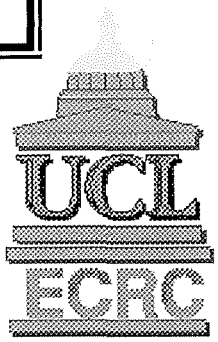
ISSN 1366-7300



ENVIRONMENTAL CHANGE RESEARCH CENTRE

University College London

RESEARCH REPORT

No. 38

Assessing the impact of diatom dissolution in biasing quantitative salinity reconstructions from saline lake sediments

D.B. Ryves & R.W. Battarbee

Final Report for the NERC Research Grant: GR9/02033

1997

Environmental Change Research Centre
University College London
26 Bedford Way
London
WC1H 0AP

Final report for the NERC Research Grant GR9/02033

Assessing the impact of diatom dissolution in biasing quantitative salinity reconstructions from saline lake sediments

Ryves, D.B. & Battarbee, R.W.

Address for correspondence: Environmental Change Research Centre, University College London, 26 Bedford Way, London WC1H 0AP

1. Project objectives and introduction

Current concern with climatic change has highlighted the need for accurate reconstructions of past climate to provide a context for natural variation and as an independent means of testing computer model hindcasts. In semi-arid and arid regions, salt-lake sediments are often excellent archives of past environmental change from both biotic and abiotic proxies. Diatoms in particular have been shown to be highly sensitive indicators of lake water salinity, and so effective moisture (Gasse *et al.* 1987, Fritz *et al.* 1991, Wilson *et al.* 1994, Gasse *et al.* 1995) but reconstructions of such parameters may be compromised by taphonomic factors, particularly diatom dissolution (Barker 1992).

Research within the NERC-funded CASPIA project (Juggins *et al.* 1994) has focused on taxonomic harmonisation and validation of inter-regional diatom datasets to improve and extend training sets and transfer functions derived from them (Carvalho *et al.* in prep). An important adjunct to this is the assessment of diatom dissolution in biasing such models. This project develops earlier work (NERC-funded PhD to D. Ryves) on diatom dissolution in the Northern Great Plains of America (NGP) which weighted samples and species in the NGP salinity transfer function according to experimentally-derived ranking of species to dissolution (Ryves 1994). Specifically, the objectives of this NERC small grant were:

(i) to recount the training set diatom slides to incorporate dissolution
(ii) to calculate dissolution indices and apply these to the transfer function
(iii) to test the method by comparing conventional and dissolution-adjusted transfer functions at a site with an historical record of lake salinity (Devils Lake, Ramsey County, North Dakota).

2. The Northern Great Plains datasets

Various combinations of the 64 lakes originally published by Fritz (1990) have been used in the construction of salinity transfer functions. The first salinity calibration dataset included 55 of the original 64 lakes, sediments from the other lakes containing very few or no diatoms. This dataset, referred to here as NGP55, was used to create a transfer function (NGP55: $\mathbf{r}^2$=0.845, $RMSE_{class}$=0.228) to reconstruct the recent history of salinity at Devils Lake, for which records exist of measured salinity for much of the present century. While diatom-inferred midsummer salinity was found to agree well with observed values at the lower salinity range (<8 g/l) since about 1970, and could distinguish higher salinity episodes (above 10 g/l) from fresher periods, it appeared to underestimate true salinity at the highest point and overestimate it at the beginning of this century (Fritz 1990). The current study seeks to assess the role that diatom dissolution has in biasing this model by incorporating dissolution both directly and indirectly into the transfer function.

Further work refined the model and extended the reconstruction to a Holocene record also from Devils Lake, North Dakota (Fritz *et al.* 1991). This model differed from that published earlier (Fritz 1990) by employing an inverse deshrinking regression during calibration. This method deshrinks less than classical techniques (Birks *et al.* 1990) and tends to reduce overall model RMSE ($RMSE_{inv}$=0.209), although values at the ends of the salinity gradient may lose accuracy compared to classical methods. Subsequently, Fritz *et al.* (1993) produced a 53 lake dataset (NGP53a) from which species salinity optima were calculated, using the weighted averaging (WA) procedure. Two lakes were excluded from the previous models (East Stump and Alkaline) for which diatom counts were less than 100 valves. A salinity transfer function (NGP53a; $\mathbf{r}^2$=0.836, $RMSE_{inv}$=0.227) derived from this dataset has recently been applied to cores from Moon Lake,

North Dakota to infer drought frequency and intensity over the last 11,000 years (Laird *et al.* 1996a & 1996b).

Unfortunately, it was discovered that in NGP53a (but not NGP55) the water chemistry of Bitter Lake, Saskatchewan (salinity 268.04 g/l), had been inadvertently linked to the diatom count for Bitter Lake, South Dakota (salinity 23.53 g/l). After this was corrected, this dataset was chosen as the test set of lakes for this project as it represents the best and current state of the salinity transfer function for the Northern Great Plains. It is hereafter referred to as the NGP53 dataset.

3. Progress

The original slides counted by R.W.Battarbee (28) and S.C.Fritz (25) were recounted to a minimum of 300 valves each, except in those cases of poor diatom preservation where a minimum of 100 valves were counted. These counts were made using the concept of dissolution stages developed previously (Ryves 1994), where each valve encountered is assigned to a dissolution category ("stage"). Stage 1 represents the pristine condition, with higher stages reflecting progressively greater dissolution. Up to 4 stages were recognised for some taxa (such as *Cyclotella* spp., *Amphora libyca* and *Navicula oblonga*, compared to 2 stages for many fine *Nitzschia* spp). These stages are those identified by LM and SEM micrography during previous dissolution experiments of NGP material (Ryves 1994) and represent, as far as possible, a sequence of objectively identifiable valve morphologies followed by species as they dissolve. It is a means of breaking the continuum of the dissolution process into definable and recurrent stages through empirical experiment to produce a dissolution taxonomy. These data were used both indirectly and directly to adjust the NGP53 salinity transfer function, as outlined below.

