshwater lakes (with well preserved surficial sediment diatom assemblages) distributed predominantly in the southern half of the province (Fig. 1, Table 1). The lakes were specifically chosen to reflect a broad trophic gradient, and they span a wide range of TP ($5\text{--}85\mu\text{g l}^{-1}$). In order to minimize the influence of pH and pH-related variables on the distribution of diatom taxa, lakes were selected to include a relatively narrow pH range (7.5–8.5). The study area includes 4 biogeoclimatic zones (Interior Douglas Fir, Coastal Douglas Fir, Interior Western Hemlock, and Cariboo Aspen - Lodgepole Pine). These zones are defined primarily in terms of temperature and precipitation (Beil, Taylor and Guppy, 1976), but also differ in vegetation and soil types.

Methods

Sampling methods

All surface sediment and water chemistry samples were collected by the British Columbia Ministry of the Environment (MOE) from the deepest portion of each lake. The lakes were sampled once during June to September, 1983 ($n=17$); May to October, 1984 ($n=27$); or May to August, 1986 ($n=2$). Water chemistry samples were collected from 1 metre below the water surface using a Van Dorn sampler. Surface sediment samples were collected with an Ekman dredge. The top 1cm of sediment was carefully removed before the Ekman sampler was pulled from the water in order to prevent sediment disturbance or loss.

Laboratory methods

Water chemistry analyses were performed by the British Columbia MOE. Details of the analytical methods are described fully by McQuaker (1976). Chemical analyses were performed on unfiltered water samples and water chemistry data are measurements of total concentrations (which include both dissolved and particulate fractions).

Diatom slides were prepared using standard techniques (Smol, 1983). Small aliquots of the sediment were digested with strong acid (potassium dichromate and sulphuric acid, 5% w/v), and the resulting slurries were centrifuged and rinsed at least 8 times with distilled water. The cleaned diatom suspensions were pipetted onto coverslips and dried overnight on a warming tray. The coverslips were mounted in Hyrax[®] onto glass microscope slides, and diatoms were counted along transects. A minimum of 500 diatom valves were identified and enumerated from each sample under oil immersion using an Olympus BHS microscope (1000x magnification, N.A. = 1.40). Taxonomically difficult species were studied using a Hitachi S-2500 scanning electron microscope at 20kV. Diatom taxonomy follows Hustedt (1930), Patrick and Reimer (1966, 1975), Koppen (1975), Germain (1981), Håkansson and Stoermer (1984 a,b), Håkansson (1986), Stoermer and

diatom assemblages has vastly improved over the past decade, with the development of direct-gradient and unimodal statistical models that are based on weighted-averaging (WA) computational techniques. These models (e.g. canonical correspondence analysis (CCA), and WA regression and calibration) have only recently become available, and do not share many of the shortcomings of multiple linear regression. Ordination by CCA and WA reconstruction techniques assume that individual taxa respond in a unimodal manner over long environmental gradients. For this reason, they are theoretically superior to multiple linear regression models (ter Braak and van Dam, 1989) and they appear to perform better (Birks *et al.*, 1990). The development of these direct-gradient, multivariate methods of analysis and WA reconstruction techniques were largely stimulated by research on lake acidification (Birks *et al.*, 1990; Stevenson *et al.* 1991), where algal microfossils were used to reconstruct changes in lakewater pH (Birks, Juggins and Line, 1990) and metal concentrations (Birks *et al.*, 1990; Cumming *et al.*, 1992; Kingston *et al.*, 1992). More recently, these approaches have been expanded to paleolimnological inferences of lakewater salinity (Fritz, 1990) and temperature (Walker *et al.*, 1991). These techniques are directly applicable to the study of lake eutrophication, although calibration sets are only now being developed. For example, Christie (1988) used CCA to establish that the distributions of surficial sediment diatom taxa are related to lake depth, and nitrogen, phosphorus and chlorophyll *a* concentrations in 37 southeastern Ontario, lakes (Canada). And she developed a WA model to infer total nitrogen concentration from diatom microfossils. Similarly, Agbeti (1992) has developed a WA model for inferring total phosphorus and Secchi from diatom microfossils in lakes of eastern Canada.

Lake eutrophication is severe in some areas of British Columbia as a result of intensive agriculture (e.g. cattle ranching), and industrial and saw mill activity (Northcote and Larkin, 1963). These human activities often impair the quality of drinking water (Bothwell, 1985, 1992) and adversely affect the production of salmonid fishes (Northcote and Larkin, 1963). In order to assess these trends and provide data regarding their sources, we need an historical measure of lake trophic status. A diatom-based inference model would allow us to answer several important questions concerning eutrophic lakes. For example, was the lake naturally eutrophic? Did nutrient levels increase? And, if so, when did the nutrient increase occur? What type of land-use practice is associated with the nutrient increase? Are lakes presently recovering in response to abatement efforts?

The aim of this paper is to develop an ecologically sound, quantitative inference model that will provide a useful and accurate means for inferring past lake trophic status. To this end, we: (1) examine the relationship between the distribution of diatom taxa from the surface sediments of 46 lakes in British Columbia and 19 measured environmental variables to determine which variables exert an independent and significant influence on species distributions; and (2) develop a WA model for the quantitative inference of lakewater total phosphorus concentration (TP) from sediment diatom assemblages.

Introduction

Paleolimnology is the branch of aquatic science that deals with reconstructing lake histories, and paleolimnological techniques are particularly useful for inferring limnological data in cases where direct human measurements are unavailable (Smol, 1992). In North America many lakes have never been sampled, and detailed historical records of water chemistry are rarely available (Stockner, 1978; Likens, 1989). Furthermore, techniques for water chemistry analysis have improved over time, and historical measurements are often unreliable or cannot be directly compared to present-day measurements. A paleolimnological approach is therefore often the only means available to hindcast trends in water chemistry.

Paleolimnologists may use a large number of indicators to infer information of past limnological conditions (e.g. algal pigments and their derivatives, and algal and invertebrate microfossils; Brown, 1969; Warner, 1990). For a number of reasons, diatoms (class Bacillariophyceae) are the most widely used group of paleolimnological indicators, and they are particularly well suited to studies of lake eutrophication. Diatoms have restricted nutrient requirements (e.g. Titman, 1976) and they are sensitive indicators of lake trophic status (Hustedt, 1937-49; Lowe, 1974). The siliceous cell walls (valves) of diatoms are abundant and well preserved in most lake sediments, and can usually be identified to the specific or subspecific level. Sediment diatom assemblages are composed of a large number of taxa, and hence contain considerable ecological information (Birks *et al.*, 1990).

Previous studies have used diatoms to infer past changes in lake trophic status (Brugam, 1978; Ennis, Northcote and Stockner, 1983; Boucherle *et al.*, 1986). The majority of past studies have been restricted to qualitative descriptions of trophic history, with inferences based on a relatively small number of indicator species and/or indices (e.g. the Centrales:Pennales ratio (Nygaard, 1956) or the Araphidineae:Centrales ratio (Stockner and Benson, 1967; Stockner, 1971)).

Recent efforts to provide quantitative inferences of lake trophic status have involved the development of multiple linear regression models (Agbeti and Dickman, 1989; Whitmore 1989). These models involve grouping diatom taxa into a number of ecological categories, based either on information from literature sources (Agbeti and Dickman, 1989), or on their present-day distributions (Whitmore, 1989). Multiple linear regression models are useful, but we believe they have several shortcomings. These models assume a linear relationship between diatom taxa (or groups) and trophic status; this assumption is inconsistent with observations that species tend to respond to environmental variables in a unimodal fashion (Gause, 1930; Whittaker, 1956; ter Braak and van Dam, 1989). The coefficients of multiple linear regression equations are often unstable because the abundances of ecologically similar taxa tend to be highly correlated (Montgomery and Peck, 1982; ter Braak and van Dam, 1989). Furthermore, ecological information is lost when a large number of taxa are clumped into a small number of ecological groups.

