This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

The dilemma of disappearing diatoms: incorporating diatom dissolution data into palaeoenvironmental modelling and reconstruction

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1016/j.quascirev.2008.08.021>

PUBLISHER

© Elsevier

VERSION

AM (Accepted Manuscript)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Ryves, David B., Richard W. Battarbee, and Sherilyn C. Fritz. 2019. "The Dilemma of Disappearing Diatoms: Incorporating Diatom Dissolution Data into Palaeoenvironmental Modelling and Reconstruction". figshare. <https://hdl.handle.net/2134/14504>.

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Manuscript Number: JQSR-D-08-00024

Title: The dilemma of disappearing diatoms: Incorporating diatom dissolution data into palaeoenvironmental modelling

Article Type: Research and Review Paper

Corresponding Author: Dr David Ryves,

Corresponding Author's Institution: Loughborough University

First Author: David B Ryves

Order of Authors: David B Ryves; Richard W Battarbee; Sherilyn C Fritz

Abstract: The problem of microfossil preservation, specifically diatom dissolution, remains an important, but often overlooked, source of error in both qualitative and quantitative reconstructions of key variables from fossil samples, especially those using relative abundance data. A first step to tackling this complex issue is establishing an objective method of assessing preservation (here, diatom dissolution) that can be applied by different analysts and incorporated into routine counting strategies. Here, we establish a methodology for assessment of diatom dissolution under standard light microscopy (LM) illustrated with morphological criteria for a range of major diatom valve shapes. Dissolution data can be applied to numerical models (transfer functions) from contemporary samples, and to fossil material to aid interpretation of stratigraphic profiles and taphonomic pathways of individual taxa. Using a surface sediment diatom-salinity training set from the Northern Great Plains (NGP) as an example, we explore a variety of approaches to including dissolution data in salinity inference models indirectly and directly. Results show that dissolution data can improve models, with apparent dissolution-adjusted error (RMSE) up to 15% lower than their unadjusted counterparts. Internal validation suggests improvements are more modest, with bootstrapped prediction errors (RMSEP) up to 10% lower. When tested on a short core from Devils Lake, North Dakota, which has a historical record of salinity, dissolution-adjusted models infer higher values compared to unadjusted models during peak salinity of the 1930s-40s Dust Bowl but nonetheless significantly underestimate peak values. Site-specific factors at Devils Lake associated with effects of lake level change on taphonomy (preservation and re-working, implied by dissolution data) may override model improvements incorporating dissolution.

Dissolution-adjusted salinity models are also applied to a 150-yr sediment record from Spiritwood Lake, North Dakota, which suggests that this lake has a damped and lagged response to major regional climate forcing of salinity during the Dust Bowl. At this site, dissolution data also suggest different taphonomic behaviour of taxa related to their seasonal patterns of growth and sedimentation. Thus, dissolution data can improve models, and aid interpretation of sedimentary profiles as records of limnological, ecological and environmental change, filtered by taphonomy.

The Editor
Quaternary Science Reviews

David Ryves
Department of Geography
Loughborough University
Loughborough
Leicestershire LE11 3TU
United Kingdom

January 21st 2008

Dear *Quaternary Science Reviews*,

Please find enclosed a copy of our manuscript (*The dilemma of disappearing diatoms: Incorporating diatom dissolution data into palaeoenvironmental modelling*; D.B Ryves, R.W. Battarbee & S.C Fritz) in consideration for review by *Quaternary Science Reviews*. We feel *QSR* to be the most appropriate journal for this research as the manuscript discusses a novel approach to the widespread, though underinvestigated, problem of microfossil preservation in Quaternary research. This is an important topic which we feel will be of interest to a wide audience of Quaternary scientists, as it affects the accuracy, and utility, of palaeoenvironmental inferences themselves. The paper also continues the tradition that *QSR* has of publication on fundamental aspects of taphonomy in Quaternary research.

While the paper focuses on diatom preservation and diatom-salinity transfer functions, we feel the approach can be generalised to other proxies and inferential techniques (both qualitative and quantitative), especially those that rely on abundance data in fossil and modern (training set) samples. In particular the manuscript discusses methods of incorporating diatom dissolution data into numerical techniques (transfer functions) and stratigraphic profiles, to generate new models for quantitative inference and improve interpretations of sedimentary archives for inferring past environments. We explore the implications of these by application to two short sedimentary profiles, covering the last ~150 years, from Devils Lake and Spiritwood Lake (both in North Dakota, USA).

There are some 20 additional plates (similar to Appendix 1: Figs. A1-A3 in the manuscript) covering a wide range of diatom taxa and morphologies that would perhaps be suitable as part of the online area (Multi Media Components). While I have not included these at this stage, I can send these in if needed for review.