The uppermost 19 samples from the Devils Lake short core taken by Fritz in 1985 were also recounted using stage counting. The surface sample for this core is also that used in the training set. Between 150-300 valves were recounted from these samples and formed the basis of a stage-counted test set for the salinity transfer function.

One visit was made to Dr. S.C.Fritz at Lehigh University, Pennsylvania to discuss taxonomy and other issues concerning the NGP dataset in February 1997. The implications of salinity reconstructions at Devils Lake under various different transfer functions involving dissolution data were also discussed and compared to the original model. Discussions have also been held with Dr. Steve Juggins (University of Newcastle, UK) about the statistical treatment of the data. All analyses were carried out using CANOCO 3.12 (ter Braak 1988, 1990) and CALIBRATE 0.7 (Juggins & ter Braak 1997). The technique of weighted averaging partial least squares (WA-PLS; ter Braak & Juggins 1993) was also applied to both conventional and dissolution models but always generated larger errors of prediction (higher RMSE) after jackknifing than WA methods and is therefore not discussed further here (cf. Wilson *et al.* 1996). All RMSE values quoted are in $\log_{10}$ salinity units.

4. Results

4.1 Diatom dissolution indices (DDIs)

A first step in quantifying diatom dissolution can be made by calculating indices for individual species and assemblages based on the proportions of valves for a species (or assemblage) for each stage. Three indices were applied to the recounted data, and calculated for each sample in NGP53: Flower's index (**F**), a weighted index (**W**) and a square weighted index ($\mathbf{W^2}$). Flower's index is developed from ideas presented in Flower & Likhoshway (1993) and calculates the proportion of

stage 1 (pristine) valves to all valves in an assemblage. **F** varies between 0 (no pristine valves) to 1 (perfectly preserved sample) and allows direct comparison of any samples regardless of species composition. It has the disadvantage that more subtle patterns of dissolution may be overlooked and does not consider differences in numbers in stages 2-4 between samples. In extreme cases, with no stage 1 valves, samples are equally treated as badly preserved (**F**=0). For this reason, two weighted indices were developed (Ryves 1994) to take into account higher dissolution profiles, as follows:

$$\textit{Weighted index: } W_i = \frac{\sum_{s=1}^{s=4} N_{is} \cdot S}{\sum_{s=1}^{s=4} N_{is}} \qquad \textit{Square Weighted index: } W^2_i = \frac{\sum_{s=1}^{s=4} N_{is} s^2}{\sum_{s=1}^{s=4} N_{is}}$$

where

$\mathbf{N_{is}}$ is the number of valves in sample **i** in stage **s**
s is the stage number (an integer from 1 to 4)

The square weighted index ($\mathbf{W^2}$), compared to the weighted index (**W**), emphasises the non-linearity between the degree of sample dissolution and the proportion of diatoms in the highest (end point) dissolution stages. **W** and $\mathbf{W^2}$ vary between 1 (perfect preservation) to $\mathbf{s_{max}}$ and $\mathbf{s^2_{max}}$ respectively and as such are only directly comparable between samples where the initial proportions of stage 2, 3 and 4 taxa are the same. One or both of the weighted indices can be used in tandem with **F** to cover all dissolution situations, as information lost using one can be captured with another.

4.2 Diatom dissolution within the Northern Great Plains (NGP53)

Two Flower's indices (not shown here) were calculated for samples including *Chaetoceros* cysts, as this is the single taxon which during experimental dissolution did not exhibit distinguishable dissolution stages. Treating *Chaetoceros* cysts as stage 1 valves tends to overestimate the preservation status of these samples, as experiments and observation have shown these are very resistant morphotypes, as distinct from *Chaetoceros* valves which were only found in fresh material and have not been recorded in these surface sediments. For this reason, *Chaetoceros* cysts are excluded from the diatom sum in calculations of Flower's **F** (denoted $\mathbf{F_{adj}}$) unless otherwise specified by $\mathbf{F_{cyst}}$, and for both weighted indices, **W** and $\mathbf{W^2}$, in all further figures and analyses presented here.

The distributions of weighted diatom dissolution indices **W** and $\mathbf{W^2}$ across the 53 lake dataset are shown in figure 1, plotted against sites ordered according to salinity. Dissolution state varies widely amongst NGP samples, from Albert Lake (salinity 0.71 g/l, $\mathbf{F_{adj}}$=0.86, **W**=1.2, $\mathbf{W^2}$=1.75) to Bitter Lake (salinity 23.53 g/l, $\mathbf{F_{adj}}$=0.01, **W**=3.5, $\mathbf{W^2}$=11.9). There is a tendency for dissolution to increase with increasing salinity but this is by no means universal. There are several saline sites with very good preservation, although it is notable that several of these are meromictic (labelled **M**). Conversely there are very few freshwater sites (salinity <3 g/l) with poor diatom preservation. Although the highest species diversity within samples (as measured by Hill's N_2; ter Braak 1988) are found in freshwater lakes, there is a trend for diversity to decrease as sample DDI increases to an extent independent of salinity. This may in part explain the observation that diatom species diversity in sediments decreases with increasing salinity (Fritz *et al.* 1993, Wilson *et al.* 1996).