Our ability to make quantitative inferences of lakewater chemistry from

Summary

1. The relationship between surficial sediment diatom taxa (Bacillariophyceae) and measured limnological variables in 46 British Columbia lakes was explored using canonical correspondence analysis (CCA). Lakewater total phosphorus concentration (TP), maximum lake depth, conductivity, and calcium concentration each account for independent and statistically significant directions of variation in the distribution of diatom taxa.
2. Weighted averaging (WA) models were developed to infer lakewater TP from the relative abundances of 131 diatom taxa in the surficial sediments of 37 lakes. WA regression and calibration with classical deshrinking provides the best model for TP reconstructions.
3. Our quantitative inference model has two major advantages over existing multiple linear regression models: (i) inferences are based on the responses of individual taxa to TP, and do not involve grouping the taxa into a small number of ecological categories; and (ii) the model assumes that diatoms respond to TP in a unimodal, rather than a linear, fashion.
4. The WA model can now be used to hindcast past lakewater TP, within the range $5\text{--}28\mu\text{g l}^{-1}$, from diatoms preserved in the sediments of British Columbia lakes. The model can provide quantitative estimates of the onset, rate, and magnitude of lake eutrophication in response to natural processes and human disturbances.

APPENDIX A

A RECENT PUBLICATION DETAILING THE
TOTAL PHOSPHORUS INFERENCE MODEL

File C:\WP1\MOE\BC Papers.RE1.
13/02/91

A WEIGHTED-AVERAGING REGRESSION AND CALIBRATION MODEL FOR
INFERRING TOTAL PHOSPHORUS CONCENTRATION FROM DIATOMS IN BRITISH
COLUMBIA (CANADA) LAKES

Roland I. Hall* and John P. Smol

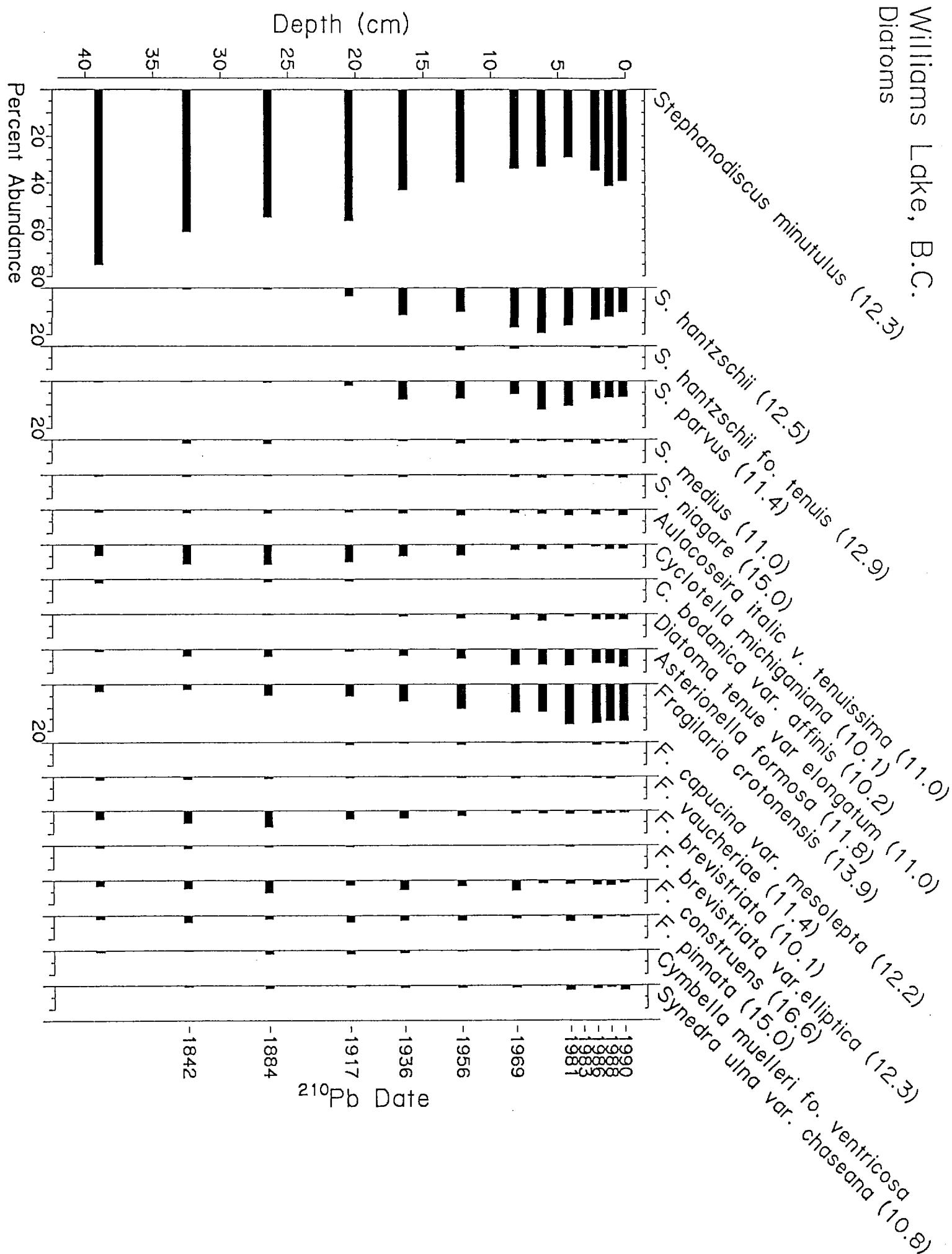
Paleoecological Environmental Assessment and Research Lab (PEARL), Department
of Biology, Queen's University, Kingston, Ontario, K7L 3N6, Canada.

* Author to whom proofs should be sent

Abbreviated title: Inferring TP from diatoms.

Revised January 13, 1992

Williams Lake, B.C. Diatoms



876.