We can suggest 4 referees for review:

Dr. Philip Barker

Department of Geography
Lancaster Environment Centre
Lancaster University
Lancaster LA1 4YQ
UK

Email: p.barker@lancaster.ac.uk

Prof. Brian Cumming

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

Email: cummingb@QueensU.Ca

Prof. John Birks

Botanical Institute
University of Bergen
Allégaten 41
N-5007 Bergen
Norway

Email: john.birks@bio.uib.no

Assoc. Prof. Roland Hall

Department of Biology

University of Waterloo
200 University Avenue West
Waterloo
Ontario N2L 3G1
Canada

Email: rihall@sciborg.uwaterloo.ca

Yours faithfully,

David Ryves

**Department of Geography
Loughborough University
Loughborough
Leicestershire LE11 3TU
United Kingdom**

**Phone: +44 (0)1509 228192
Fax: +44 (0)1509 223930
Email: D.B.Ryves@Lboro.ac.uk**

The dilemma of disappearing diatoms: Incorporating diatom dissolution data into palaeoenvironmental modelling

David B. Ryves^{a*}, Richard W. Battarbee^b and Sherilyn C. Fritz^c

^a Department of Geography, Loughborough University, Loughborough, Leicestershire, LE11 3TU, United Kingdom.

^b Environmental Change Research Centre, University College London, 26 Bedford Way, London, WC1H 0AP, United Kingdom. Email: ucfabat@ucl.ac.uk

^c Department of Geosciences and School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68588—0340, USA. Email: sfritz2@unl.edu

* author to whom correspondence should be addressed.

Email: d.b.ryves@lboro.ac.uk; telephone: +44 (0)1509 228192; fax: +44 (0)1509 223930

Running head: Diatom dissolution and transfer functions

Abstract

The problem of microfossil preservation, specifically diatom dissolution, remains an important, but often overlooked, source of error in both qualitative and quantitative reconstructions of key variables from fossil samples, especially those using relative abundance data. A first step to tackling this complex issue is establishing an objective method of assessing preservation (here, diatom dissolution) that can be applied by different analysts and incorporated into routine counting strategies. Here, we establish a methodology for assessment of diatom dissolution under standard light microscopy (LM) illustrated with morphological criteria for a range of major diatom valve shapes. Dissolution data can be applied to numerical models (transfer functions) from contemporary samples, and to fossil material to aid interpretation of stratigraphic profiles and taphonomic pathways of individual taxa. Using a surface sediment diatom-salinity training set from the Northern Great Plains (NGP) as an example, we explore a variety of approaches to including dissolution data in salinity inference models indirectly and directly. Results show that dissolution data can improve models, with apparent dissolution-adjusted error (RMSE) up to 15% lower than their unadjusted counterparts. Internal validation suggests improvements are more modest, with bootstrapped prediction errors (RMSEP) up to 10% lower. When tested on a short core from Devils Lake, North Dakota, which has a historical record of salinity, dissolution-adjusted models infer higher values compared to unadjusted models during peak salinity of the 1930s-40s Dust Bowl but nonetheless significantly underestimate peak values. Site-specific factors at Devils Lake associated with effects of lake level change on taphonomy (preservation and re-working, implied by dissolution data) may override model improvements incorporating dissolution. Dissolution-adjusted salinity models are also applied to a 150-yr sediment record from Spiritwood Lake, North Dakota, which suggests that this lake has a damped and lagged response to major regional climate forcing of salinity during the Dust Bowl. At this site, dissolution data also suggest different taphonomic behaviour of taxa related to their seasonal patterns of growth and

sedimentation. Thus, dissolution data can improve models, and aid interpretation of sedimentary profiles as records of limnological, ecological and environmental change, filtered by taphonomy.

1. Introduction

Current concern with global environmental change (IPCC, 2007) has highlighted the need for accurate reconstructions of past conditions to describe natural patterns of variation, and as an independent means of testing climate model hindcasts (Battarbee, 2000; Birks and Birks, 2006). Biological proxies preserved in sedimentary archives provide powerful tools for making quantitative reconstructions, often using multivariate modelling approaches. While sophisticated methods exist for such analyses, remarkably little attention has been paid to the quality of the primary data on which such models, and reconstructions, are based (Barker, 1992; Mackay et al., 2005; Ryves et al., 2006). Poor microfossil preservation has long been recognised as affecting all biogenic materials in certain circumstances (e.g. silica, calcite and sporopollenin; Lewin, 1961; Berger, 1968; Birks and Birks, 1980), to the extent that perfect preservation is often the exception rather than the rule. Despite much progress on understanding processes and rates of silica dissolution and biogeochemical cycling in natural systems (Conley et al., 1993; Dixit et al., 2001; Van Cappellen et al., 2002), little work has been carried out either to quantify the effects of preservation on inference models (Barker, 1992; Ryves et al., 2006) or explicitly to consider what approaches can be taken to minimise the impacts on inferences from fossil data.