There are positive relationships between sample DDIs and salinity, except at the highest salinities (linear correlation **r**=+0.36-0.37, p=0.005), but there is little correlation between pH or depth (except the very shallowest sites, which have an inverse relationship to DDI; neither variable is significantly correlated at p=0.01). The factors which show the greatest correlation with DDIs of 18 measured parameters (Fritz 1990, Fritz *et al.* 1993) are Ca^{2+} as percentage of total cations (%Ca^{2+}; **r**=+0.38-0.46, p=0.005) and CO_3^{2-}-HCO_3^- concentration (log-transformed mg/l; **r**=+0.41-0.44, p=0.005), although scatterplots show considerable variation along the chemical gradient. Other variables highly collinear with these are also correlated with DDIs but to a lesser extent (e.g. conductivity, major cation concentrations), but no other anions.

Finally, separate redundancy analyses (RDAs) were carried out on the transformed DDIs individually, to determine if diatom dissolution could be predicted from measured physico-chemical variables. The DDIs (response variables) were constrained by the 18 measured environmental parameters, with an additional dummy variable (0/1) identifying meromictic lakes, including all variables and using forward selection (with Bonferroni adjustment of significance) to provide a more parsimonious model. Results of these are shown in table 1.

Only one or two variables were actually found to offer significant non-collinear explanation of DDI, and only 16-31% of variance in DDI could be explained by lake parameters measured. There is much confusion and often contrary anecdotal evidence in the literature about the causes of diatom dissolution in lakes, but it is interesting that in the NGP, %Ca^{2+} is the dominant factor, with meromixis often playing an important secondary role (cf. Meriläinen 1971, 1973). Other factors may become important in different systems; for example, although pH does not appear to be an important variable in the NGP, this may well reflect the high pH (>8.3) for all these lakes.

4.3 Applying diatom dissolution data to transfer functions

4.3.1 Sample weighting methods

Diatom dissolution indices can be used in two ways to adjust the original NGP53 salinity transfer function, by weighting samples within the training set. Firstly, sample diatom dissolution indices can be used as a means of screening the dataset and removing samples below a certain threshold, effectively imposing a weight of 0. A 49 lake dataset (NGP49) was created by deleting 4 samples with $\mathbf{F_{adj}}$ below 0.2 (Bitter, George, Rabbit and Mission), all saline and three above 20 g/l. A more sophisticated strategy is to downweight each sample in some proportion to sample DDI, either directly or by dividing the dataset into several classes of sample preservation, to reduce the variation in weights ($NGP53_{ddi}$ using $\mathbf{F_{adj}}$). The underlying logic in both cases involves the assumption that poorly preserved samples are unreliable repositories of environmental information, and specifically that species optima derived from them are unsound. Both approaches were used to create two different models and the resultant transfer function applied to the Devils Lake core, using a classical WA for comparison with the original model.

Model diagnostics suggest that the weighted transfer functions will perform better, with larger $\mathbf{r^2}$ (0.873 and 0.880 for NGP49 and $NGP53_{ddi}$ respectively) and lower RMSE (0.195 and 0.204) compared to NGP53. At Devils Lake, however, this is not the case, as the new models behave similarly at lower salinity, but underestimate the higher salinity excursion of the 1920s-1940s to an even greater degree. This disappointing performance may in part be explained by the diatom flora of the Devils Lake short core, which is dominated in pre-1970 samples by *Cyclotella quillensis*, but is also an artifact of the method. Downweighting samples according to dissolution index will tend to dampen the response of the model at higher salinities as there is a bias towards poor dissolution

at higher salinities within the dataset. If species in these samples are not represented in well-preserved saline samples, their estimated WA optima will be underestimated relative to the unadjusted model. This is especially so for *C. quillensis* which has its greatest abundances in generally badly preserved samples. It should be noted that the unadjusted ("original") NGP53 model reconstructs lower values during the high salinity events for the Devils Lake short core than that published in Fritz (1990) because the deleted East Stump (salinity 82.25 g/l) had a high proportion of *C. quillensis*, and so a higher estimated optimum for this species.

4.3.2 Creating dissolution training sets

Diatom dissolution data can also be used more directly to create a new training set for environmental calibration. Each species dissolution stage can be considered as a taxon (with associated optima and tolerances) in its own right. Such "dissolution models" decrease the salinity bias of downweighting according to assemblage DDI, by incorporating the effect that salinity has on dissolution, for those species which exhibit progressively more dissolution as salinity increases. It is also a more sensitive approach; whereas sample DDIs represent the average dissolution state of the constituent taxa, this considers each species occurrence in the dataset individually.

As counts of the original NGP slides were larger than was possible in this study (generally from 400-600 valves), and to allow direct comparison with the original dataset, dissolution counts were used to find the proportion of each species in each dissolution stage for each sample. These values were then multiplied by the original percentage of each species so that, as far as possible, the original relative species composition was conserved within each sample. This procedure inevitably produces a dataset with many more taxa. The full dissolution NGP53 dataset contains 537 taxa ($NGP53_{537}$), 494 above 0.2% ($NGP53_{494}$) and 207 above 2% in any sample ($NGP53_{207}$), compared to 149 over 1% in NGP53 (Fritz *et al.* 1993). Several different salinity transfer functions were developed using the full dataset, and minimum species abundance cutoffs of 0.2% and 2%, with both classical and inverse regression deshrinking, validated using the jackknife technique. Results of these trials are shown in table 2.