- Likens G.E. (Ed.) (1989) Long-term studies in ecology. Springer-Verlag, New York. pp.214.
- Line J.M. & Birks H.J.B. (1990) WACALIB v. 2.1. A computer program to reconstruct environmental variables from fossil assemblages by weighted averaging. Journal of Paleolimnology, **3**, 170-173.
- Lowe R. (1974) Environmental requirements and pollution tolerances of freshwater diatoms. U.S. Environmental Protection Agency. Environmental Monitoring Series #EPA-67014-74-005: 1-333.
- McQuaker N.R. (1976) A laboratory manual for the chemical analysis of waters, wastewaters, sediments and biological tissues (2nd Ed.) Department of Environment, Water Resources Service, The Province of British Columbia.
- Montgomery D.C. & Peck E.A. (1982) Introduction to linear regression analysis. J. Wiley & Sons, New York, pp. 504.
- Northcote T.G. & Larkin P.A. (1963) Western Canada. In Frey D.G. (Ed.) Limnology in North America. University of Wisconsin Press, Madison, Wisconsin. pp. 451-485.
- Nygaard G. (1956) Ancient and recent flora of diatoms and Chrysophyceae in Lake Gribso. Folia Limnologica Scandinavica, **8**, 50-62.
- Patrick R. & Reimer C. (1966) The diatoms of the United States vol. 1. Academy of Natural Sciences, Philadelphia, Monograph, **3**, 1-668.
- Patrick R. & Reimer C. (1975) The diatoms of the United States vol 2 part 1. Academy of Natural Sciences, Philadelphia, Monograph, **13**, 1-213.
- Reynolds C.S. (1980) Phytoplankton assemblages and their periodicity in stratifying lake systems. Holarctic Ecology, **3**, 141-159.
- Smol J.P. (1983) Paleophycology of a high arctic lake near Cape Herschel, Ellesmere Island. Canadian Journal of Botany, **61**, 2195-2204.
- Smol J.P. (1992) Paleolimnology: An important tool for effective ecosystem management. Journal of Ecosystem Health (in press).
- Stevenson A.C. Juggins S., Birks H.J.B., Anderson D.S., Anderson N.J., Battarbee R.W., Berge F., Davis R.B., Flower R.J., Haworth E.Y., Jones V.J., Kingston J.C., Kreiser A.M., Line J.M., Munro M.A.R. & Renberg I. (1991) The surface waters acidification project palaeolimnology programme: modern diatom/lake-water chemistry data-set. ENSIS Publishing, London, England. pp. 86.
- Stockner J.G. (1971) Preliminary characteristics of lakes in the Experimental lakes Area, Northwestern Ontario, using diatom occurrence in the sediments. Journal of the Fisheries Research Board of Canada, **28**, 265-275.
- Stockner J.G. (1978) Man's impact on lakes and watersheds in Canada as assessed from paleolimnological studies. Polskie Archiwum Hydrobiologii, **25**, 393-402.
- Stockner J.G. & Benson W.W. (1967) The succession of diatom assemblages in the recent sediments of Lake Washington. Limnology and Oceanography, **12**, 513-532.

- Stoermer E.F. & Håkansson H. (1984) Stephanodiscus parvus: validation of an enigmatic and widely misconstrued taxon. Nova Hedwigia, **39**, 497-511.
- Stumm W. & Morgan J.J. (1981) Aquatic Chemistry. (2nd Ed.) John Wiley, New York
- ter Braak C.J.F. (1985) Correspondence analysis of incidence abundance data: properties in terms of a unimodal response model. Biometrics, **41**, 859-873.
- ter Braak C.J.F. (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology, **67**, 1167-1179.
- ter Braak C.J.F. (1987a). Calibration. In Jongman R.H.G., ter Braak C.J.F. & van Tongeren, O.F.R. (Eds.). Data analysis in community and landscape ecology. Pudoc, Wageningen, pp.78-90.
- ter Braak C.J.F. (1987b) Ordination. In Jongman R.H.G., ter Braak C.J.F. & van Tongeren, O.F.R. (Eds.). Data analysis in community and landscape ecology. Pudoc, Wageningen, pp.91-173.
- ter Braak C.J.F. (1988a) CANOCO - A FORTRAN program for canonical community ordination by (partial) (detrended) (canonical) correspondence analysis, principal components analysis, and redundancy analysis (version 2.1). Institute of Applied Computer Science, Statistical Department Wageningen, 6700 AC Wageningen, The Netherlands. Technical Report LWA-88-02, Wageningen, pp.95.
- ter Braak C.J.F. (1988b) Partial canonical correspondence analysis. In Brock H.H. (Ed.) Classification and related methods of data analysis. North Holland, Amsterdam. pp. 551-558.
- ter Braak C.J.F. (1990a) CANOCO - version 3.10. Unpublished computer program, Agricultural Mathematics Group, 6700 AC Wageningen.
- ter Braak C.J.F. (1990b) Update notes: CANOCO - version 3.10. Unpublished computer program, Agricultural Mathematics Group, 6700 AC Wageningen, pp.35.
- ter Braak C.J.F. & Barendregt L.G. (1986) Weighted averaging of species indicator values: its efficiency in environmental calibration. Mathematical Biosciences, **78**, 57-72.
- ter Braak C.J.F. & Looman C.W.N. (1986) Weighted averaging logistic regression and the Gaussian response model. Vegetatio, **65**, 3-11.
- ter Braak C.J.F. & Prentice I.C. (1988) A theory of gradient analysis. Advances in Ecological Research, **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.
- Titman D. (1976) Ecological competition between algae: Experimental confirmation of resource-based theory. Science, **192**, 463-465.
- Walker I.R., Smol J.P., Engstrom D.R. & Birks H.J.B. (1991) An assessment of Chironomidae as quantitative indicators of past climatic change. Canadian Journal of Fisheries and Aquatic Science, **48**, 975-987.
- Wallach D. & Goffinet B. (1989) Mean squared error of prediction as a criterion for evaluating and comparing systems models. Ecological Modelling, **44**, 299-

306.

- Warner B.G. (Ed.) (1990) Methods in Quaternary Ecology. Geoscience Canada Reprint Series 5. Geological Association of Canada, St. John's, Newfoundland. pp. 170.
- Whitmore T.J. (1989) Florida diatom assemblages as indicators of trophic state and pH. Limnology and Oceanography, **34**, 882-895.
- Whittaker H. (1956) Vegetation of the Great Smoky Mountains. Ecological Monographs, **26**, 1-80.
- Wilkinson L. (1988) SYSTAT: The system for statistics. SYSTAT Inc., Evanston, Illinois.
- Zar J.H. (1984). Biostatistical analysis. (2nd Ed.) Prentice-Hall, Inc., Englewood Cliffs, N.J.

Table 1. List of the 46 lakes included in the ordination analyses and in the development of the diatom-based total phosphorus calibration equations, with selected geographical, morphological and limnological data. **Bold numbers** indicate median values inserted where data points were missing. Biogeoclimatic zone IDF = Interior Douglas Fir, CDF = Coastal Douglas Fir, IWH = Interior Western Hemlock, and CALP = Cariboo Aspen - Lodgepole Pine.

Lake Name	Lake Number	Location (N°,W°)	Biogeoclimatic zone	Total Phosphorus ($\mu\text{g l}^{-1}$)	Maximum Depth (m)	Conductivity ($\mu\text{S cm}^{-1}$)	Calcium (mg l^{-1})
Langford	1	48°30',123°40'	CDF	9	15	160	23.9
King West	2	51°45',120°50'	IDF	10	18	161	25.8
Hilltout	3	55°00',127°00'	CALP	9	13	75	10.4
King East	4	51°45',120°50'	IDF	11	13	147	22.2
Kootenay	5	49°40',116°50'	IWH	12	154	157	21.7
Gardom	6	50°40',119°10'	IDF	16	24	455	43.9
Davis	7	49°50',120°45'	IDF	13	17	318	44.5
McCaffrey	8	49°30',120°30'	IDF	9	11	399	55.0
Pavillion	9	50°55',121°40'	IDF	10	56	392	40.7
Laird	10	49°30',120°30'	IDF	9	11	390	53.0
Boss	11	49°50',120°45'	IDF	14	11	227	29.7
Echo	12	51°10',118°45'	IDF	8	50	244	42.4
Tyhee	13	54°45',127°15'	CALP	21	22	261	33.2
Champion #3	14	49°15',117°30'	IWH	5	24	101	13.8
McQuarrie	15	54°45',126°50'	CALP	9	11	92	13.8