In terms of $\mathbf{r}^2$ and RMSE, all models perform better under inverse regression in all situations. It also appears that dissolution models outperform the respective controls. The "best" transfer function uses weighted averaging with species downweighted according to salinity tolerance (WA_{tol}) for the 537 taxa dissolution dataset under inverse regression ($NGP53_{537}$: $\mathbf{r}^2$=0.932, RMSE=0.1352; control NGP53: $\mathbf{r}^2$=0.906, RMSE=0.156). Jackknifing, however, demonstrates that this cannot be expected outside the dataset (table 2). Under validation the control models fare better than their dissolution counterparts, although the differences are small. Models also tend to improve as the number of taxa increases, in agreement with other studies (Birks 1994, Wilson *et al.* 1996).

All the models show a similar trend in observed against predicted salinity and in the residual structure. Figure 2 shows typical model results (from the $NGP53_{537}$ dissolution dataset under classical regression) with a LOWESS (locally weighted regression smoothing) curve plotted, emphasising bias as value-dependant deviation. Salinity tends to be slightly underestimated at low levels (<3 g/l), overestimated in the mid-range (3-12 g/l) and increasingly underestimated above about 20 g/l (though these residuals are smaller with dissolution datasets). The salinity gradient has shrunk as high salinity lakes have been removed from the datasets, from 0.65-82.25 g/l (NGP55), 0.65-268.04 g/l (NGP53a) to 0.65-41.41 g/l (NGP53), which imposes a limit on inferred salinities. Within this range, the NGP salinity transfer function compares favourably with a 219 lake training

set from British Columbia (Wilson *et al.* 1996; salinity range 0.02-620.29 g/l, no. taxa 204, r^2=0.87, $RMSE_{boot}$=0.371).

If salinity is important in diatom dissolution, salinity optima calculated for successive dissolution stages should increase within an individual species. Figure 3 shows the WA salinity optima for all dissolution stages for the 39 species in NGP53 with a maximum abundance of at least 10% and which occur in 6 lakes or more. For many species, less dissolved valves (lower stage numbers) have lower salinity optima than more dissolved valves, often in numeric order (e.g. *Cyclotella quillensis*, species 7 in figure 3, for which the species optimum is largely defined by badly dissolved samples). Dissolution stage salinity optima can span the species optimum by 0.5 log-salinity units or more. Many of these dissolution taxa exhibit a unimodal response to salinity as does the species from which they derive, an assumption of WA methods.

During the dissolution recounting, raphid and araphid valves were also separated (*Achnanthes* spp., *Cocconeis* spp. and *Rhoicosphenia curvata*). These are given the same species optimum but plotted as raphid (**R**) and araphid (**A**) forms. Assuming equal production of both valve types, differences in salinity optima can only reflect taphonomic processes biasing the proportions in samples. Dissolution experiments have demonstrated that araphid valves are more resistant to dissolution than their raphid counterparts (Ryves 1994), but figure 3 suggests only minor difference in salinity range of raphid and araphid valves. It is important to show that differential dissolution should not affect models with counting strategies that do not separate these valve types.

4.4 Testing the dissolution models: Devils Lake short core

As a test of dissolution model performance, the four best dissolution salinity transfer functions (537 taxa) were applied to fossil samples from the Devils Lake short core, and compared with both the measured salinity and the original NGP53 model (without stage counting). These reconstructions are shown in figure 4, covering the period from about 1900 to 1985. Fossil sample counts were treated in a similar way to the surface sediments to preserve species proportions originally based on counts of between 350 and 600 valves (Dr. S.C.Fritz pers.comm.).

All perform similarly well at the lower salinity values (<8 g/l) but all underestimate the higher salinity excursions of 1960-1970 and the major 1920-1948 "dust bowl" drought period. There is also a suggestion that values are overestimated during the 1950s and again at the start of this century. There is however less error than with the earlier weighting methods and the classical models (particularly WA) predict higher values than NGP53 during the period of greatest salinity, which represents a modest but real improvement. There is no systematic bias in dissolution transfer functions to infer either always higher or lower values than the original model, as they are responding to more subtle differences in sample dissolution profiles. Intriguingly, simple WA appears a better model at this site than WA_{tol}, despite the better performance of WA_{tol} under jackknifing (table 2).

Fossil sample dissolution indices can be used in conjunction with the inferred values as a diagnostic measure of their reliability (figure 5). Trends in sample DDI tend to parallel salinity; samples with the greatest discrepancy between measured and inferred salinity are the most dissolved, but values are not outside the range found within the surface sediments. DDIs provide another means of flagging samples whose inferred values may be suspect, in addition to analogue matching (Birks *et al.* 1990) and goodness-of-fit measures (Kingston *et al.* 1992, Laird *et al.* 1996a). For instance, no fossil sample had squared residual distances from a CCA salinity axis greater than the 90% limit of the distribution of such distances in the surface sediments (log transformed for normality). Samples

in the 90-95% range are considered to have a "poor fit" to the reconstructed variable and those above 95% a "very poor fit" (Birks *et al.* 1990, Laird *et al.* 1996a), although by this method none would have been flagged.

Although individual species abundances in Devils Lake short core samples are not unusual within the NGP53 dataset, the coincidence of the freshwater taxa *Stephanodiscus niagarae* (optimum 1.24 g/l) and *S. minutulus* (2.06 g/l) with the saline taxon *C. quillensis* (12.53 g/l) for many of the samples is. This "no-analogue" problem may be a major factor affecting salinity reconstruction at this site, as well as sample dissolution and problems of taphonomy and sediment reworking linked to substantial lake level fluctuation. Considering the DDI for these species across their NGP distributions, it can be shown that the lower limit for *C. quillensis* and the upper limits for the *Stephanodiscus* species are not related to poor preservation and support is given to the argument that these taxa do not co-occur in the same salinity range. As an experiment, these freshwater taxa were deleted from all fossil samples with measured salinity above 10 g/l and salinity recalculated using the classical WA and WA_{tol} $NGP53_{537}$ model. As expected, inferred salinity values increase but show a very flat response to measured fluctuations compared to the control model. For this site this is an unsatisfactory approach to the problem.

5. Conclusions: future work

Diatom dissolution indices allow diatom preservation to be quantified and samples to be compared objectively rather than on an *ad hoc* basis. Sample DDIs can be used diagnostically (as a data screening process, and to flag samples) and can be used to adjust inferential models by supplying an impartial means of weighting samples. Changes in dissolution profiles in marine sediments have been correlated with climatically-driven variations in dissolution rate (Farrell & Prell 1989) and could provide similar information independent of species assemblage change in palaeolimnological studies.

Transfer functions created using "dissolution taxa" appear to perform better than orthodox models but are in fact poorer under validation. All suffer from the bias in reconstructed salinity at the highest salinity values. This study has also shown that there is a paucity of higher salinity sites with well preserved diatom assemblages. Targeting such gaps (perhaps meromictic lakes) in the dataset can be expected to improve the transfer function at higher salinities as species parameters will be improved.

Diatom dissolution is important in the NGP and there can be modest improvements in model performance if dissolution is incorporated into transfer functions. Factors unrelated to diatom dissolution may be responsible for the poor fit at high salinity within the Devils Lake short core but a different strategy may be more successful. Previous research has suggested that many species exhibit a predictable relationship between valve loss and species DDI (Ryves 1994), thus providing a procedure for adjusting taxon proportions within a dissolved sample prior to model building. Although outside the scope of this project, this represents another approach to dissolution problems.

Errors in reconstructed salinity are likely for high values (underestimation) but additionally sites that fluctuate across ecological thresholds (especially the freshwater-saline) may contain incongruous assemblages. Species DDIs can be used to validate lower or upper species distributions, and comparison of species DDIs within a sample may help to indicate valves of a different taphonomic provenance (e.g. reworked taxa) in a sample. This can provide a sounder basis for dealing with such no-analogue cases.

These ideas may have applications in other situations where the luxury of selecting only well-preserved samples does not exist. The quantitative treatment of microfossil preservation may extend the range over which valid palaeoenvironmental reconstruction can be made.

6. Acknowledgements

We would like to thank Sheri Fritz and Steve Juggins for their substantial help during this project providing original NGP slides, hospitality and useful discussion of the issues involved with the project. Thanks are also due to Laurence Carvalho, Martin Kernan, Anson Mackay and Roger Flower for support and advice.

7. References

Publications in prep arising from the project

Ryves, D.B. & Battarbee, R.W. "Assessing the impact of diatom dissolution in biasing quantitative salinity reconstructions from saline lake sediments", ECRC Research Report no.38, University College London.

Ryves, D.B., Battarbee, R.W. & Juggins, S. "Experimental quantification of diatom dissolution in lake sediments".

Ryves, D.B., Battarbee, R.W., Fritz, S.C. & Juggins, S. "Applications of diatom dissolution data to palaeoenvironmental reconstruction".

Other references

Barker, P. (1992) 'Differential diatom dissolution in Late Quaternary sediments from Lake Manyara, Tanzania: An experimental approach', *J. Paleolim.*, **7**, 235-251.

Birks, H.J.B. (1994) 'The importance of pollen and diatom taxonomic precision in quantitative palaeoenvironmental reconstructions', *Rev. Palaeobot. Palynol.*, **83**, 107-117.

Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C. & ter Braak, C.J.F. (1990) 'Diatoms and pH reconstruction', *Phil. Trans. R. Soc. Lond.*, **B 327**, 263-278.

Carvalho, L.R., Cox, E.J., Ryves, D., Juggins, S., Fritz, S.C., Gasse, F., Sims, P.A., Mitlehener, A. & Battarbee, R.W. 'Diatom flora of African and North American inland saline waters', *in prep.*

Farrell, J.W. & Prell, W.L. (1989) 'Climatic change and $CaCO_3$ preservation: An 800,000 year bathymetric reconstruction from the central equatorial Pacific Ocean', *Paleoceanography*, **4**, 447-466.

Flower, R.J. & Likhoshway, Y. (1993) 'An investigation of diatom preservation in Lake Baikal', *Fifth workshop on diatom algae*, March 16-20, 1993, Irkutsk, Russia, 77-78.

Fritz, S.C. (1990) 'Twentieth-century salinity and water-level fluctuations in Devils Lake, North Dakota: A test of a diatom-based transfer function', *Limnol. Oceanogr.*, **35**, 1771-1781.

Fritz, S.C., Juggins, S., Battarbee, R.W. & Engstrom, D.R. (1991) 'Reconstruction of past changes in salinity and climate using a diatom-based transfer function', *Nature*, **352**, 706-708.

Fritz, S.C., Juggins, S. & Battarbee, R.W. (1993) 'Diatom assemblages and ionic characterization of lakes of the northern Great Plains, North America: A tool for reconstructing past salinity and climate fluctuations', *Can. J. Fish. Aquat. Sci.*, **50**, 1844-1855.

Gasse, F., Juggins, S. & Ben Khelifa, L. (1995) 'Diatom-based transfer functions for inferring past hydrochemical characteristics of African lakes', *Palaeogeog., Palaeoclim., Palaeoecol.*, **117**, 31-254.