Table 1 continued

Lake Name	Lake Number	Location (N°,W°)	Biogeo-climatic zone	Total Phosphorus ($\mu\text{g l}^{-1}$)	Maximum Depth (m)	Conductivity ($\mu\text{S cm}^{-1}$)	Calcium (mg l^{-1})
Florence	16	48°30',123°30'	CDF	15	6	132	16.2
Wilgress	17	49°10',118°40'	IDF	8	18	410	55.7
Green	18	52°30',121°00'	IDF	15	17	239	36.1
Jewel	19	49°10',118°40'	IDF	13	21	158	19.9
Champion #2	20	49°15',117°30'	IWH	7	11	120	19.5
Lac de Roche	21	51°45',120°45'	IDF	10	47	229	34.3
Columbia	22	50°15',115°15'	IDF	12	6	320	31.6
Three Valley	23	50°50',118°15'	IDF	7	17	63	23.8
Whiteswan	24	50°15',115°30'	IDF	10	20	268	36.2
Alces	25	50°15',115°50'	IDF	7	15	330	44.1
Holden	26	49°05',123°45'	CDF	45	6	110	8.6
Elk	27	48°30',123°25'	CDF	15	15	159	15.6
Windermere	28	50°30',115°50'	IDF	8	8	240	23.8
Glen	29	48°30',123°30'	CDF	12	15	137	10.5
Durrance	30	48°30',123°25'	CDF	9	17	154	21.4
Sneezie	31	51°55',121°10'	CALP	9	24	252	41.8
Tahla	32	49°50',120°45'	IDF	12	8	278	39.7
Spanish	33	52°35',121°20'	IDF	8	73	97	13.6

Table 1 continued

Lake Name	Lake Number	Location (N°,W°)	Biogeo-climatic zone	Total Phosphorus ($\mu\text{g l}^{-1}$)	Maximum Depth (m)	Conductivity ($\mu\text{S cm}^{-1}$)	Calcium (mg l^{-1})
Spout	34	52°00',121°20'	CALP	17	18	225	31.5
Lang	35	52°00',121°00'	IDF	14	73	95	10.9
Hawkins	36	51°55',120°55'	IDF	24	18	193	23.9
Morehead	37	52°35',121°45'	IDF	26	17	82	11.1
Little Bobtail	38	53°40',123°40'	CALP	85	15	167	15.0
Keno	39	52°20',121°20'	IWH	12	33	111	15.9
Woodcock	40	53°40',123°40'	CALP	28	15	116	13.4
Elkin	41	51°40',123°45'	CALP	17	35	71	9.0
Dahl	42	53°50',123°15'	CALP	19	24	131	11.0
Milburn	43	53°00',122°40'	CALP	23	8	191	23.7
Murphy	44	52°00',121°15'	IDF	19	23	164	20.0
Aid	45	51°40',117°20'	IDF	8	5	228	29.2
Comfort	46	51°40',117°20'	IDF	5	6	269	39.2

Table 2. Matrix of Pearson product-moment correlations between environmental variables (* indicates significant correlations at $\alpha = 0.05$, based on Bonferroni adjusted probabilities).

	Alt.	Water-shed	Surf. area	Mean depth	Max. depth	Secchi	pH	TDS	Cond.	Hard.	TP	Ca	Fe	Mg	Mn	Al
Altitude	1.000															
Watershed	0.176	1.000														
Surface area	0.040	0.722*	1.000													
Mean depth	0.005	0.191	0.344	1.000												
Max. depth	0.117	0.241	0.391	0.919*	1.000											
Secchi	0.284	-0.187	-0.207	0.323	0.329	1.000										
pH	0.169	0.113	0.167	0.066	-0.027	0.310	1.000									
TDS	0.132	-0.043	-0.152	0.005	-0.104	0.360	0.713*	1.000								
Conductivity	0.151	-0.033	-0.140	-0.003	-0.131	0.376	0.737*	0.984*	1.000							
Hardness	0.220	0.060	-0.108	-0.018	-0.122	0.373	0.704*	0.936*	0.956*	1.000						
TP	-0.247	0.177	0.287	-0.043	-0.056	-0.532*	-0.124	-0.191	-0.248	-0.331	1.000					
Ca	0.223	-0.113	-0.256	0.024	-0.076	0.440	0.666*	0.922*	0.929*	0.952*	-0.378	1.000				
Fe	-0.240	-0.142	-0.195	-0.307	-0.304	-0.520*	-0.409	-0.245	-0.307	-0.314	0.342	-0.356	1.000			
Mg	0.157	0.312	0.166	-0.041	-0.121	0.107	0.611*	0.765*	0.780*	0.861*	-0.100	0.690*	0.149	1.000		
Mn	-0.337	-0.078	-0.154	-0.347	-0.382	-0.522*	-0.217	-0.066	-0.217	-0.188	0.577*	-0.193	0.571*	-0.077	1.000	
Al	0.089	0.037	0.113	-0.058	-0.057	-0.023	0.059	0.068	0.066	0.012	0.224	-0.018	0.096	0.064	0.145	1.000

Table 3. Canonical coefficients of the 8 environmental variables included in the Canonical Correspondence Analysis and their approximate t-values (* indicates t-test significant at $p = 0.05$; $n = 41$, d.f. = 32).

Environmental Variable	Canonical coefficients Axis 1	Canonical coefficients Axis 2	t-values of canonical coefficients Axis 1	t-values of canonical coefficients Axis 2
Altitude	0.16	0.00	1.72	0.03
Surface area	0.27	0.13	2.36*	0.98
Maximum depth	-0.56	0.37	-5.02*	2.83*
Secchi	0.23	0.45	1.88	3.18*
Conductivity	1.65	0.48	5.91*	1.47
Total phosphorus	-0.30	0.64	-2.77*	4.99*
Calcium	-1.37	0.20	-4.81*	0.59
Aluminum	0.16	-0.29	1.73	-2.67*

Table 4. Total phosphorus optima and number of occurrences for the 131 diatom taxa included in the weighted-averaging model, sorted by ascending TP optimum. Taxa with numbers followed by "BC" are taxa of unknown taxonomic position that were retained as distinct in this study.

Number	Diatom Taxon	TP Optimum ($\mu\text{g l}^{-1}$)	Number of Occurrences
1	<u>Aulacoseira</u> sp.12 PIRLA	6.8	2
2	<u>Achnanthes</u> <u>levanderi</u> Hust.	6.9	3
3	<u>Anomoeoneis</u> <u>vitrea</u> (Grun.) R. Ross	7.3	8
4	<u>Achnanthes</u> <u>bioreti</u> Germ.	7.3	3
5	<u>Cyclotella</u> <u>comensis</u> Grun.	7.4	7
6	<u>Aulacoseira</u> <u>distans</u> (Ehrenb.) Simonsen	7.8	10
7	<u>Fragilaria</u> <u>virescens</u> var. <u>exigua</u> Grun.	7.8	2
8	<u>Cyclotella</u> <u>kuetzingiana</u> var. <u>planetophora</u> Fricke	7.9	2
9	<u>Neidium</u> <u>ampliatum</u> (Ehrenb.) Krammer	8.0	1
10	<u>Cyclotella</u> <u>kuetzingiana</u> var. <u>radiosa</u> Fricke	8.3	5
11	<u>Cymbella</u> <u>delicatula</u> Kütz.	8.3	7
12	<u>Amphora</u> <u>parallelstriata</u> Manguin	8.4	3
13	<u>Gomphonema</u> <u>angustum</u> Agardh.	8.6	4
14	<u>Achnanthes</u> <u>microcephala</u> (Kütz.) Cleve	8.8	9
15	<u>Aulacoseira</u> <u>lirata</u> (Ehrenb.) R. Ross	8.9	3
16	<u>Cyclotella</u> <u>meneghiniana</u> Kütz.	9.0	2
17	<u>Synedra</u> <u>radians</u> Kütz.	9.0	2
18	<u>Cymbella</u> <u>cesatii</u> (Rabenh.) Grun.	9.0	2
19	<u>Stephanodiscus</u> sp.#5 BC	9.0	2