Gasse, F., Fontes, J.-Ch., Plaziat, J.C., Carbonel, P., Kaczmarska, I., De Deckker, P., Soulie-Marsche, I., Callot, Y. & Dupeuble, P.A. (1987) 'Biological remains, geochemistry and stable isotopes for the reconstruction of environmental and hydrological changes in the holocene lakes from northern Sahara', *Palaeogeog., Palaeoclim., Palaeoecol.*, **60**, 1-46.

Juggins, S., Battarbee, R.W., Fritz, S.C. & Gasse, F. (1994) 'The CASPIA project: diatoms, salt lakes, and environmental change', *J. Paleolim.*, **12**, 191-196.

Juggins, S. & ter Braak, C.J.F. (1997) *CALIBRATE v 0.7 (Beta Test) - a computer program for species-environment calibration by [weighted-averaging] partial least squares regression*, ECRC, University College London.

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', *Can. J. Fish. Aquat. Sci.*, 49, 116-127.

Laird, K.R., Fritz, S.C., Grimm, E.C. & Mueller, P.G. (1996a) 'Century-scale paleoclimatic reconstruction from Moon Lake, a closed-basin lake in the northern Great Plains', *Limnol. Oceanogr.*, **41**, 890-902.
Laird, K.R., Fritz, S.C., Maasch, K.A. & Cumming, B.F. (1996b) 'Greater drought intensity and frequency before A.D. 1200 in the Northern Great Plains, USA', *Nature*, **384**, 552-554.
Meriläinen, J. (1971) 'The recent sedimentation of diatom frustules in four meromictic lakes', *Ann. Bot. Fennici S.*, 160-176.
Meriläinen, J. (1973) 'The dissolution of diatom frustules and its paleoecological interpretation', *Symposium, Univ. Lund*, 91-95.
Ryves, D.B. (1994) 'Diatom dissolution in saline lake sediments: an experimental study in the Great Plains of North America', Unpublished Ph.D. thesis, University College London.
ter Braak, C.J.F. (1988) *'Canoco - A Fortran Program for Canonical Community Ordination by (Partial)(Detrended)(Canonical) Correspondence Analysis, Principal Components Analysis (Version 2.1)'*, Agricultural Mathematics Group, Wageningen.
ter Braak, C.J.F. (1990) *'Update Notes: Canoco Version 3.10'*, Agricultural Mathematics Group, Wageningen.
ter Braak, C.J.F. & Juggins, S. (1993) 'Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages', *Hydrobiologia*, **269/270**, 485-502.
Wilson, S.E., Cumming, B.F. & Smol, J.P. (1996) 'Assessing the reliability of salinity inference models from diatom assemblages: an examination of a 219-lake data set from western North America', *Can. J. Fish. Aquat. Sci.*, **53**, 1580-1594.

Notes

general problem
objectives statement
(short?)
WA-PLs tried and failed (Wilson et al 1996)
N2 vs salinity vs DDIs
use of canoco, cali

correlations of ddi's vs env vars (esp salinity) - fig 1
graphs of ddi's vs salinity - fig 2 3
RDAs of ddi (response) to env vars

ie ddi d/w
remove 4 lakes fig 4

ie building a new ngp dataset
table 1 - output of models
exp vs obs - fig 5
residuals - fig 6
optima vs stages fig 7
plot of cyclotella q & stages fig 8a 8b

application to devils - fig 9 4 models
compare fossil ddi's vs salinity (& model performance) fig 10
removal of stephs fig 11

Dissolution can now be evaluated - sites compared;
Some improvement - need to test hypothesis that dissln is important in affecting recon.
can put confidence limit (flag) samples with poor pres = bad recon.
can suggest residuals where likely errors (false recons0 likely to occur - fresh/saline dissolved assemblage

Table 1 - RDA model results (DDI)