Number	Diatom Taxon	TP Optimum ($\mu\text{g l}^{-1}$)	Number of Occurrences
20	<u>Achnanthes minutissima</u> Kütz.	9.0	27
21	<u>Diatoma hiemale</u> var. <u>mesodon</u> (Ehrenb.) Grun.	9.2	5
22	<u>Navicula utermoehlii</u> Hust.	9.2	6
23	<u>Achnanthes conspicua</u> A. Mayer	9.3	21
24	<u>Cymbella diluviana</u> (Krasske) M.B. Florin	9.3	7
25	<u>Achnanthes linearis</u> (W. Smith) Grun.	9.3	14
26	<u>Aulacoseira perglabra</u> var. <u>florinae</u> Camburn	9.3	4
27	<u>Aulacosiera distans</u> var. <u>nivalis</u> (W. Smith) Kirchner	9.5	3
28	<u>Tabellaria fenestrata</u> (Lyngb.) Kütz.	9.5	2
29	<u>Fragilaria</u> sp.3 PIRLA	9.5	3
30	<u>Stauroneis phoenicenteron</u> (Nitzsch) Ehrenb.	9.6	6
31	<u>Amphora thumensis</u> (Mayer) Krieger	9.7	12
32	<u>Cyclotella stelligera</u> (Cleve & Grun.) Van Heurck	9.7	24
33	<u>Fragilaria</u> sp.4 PIRLA	9.9	7
34	<u>Mastogloia smithii</u> Thwaites ex W. Smith	10.0	6
35	<u>Cyclotella michiganiana</u> Skvortzow	10.1	19
36	<u>Fragilaria brevistriata</u> Grun.	10.1	36
37	<u>Synedra parasitica</u> (W. Smith) Hust.	10.1	15
38	<u>Cymbella minuta</u> Hilse ex Rabenh.	10.2	14
39	<u>Gomphonema parvulum</u> (Kütz.) Kütz.	10.2	4
40	<u>Fragilaria brevistriata</u> var. <u>inflata</u> (Pant.) Hust.	10.2	22
41	<u>Navicula laevissima</u> Kütz.	10.2	12
42	<u>Cyclotella bodanica</u> var. <u>affinis</u> Grun.	10.2	16

Number	Diatom Taxon	TP Optimum ($\mu\text{g l}^{-1}$)	Number of Occurrences
43	<u>Navicula</u> <u>radiosa</u> Kütz.	10.3	9
44	<u>Amphora</u> <u>pediculus</u> (Kütz.) Grun. ex A. Schmidt	10.4	32
45	<u>Navicula</u> <u>subminiscula</u> Manguin	10.4	13
46	<u>Fragilaria</u> sp.2 PIRLA	10.4	14
47	<u>Achnanthes</u> <u>biasolettiana</u> Grun.	10.5	16
48	<u>Navicula</u> <u>explanata</u> Hust.	10.5	2
49	<u>Achnanthes</u> <u>pinnata</u> Hust.	10.7	4
50	<u>Synedra</u> <u>ulna</u> var. <u>chaseana</u> Thomas	10.8	4
51	<u>Navicula</u> <u>cryptocephala</u> Kütz.	10.8	12
52	<u>Navicula</u> <u>minima</u> Grun.	10.9	17
53	<u>Aulacoseira</u> <u>italica</u> var. <u>tenuissima</u> (Grun.) Simonsen	11.0	11
54	<u>Diatoma</u> <u>tenue</u> var. <u>elongatum</u> Lyngb.	11.0	7
55	<u>Rhopalodia</u> <u>gibba</u> (Ehrenb.) O. Müll.	11.0	5
56	<u>Stephanodiscus</u> <u>medius</u> Håkansson	11.0	10
57	<u>Gomphonema</u> <u>angustatum</u> (Kütz.) Rabenh.	11.1	6
58	<u>Fragilaria</u> sp.1 BC	11.2	16
59	<u>Amphora</u> <u>libyca</u> Ehrenb. ex Kütz.	11.2	23
60	<u>Achnanthes</u> <u>suchlandtii</u> Hust.	11.2	7
61	<u>Cyclotella</u> cf. <u>radiosa</u> (Grun.) Lemm.	11.3	3
62	<u>Stephanodiscus</u> <u>alpinus</u> Hust.	11.3	13
63	<u>Fragilaria</u> <u>vaucheriae</u> (Kütz.) J.B. Petersen	11.4	8
64	<u>Synedra</u> <u>delicatissima</u> Grun.	11.4	3
65	<u>Fragilaria</u> <u>pinnata</u> var. <u>intercedens</u> (Grun.) Hust.	11.4	21
66	<u>Stephanodiscus</u> <u>parvus</u> Stoermer & Håkansson	11.4	12

Table 4 continued

Number	Diatom Taxon	TP Optimum ($\mu\text{g l}^{-1}$)	Number of Occurrences
67	<u>Nitzschia fonticola</u> Grun.	11.5	21
68	<u>Stephanodiscus</u> cf. <u>vestibulis</u> Håkansson, Theriot, & Stoermer	11.6	2
69	<u>Cocconeis diminuta</u> Pant.	11.6	21
70	<u>Synedra filiformis</u> var. <u>exilis</u> Cleve-Euler	11.6	5
71	<u>Cymbella microcephala</u> Grun.	11.6	16
72	<u>Synedra cyclopus</u> Brutschy	11.7	11
73	<u>Achnanthes clevei</u> Grun.	11.7	17
74	<u>Asterionella ralfsii</u> W. Smith	11.7	4
75	<u>Cyclotella bodanica</u> Grun.	11.7	26
76	<u>Nitzschia denticula</u> Grun.	11.7	13
77	<u>Navicula shadei</u> Krasske	11.8	6
78	<u>Fragilaria lapponica</u> Grun.	11.8	28
79	<u>Asterionella formosa</u> Hasall	11.8	25
80	<u>Navicula vitabunda</u> Hust.	11.8	19
81	<u>Synedra minuscula</u> Grun.	12.0	1
82	<u>Fragilaria construens</u> var. <u>venter</u> (Ehrenb.) Grun.	12.0	34
83	<u>Navicula pupula</u> Kütz.	12.1	17
84	<u>Fragilaria pinnata</u> var. <u>lancettula</u> (Schum.) Hust.	12.1	28
85	<u>Cocconeis placentula</u> var. <u>lineata</u> (Ehrenb.) Van Heurck	12.1	10
86	<u>Fragilaria capucina</u> var. <u>mesolepta</u> (Rabenh.) Rabenh.	12.2	9
87	<u>Achnanthes detha</u> Hohn & Hellerman	12.2	14
88	<u>Fragilaria construens</u> var. <u>binodis</u> (Ehrenb.) Grun.	12.3	12