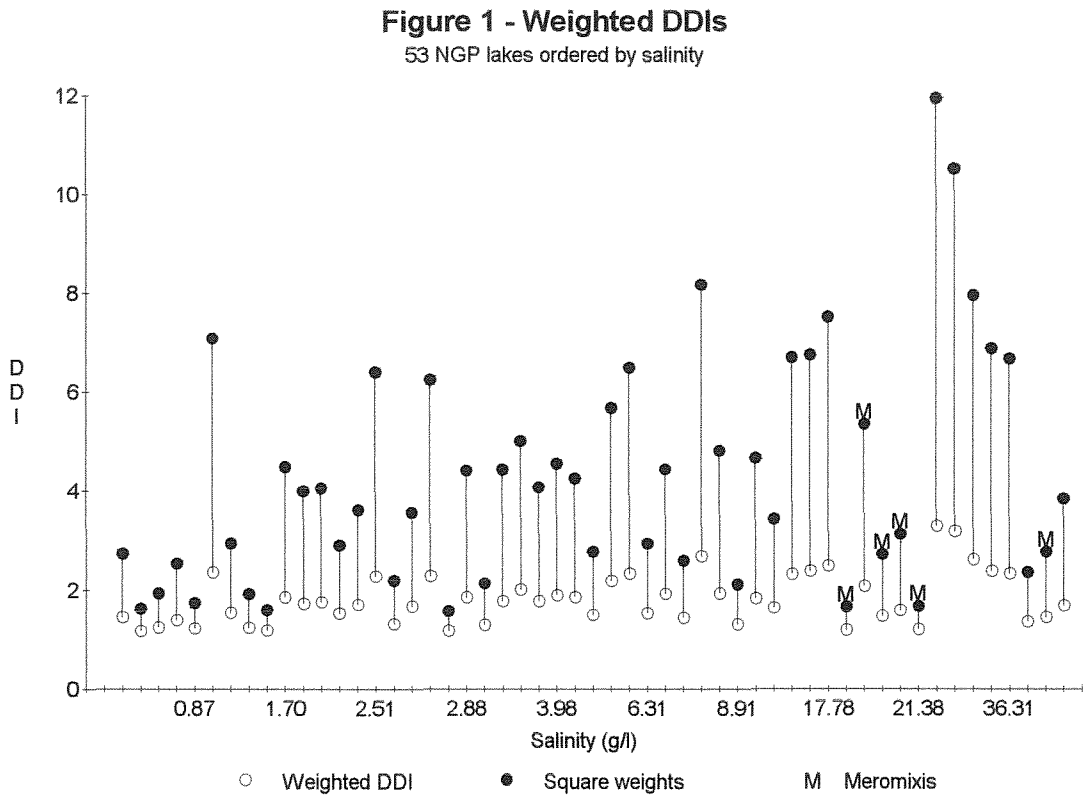
DDI[1]	All variable model (%)	Forward selection	Added fit	P^2	Variance explained (%)	P^2
$(1\text{-}F_{(cyst)})^{½}$	47.1	CO_3^{2-}+HCO_3^-	0.16	0.01	16.2	0.01
$(1\text{-}F_{(adj)})^{½}$	50.5	$\%Ca^{2+}$ meromixis	0.21 0.09	0.01 0.01	30.2	0.01
log W	53.4	$\%Ca^{2+}$ meromixis	0.20 0.10	0.01 0.02	30.3	0.01
log W^2	52.2	$\%Ca^{2+}$ meromixis	0.20 0.11	0.01 0.01	31.0	0.01

Table 2 - Salinity transfer function results

Type	Dataset	No. taxa	Cutoff %	WA		WA_{tol}		Jack WA		Jack WA_{tol}	
				r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
Classical	$NGP53_{537}$	537	---	0.8815	0.1894	0.9315	0.1401	0.7193	0.2767	0.7552	0.2619
	$NGP53_{494}$	494	0.2	0.8795	0.1913	0.9307	0.1410	0.7140	0.2799	0.7532	0.2633
	$NGP53_{207}$	207	2	0.8704	0.1994	0.9233	0.1489	0.7142	0.2820	0.7661	0.2562
	NGP53	**149**	**1**	**0.8681**	**0.2014**	**0.9055**	**0.1669**	**0.7300**	**0.2745**	**0.7858**	**0.2453**
Inverse	$NGP53_{537}$	537	---	0.8815	0.1778	0.9315	0.1352	0.7160	0.2759	0.7541	0.2575
	$NGP53_{494}$	494	0.2	0.8795	0.1794	0.9307	0.1361	0.7104	0.2784	0.7520	0.2587
	$NGP53_{207}$	207	2	0.8704	0.1860	0.9233	0.1431	0.7105	0.2781	0.7649	0.2512
	NGP53	**149**	**1**	**0.8681**	**0.1876**	**0.9055**	**0.1588**	**0.7262**	**0.2705**	**0.7843**	**0.2400**

[1] DDIs were transformed to approximate normality by taking logarithms (log **W** and log $\mathbf{W^2}$) or the square root of $(1\text{-}\mathbf{F_{cyst}})$ or $(1\text{-}\mathbf{F_{adj}})$ as $\mathbf{F_{adj}}$ and $\mathbf{F_{cyst}}$ are negatively skewed.