Number	Diatom Taxon	TP Optimum ($\mu\text{g l}^{-1}$)	Number of Occurrences
89	<u>Achnanthes lanceolata</u> (Bréb. ex Kütz.) Grun.	12.3	19
90	<u>Stephanodiscus minutulus</u> (Kütz.) Cleve & Möller	12.3	24
91	<u>Fragilaria brevistriata</u> var. <u>elliptica</u> Hérib.	12.3	15
92	<u>Navicula submuralis</u> Hust.	12.5	15
93	<u>Stephanodiscus hantzschii</u> Grun.	12.5	13
94	<u>Achnanthes lanceolata</u> var. <u>dubia</u> Grun.	12.6	18
95	<u>Navicula seminuloides</u> Hust.	12.6	16
96	<u>Achnanthes exigua</u> var. <u>heterovalva</u> Krasske	12.6	21
97	<u>Achnanthes exigua</u> Grun.	12.7	24
98	<u>Denticula elegans</u> Kütz.	12.7	6
99	<u>Navicula cari</u> Ehrenb.	12.8	13
100	<u>Cocconeis placentula</u> Ehrenb.	12.9	8
101	<u>Stephanodiscus hantzschii</u> fo. <u>tenuis</u> (Hust.) Håkansson & Stoermer	12.9	5
102	<u>Navicula pseudoventralis</u> Hust.	12.9	16
103	<u>Stephanodiscus</u> sp.1 BC	13.0	1
104	<u>Fragilaria brevistriata</u> var. <u>capitata</u> Hérib.	13.2	6
105	<u>Aulacoseira italica</u> ssp. <u>subarctica</u> (O. Müll.) Simonsen	13.2	13
106	<u>Cyclotella kuetzingiana</u> Thwaites	13.6	11
107	<u>Aulacoseira undulata</u> (Ehrenb.) Kütz.	13.7	4
108	<u>Fragilaria crotonesis</u> Kitton	13.9	21
109	<u>Cocconeis placentula</u> var. <u>euglypta</u> (Ehrenb.) Grun.	14.0	8
110	<u>Achnanthes lanceolata</u> var. <u>elliptica</u> Cleve	14.3	15
111	<u>Cyclotella ocellata</u> Pant.	14.5	8

Number	Diatom Taxon	TP Optimum ($\mu\text{g l}^{-1}$)	Number of Occurrences
112	<u>Stephanodiscus niagare</u> Ehrenb.	15.0	14
113	<u>Fragilaria pinnata</u> Ehrenb.	15.0	36
114	<u>Tabellaria flocculosa</u> str. IIIp sensu Koppen	15.1	7
115	<u>Navicula absoluta</u> Hust.	15.2	11
116	<u>Aulacoseira lirata</u> var. <u>lacustris</u> (Grun.) R. Ross	15.8	6
117	<u>Epithemia argus</u> (Ehrenb.) Kütz.	16.4	9
118	<u>Stephanodiscus</u> cf. <u>pseudoexcentricus</u> Håkansson & Stoermer	16.5	3
119	<u>Fragilaria construens</u> (Ehrenb.) Grun.	16.6	28
120	<u>Aulacoseira ambigua</u> (Grun.) Simonsen	16.9	18
121	<u>Navicula scutelloides</u> W. Smith ex Greg.	17.3	8
122	<u>Cyclostephanos</u> sp.1 BC	17.5	3
123	<u>Tabellaria flocculosa</u> str. III sensu Koppen	17.6	7
124	<u>Achnanthes delicatula</u> (Kütz.) Grun.	18.6	4
125	<u>Navicula ventralis</u> Krasske	18.7	4
126	<u>Pinnularia nodosa</u> (Ehrenb.) W. Smith	20.6	3
127	<u>Gyrosigma acuminatum</u> (Kütz.) Rabenh.	20.6	3
128	<u>Stauroneis anceps</u> Ehrenb.	21.0	6
129	<u>Gomphonema subtile</u> Ehrenb.	21.5	3
130	<u>Pinnularia braunii</u> (Grun.) Cleve	21.6	2
131	<u>Aulacoseira granulata</u> var. <u>angustissima</u> (O. Müll.) Simonsen	23.3	2

Table 5. Selected limnological characteristics of 37 British Columbia lakes used in the WA regression and calibration. Data are provided by the British Columbia Ministry of the Environment.

Environmental Variables	Minimum	Maximum	Median	Mean	Standard deviation
Altitude (m)	61	1275	810	798	342
Surface area (ha)	9	2574	45	302	519
Maximum depth (m)	5	73	15	21.6	16.0
Secchi (m)	1.3	9.9	3.3	4.6	2.0
Conductivity ($\mu\text{S cm}^{-1}$)	63	410	147	197	100
Total phosphorus ($\mu\text{g l}^{-1}$)	5	28	9	13	5.7
Calcium (mg l^{-1})	9.0	55.7	19.9	26.0	13.2
Aluminum ($\mu\text{g l}^{-1}$)	20	80	20	35.4	16.9
pH	7.5	8.5	7.9	8.1	0.3

Table 6. Apparent root mean squared error (RMSE) of prediction, and correlation between observed and inferred total phosphorus concentration for different calibration and deshrinking procedures. Values were derived using 131 diatom taxa from the 37 lake calibration set.

Calibration procedure	Deshrinking type	Apparent RMSE ($\mu\text{g l}^{-1}$)	Correlation (r)
WA	Classical	0.25	0.86
WA(tol)	Classical	0.28	0.84
WA	Inverse	0.21	0.86
WA(tol)	Inverse	0.23	0.84

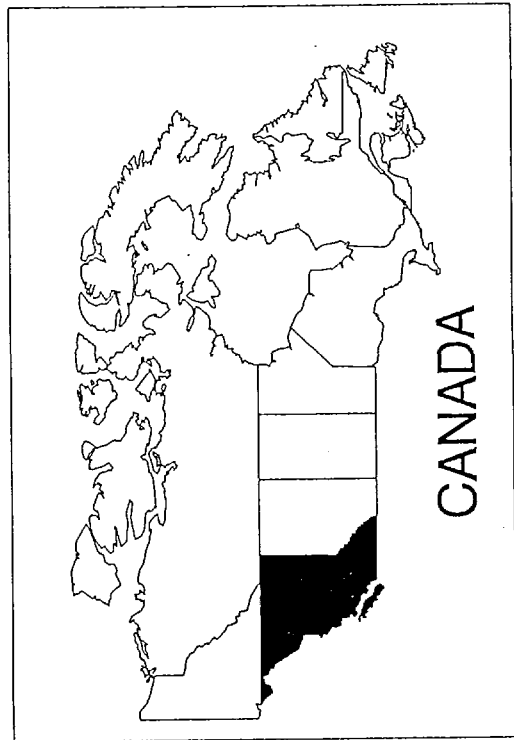
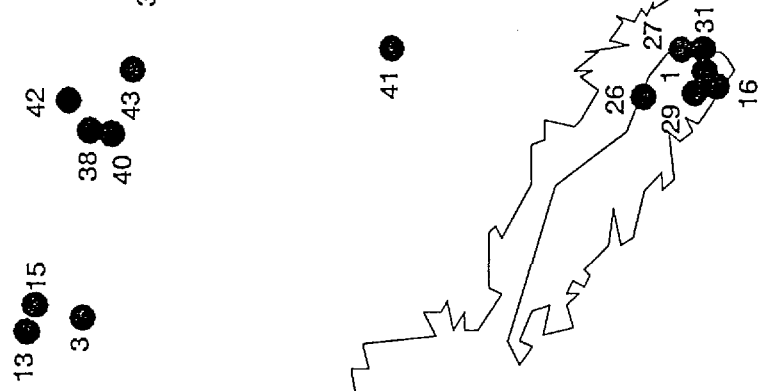
Figure legends

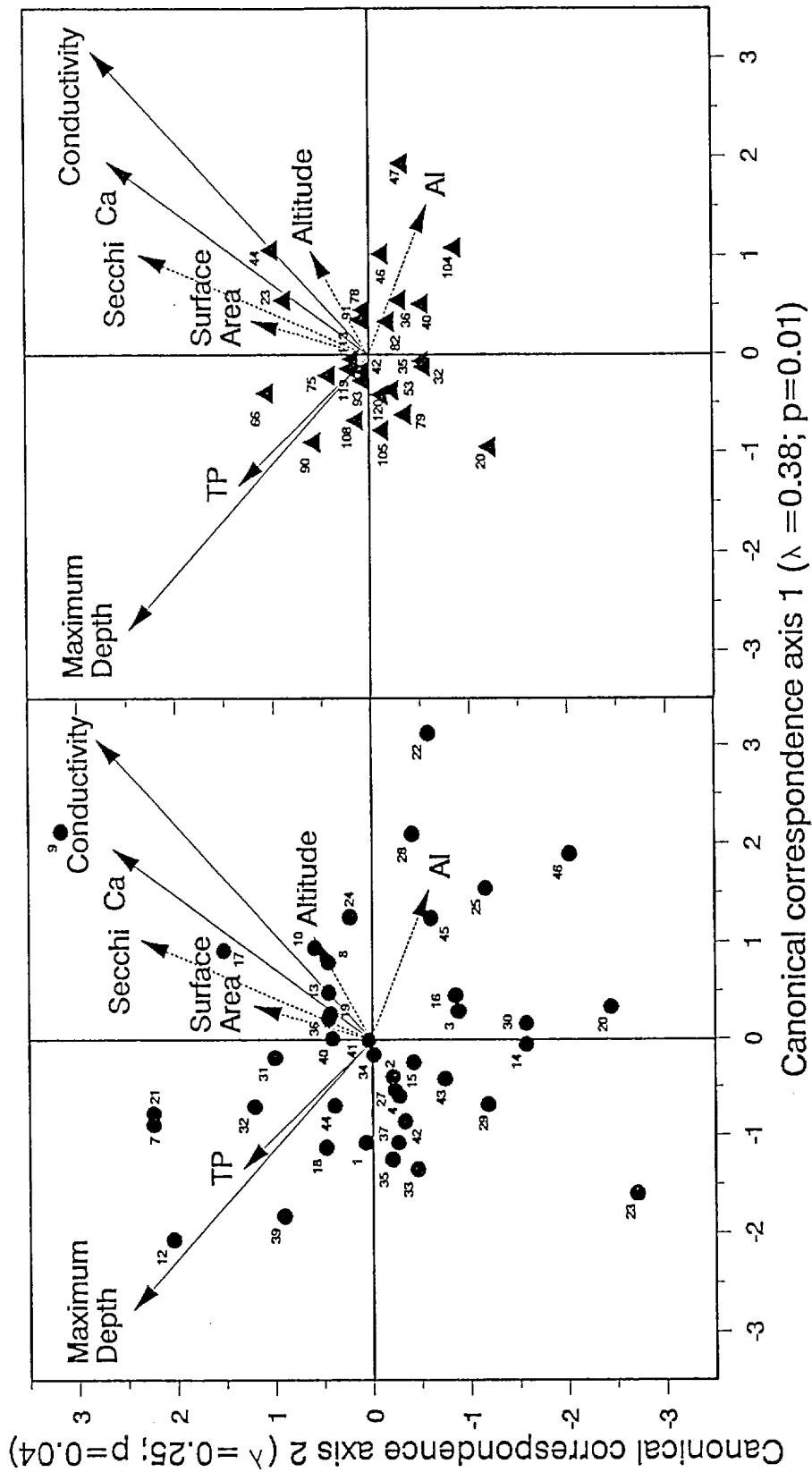
Figure 1. Map of Canada and British Columbia showing the locations of the 46 study lakes. Lake numbers are listed in Table 1.

Figure 2. Canonical correspondence analysis of the 41 lake set. Circles represent lakes and triangles represent diatom taxa. Lake names are presented in Table 1 and taxon names are presented in Table 4. Arrows with solid lines indicate environmental variables that exert significant and independent influences on the distributions of diatom taxa, as detected by forward selection in CCA. Other environmental variables included in the analysis are indicated by arrows with broken lines. Maximum depth, surface area, total phosphorus concentration (TP), and aluminum concentration (Al) were $\ln(x + 1)$ transformed.

Figure 3. Plots of observed vs. diatom-inferred total phosphorus concentration (TP), and observed vs. residual TP (or, observed minus diatom-inferred TP), based on weighted-averaging regression and calibration models using (a) inverse deshrinking and (b) classical deshrinking. The large circles indicate 2 coincident values.

BRITISH COLUMBIA

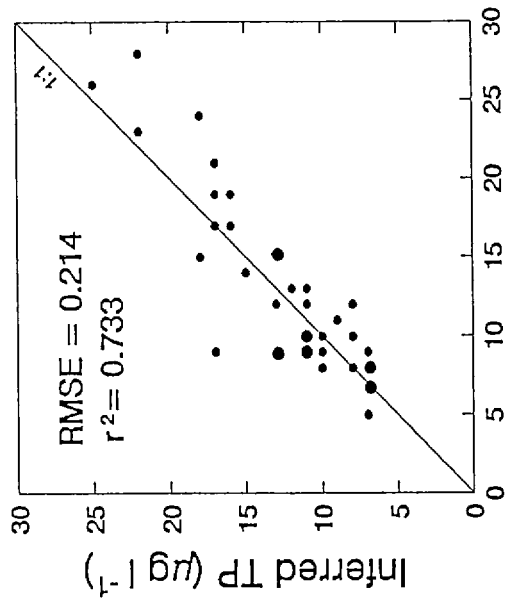




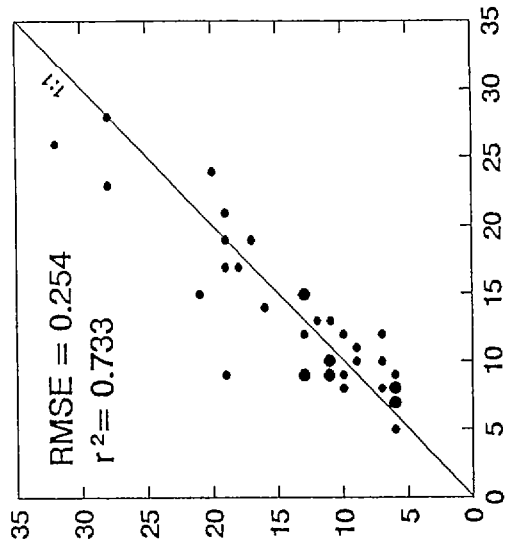
APPENDIX B

DIATOM PERCENT ABUNDANCE DATA FOR ALL TAXA
INCLUDED IN THE TOTAL PHOSPHORUS INFERENCE MODEL

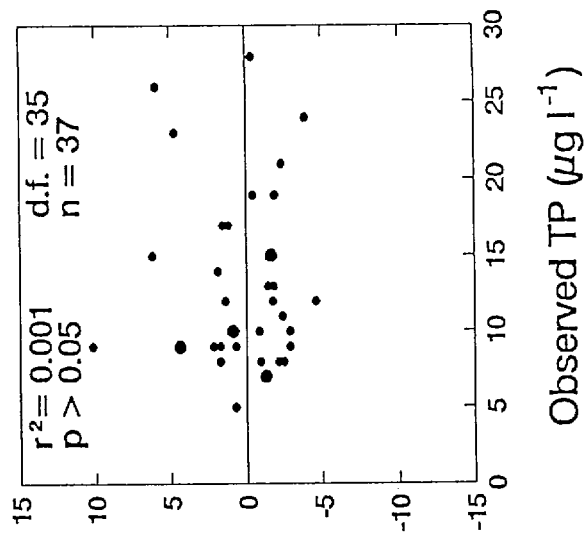
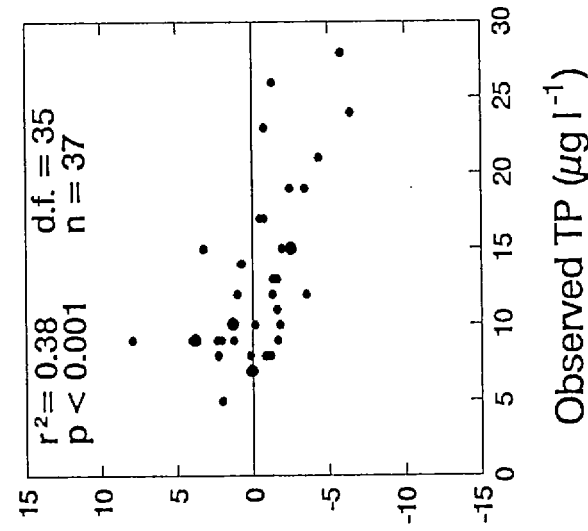
a) WA -- Inverse Deshrinking



b) WA -- Classical Deshrinking



Observed TP - Inferred TP ($\mu\text{g l}^{-1}$)