[2] Monte Carlo permutation, n=99



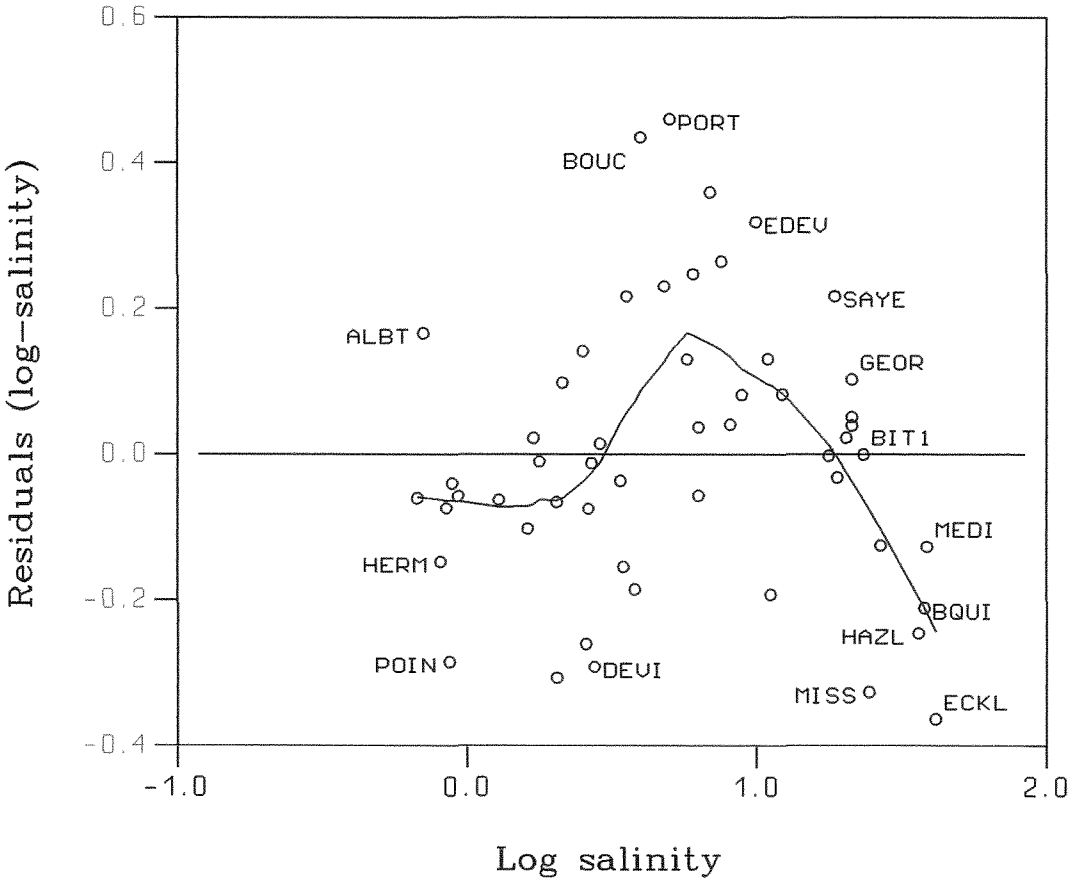


Figure 2 - Model residuals for NGP_{537} (classical WA)

Figure 3 - Salinity optima

Species and dissolution stages

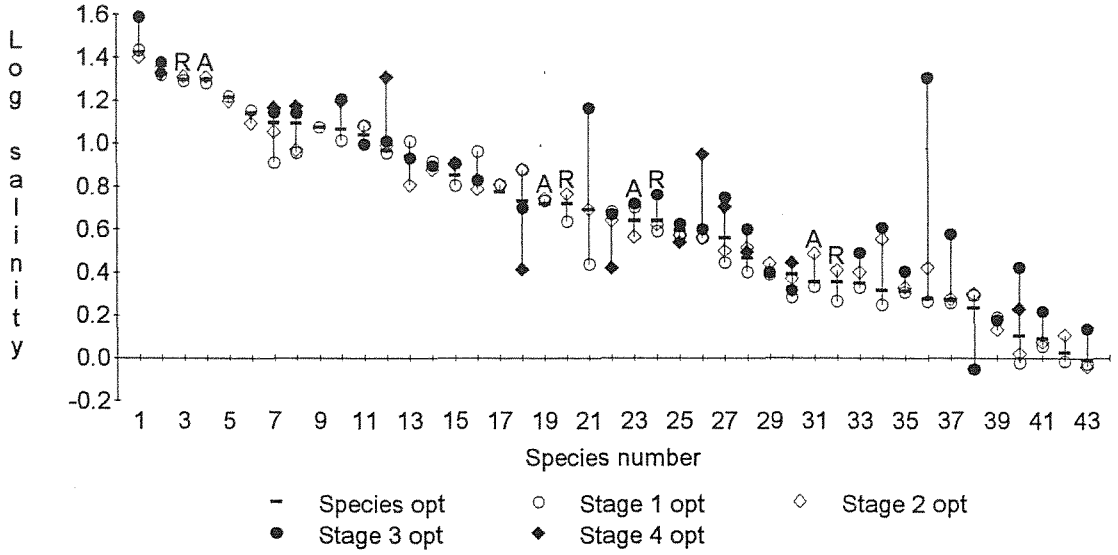


Figure 4 - Inferred & measured salinity

WA models applied to Devils Lake core

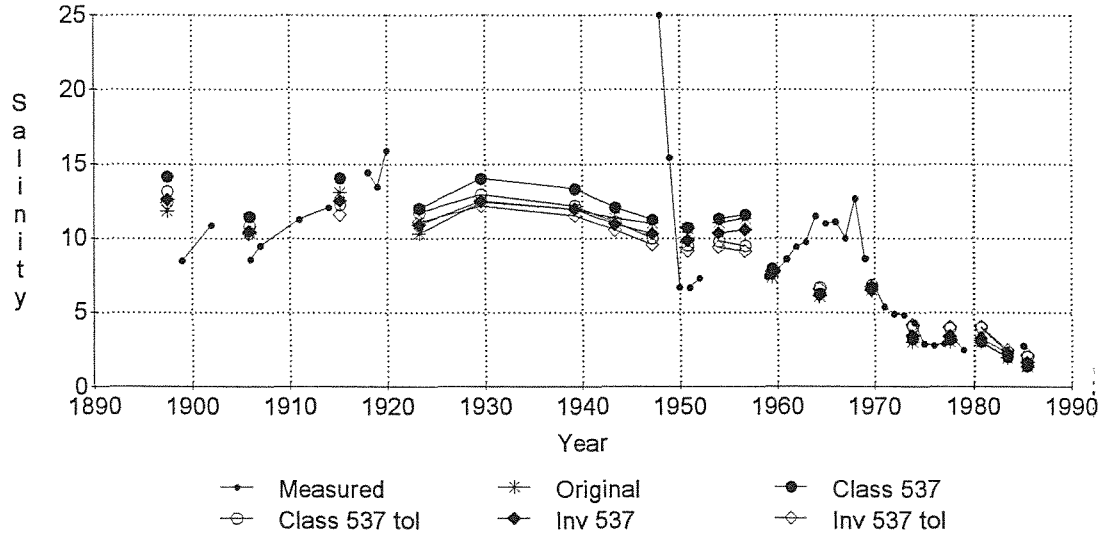


Figure 5 - DDIs and observed salinity

Devils Lake short core