Tabelaria

parasitica (W.Smith) Hust.	10.1	SYNPAR	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
radians Kütz.	9	SYNRAD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ulna var. chaseana Thomas	10.8	SYNULC	1.9	0.7	1.2	1.9	0.2	0.9	0.9	0.0	0.8	1.3	0.5	0.0
fenestrata (Lyngb.) Kütz.	9.5	TABFEN	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
fiocculosa II sensu Koppen	15.1	TABFL03	0.2	0.9	0.2	0.0	0.4	0.9	0.4	0.2	0.2	0.0	0.0	0.0
fiocculosa IIp sensu Koppen	17.6	TABFL3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	TOTAL		96.2	97.4	96.2	96.4	96.5	96.4	96.6	94.9	93.1	96.7	96.5	98.1

TAXA	WA - Optimum totalP (µg/L)	Sediment Depth (cm)	0.25	1.25	2.25	4.25	6.25	8.25	12.25	16.50	20.50	26.50	32.50	39.00
Achnanthes	bisoletiana Grun.	10.5	ABIASO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	bioret Germ.	7.3	ABIORE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	clevei Grun.	11.7	ACLEVEI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.3	0.0
	conspicua A. Mayer	9.3	ACONS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	delicatula (Kutz.) Grun.	18.6	ACHDEL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	detha Hohn and Heileman	12.2	ACDETH	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	exigua Grun.	12.7	AEXIGU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.3	0.0
	exigua var. heterovalva Krasske	12.6	AEXVHE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	lanceolata (Breb. ex Kutz.) Grun	12.3	ALANCE	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.5	0.0
	lanceolata var. dubia Grun.	12.6	ALANDU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	lanceolata var. elliptica Cleve	14.3	ALANEL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	leopardi Hust.	6.9	ACHLEV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	linearis (W. Smith) Grun.	9.3	ALINEAR	0.0	0.0	0.2	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
	microcephala (Kutz.) Cleve	6.8	AMICRO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	minutissima Kutz.	9.0	AMINUTI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Amphora	pinata Hust.	10.7	APINNA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	suchlandti Hust.	11.2	ASUCHL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	illycia Ehrenb. ex Kutz.	11.2	AMPILBY	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.2	0.0	0.2	0.0
	parallelistrata Mangun	8.4	AMPAPAR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	pedicularis (Kutz.) Grun. ex A.Schmidt	10.4	AMPEDP	0.0	0.0	0.0	0.2	0.2	0.5	0.0	0.8	0.4	0.8	0.2
	thurensis (Mayer) Krüger	9.7	AMPTHU	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	vitrea (Grun.) R.Floss	7.3	AMVITV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	formosa Hassall	11.8	ASTFOR	7.5	6.1	5.8	6.8	6.5	3.9	2.7	1.0	3.1	2.9	0.9
	raisi W. Smith	11.7	ASTRAL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	diminuta Part.	11.6	COODIM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.2
	placenta Ehrenb.	12.9	COCPPLA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
	placenta var. linearis (Ehr.) Van Heurck	12.1	CPALIN	0.6	0.9	0.3	0.2	0.7	0.5	0.5	0.4	0.4	0.2	0.0
	planorbula var. euglypta (Ehr.) Grun.	14.0	CPLEU	0.2	0.4	0.0	0.2	0.2	0.0	0.0	0.0	0.2	0.0	0.0
	bodanica Grun.	11.7	CYCBO	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Cocconeis	bodanica var. affinis Grun.	10.2	CBODAF	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	1.1	0.2
commensis Grun.		7.4	CYCOCO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
kuetzingiana Thwaites		13.6	CYCKUE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0
kuetzingiana var. radiosa Fricke		8.3	CKUEFR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
kuetzingiana var. planetophora Fricke		7.9	CKUEPL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mereghiniana Kutz.		9.0	CYOCME	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.4
michiganiana Skvortzow		10.1	CYOMIC	1.7	1.8	0.7	1.7	1.9	2.2	4.6	4.9	7.5	8.6	4.7
ocellata Part.		14.5	CYCOC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
cf. radiosa (Grun.) Lemm.		11.3	CYCRAD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
stelligera (Cleve&Grun.) Van Heurck		9.7	CYCSTE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sp. 1 BC		17.5	CYCOSP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
cesatii (Fabenb.) Grun.		9.0	CYMCCE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
delicatula Kutz.		8.3	CYMDCL	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
diluviana (Krasske) M.B. Florin		9.3	CYMDIL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
microcephala Grun.		11.6	CYMICR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cyclotella	minuta Hilse ex Fabenb.	10.2	CYMINU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	elegans Kutz.	12.7	DENELE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0
	tenue var. elongatum Lyngb.	11.0	DIENTE	2.3	2.4	2.4	1.0	2.8	1.8	0.9	0.2	0.2	0.0	0.0
	hiemale var. mesodon (Ehrenb.) Grun.	9.2	DHIEME	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	argus (Ehrenb.) Kutz.	16.4	EPIARG	0.4	0.0	0.0	0.0	0.2	0.2	0.0	0.0	0.4	0.0	0.6
	brevistriata Grun.	10.1	FRABRE	0.8	0.9	1.4	1.3	1.2	2.1	3.3	3.6	6.9	5.3	3.6
	brevistriata var. inflata (Part.) Hust.	10.2	FRBINF	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0
	brevistriata var. capitata Herth.	13.2	FRBPCA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5
	brevistriata var. elliptica Herth.	12.3	FRBREL	0.0	0.0	0.0	0.6	0.2	0.0	0.2	0.6	0.5	1.4	0.9
	capucina var. mesolepta (Fabenb.) Fabenb.	12.2	FRICAPM	0.9	0.6	1.0	0.2	0.2	1.1	0.4	1.2	0.0	0.0	0.0
	constuens (Ehrenb.) Grun.	16.6	FRACON	1.1	2.0	1.9	1.5	1.2	2.5	4.2	2.1	5.3	3.6	2.5
	constuens var. binodis (Ehrenb.) Grun.	12.3	FRACOBI	0.0	0.0	0.0	0.0	0.0	0.7	0.4	0.4	0.0	0.3	0.0
	constuens var. brevisstriata	11.2	FRCOBR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	constuens var. venter (Ehrenb.) Grun.	12.0	FRCCOV	0.0	0.0	0.0	0.0	0.4	0.2	0.0	0.4	0.8	0.5	0.3
	crotonensis Kitton	13.9	FRACRO	16.0	15.9	16.8	17.3	11.7	12.0	10.4	7.3	5.0	2.0	2.8
lapponica Grun.	11.8	FRAPAP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
pinata Ehrenb.	15.0	FRAPIN	0.6	0.4	1.4	2.5	0.2	1.1	2.0	2.7	1.3	2.9	1.3	