Diatoms are widely used to reconstruct past aquatic environments in a range of systems (freshwater, brackish and marine; Stoermer and Smol, 1999), and are key in understanding the global biogeochemistry of silica (Tréguer et al., 1995). Diatom dissolution is a widespread problem in many depositional environments, particularly in saline and marine settings, but is observed in many freshwater systems too (Round, 1964; Flower, 1993; Straub, 1993; Ryves et al., 2003). Poor diatom preservation may be one reason why diatom-salinity models may lead to inaccurate or misleading reconstructions when applied to fossil data, especially at higher salinities (Fritz et al., 1993; Reed, 1998b; Ryves et al., 2006). The majority of models for quantitative reconstruction using diatoms use the robust technique of weighted averaging (WA) and its derivatives (e.g. weighted averaging-partial least squares, WA-PLS; ter Braak and Juggins, 1993),

whereby relative abundance (%) data are used to define species' optima, and tolerances. Other methodological approaches, such as analogue matching and maximum likelihood modelling, also rely on relative abundance data. Such approaches are susceptible to any taphonomic processes (e.g. differential transport, sedimentation and dissolution; Juggins, 1992; Cameron, 1995; Ryves et al., 2003) that bias the relative representation of species within, and between, the training set and fossil data. Given that it is often not possible (or desirable, from an analogue perspective; Hutson, 1977) to create models and choose fossil samples that are always well preserved, incorporating preservation data directly or indirectly into modern training sets (transfer functions) and fossil data (reconstructions) offers an explicit route to tackle preservation issues.

Several possible approaches exist to the general problem of poor preservation using dissolution data (Table 1). Indirect methods include weighting species and samples according to relative resistance to dissolution, although situations in which this approach is suitable may be rare and may only result in minor improvements to model performance (Ryves, 1994). A more intuitively appealing, flexible, and potentially powerful, approach makes use of the direct (morphological) assessment of valve preservation in samples, whereby valves are classed into morphologically distinct dissolution categories, or simply classified as pristine or dissolved (Flower and Likhoshway, 1993; Ryves et al., 2001; Ryves et al., 2006). Earlier work has explored qualitative and quantitative use of diatom dissolution data in contemporary and palaeoenvironmental studies, to estimate the reliability of diatom-based inferences from fossil data (Barker, 1992; Ryves et al., 2001; Ryves et al., 2006), but here we incorporate dissolution data directly into numerical models and reconstructions themselves (cf. Battarbee et al., 2005; Mackay et al., 2005; Austin et al., 2007). We also make use of dissolution data for individual taxa to help understand their taphonomic behaviour and aid interpretation of stratigraphic profiles.

Dissolution indices can be used to weight species and samples directly within (or exclude them from) the training set (Table 1). Dissolution counting also can be used to create new transfer functions whereby dissolution is explicitly incorporated into the model. Each species' response

along an environmental gradient can be decomposed into several constituent “dissolution taxa”, each with an associated optimum and tolerance, especially where dissolution is driven partly by the parameter of interest (e.g. salinity; Fritz et al., 1993; Wilson et al., 1996; Reed, 1998a; Ryves et al., 2006). While species’ distributions in living communities should respond to major environmental gradients (Jongman et al., 1995), a dissolution (taphonomic) gradient may cut across these in a fossil or sub-fossil assemblage. The replacement of one species optimum by two (or more) dissolution taxa optima may improve both model performance and quantitative inferences. Approaches that directly assess dissolution data in both training sets and fossil data are especially attractive as both datasets are treated equally (Table 1).

In some instances, valve abundances for individual taxa can be recalculated across a dataset (Battarbee et al., 2005; Mackay et al., 2005) where data on taxon-specific losses through dissolution exist (Ryves et al., 2001; Ryves et al., 2003). These dissolution-corrected species’ abundances can be used to derive new models and reconstructions (both qualitative and quantitative; Mackay et al., 2005). Dissolution-adjusted transfer functions can be compared to conventional models (e.g. Mackay et al., 2003; Mackay et al., 2005) and validated both internally (by the jack-knife or bootstrap methods) and externally by comparison with an observed record of an environmental parameter, where this exists.

Here we explore the issue of a dissolution-corrected transfer function in a well-known and widely applied diatom-salinity model used for palaeoclimate reconstruction from lake sediments in the Northern Great Plains of America (the NGP; Fritz et al., 1991; Laird et al., 1996b; Laird et al., 2003; Laird et al., 2007). This paper develops earlier work on quantifying and predicting diatom dissolution from physical and chemical limnology, and its role as a source of error in palaeosalinity reconstruction at Devils Lake, North Dakota (Ryves et al., 2001; Ryves et al., 2006). This site, which has an observed record of salinity covering much of the 20th century (Fritz, 1990), provides a test case for the approaches to dissolution data discussed here. Dissolution adjusted models and data are also applied to a ²¹⁰Pb-dated short core from Spiritwood Lake, North

Dakota covering the last ~150 years.

2. Materials and methods

2.1 The Northern Great Plains datasets

Although, on an annual basis, evapotranspiration exceeds precipitation across the Northern Great Plains, topographic and groundwater conditions permit many perennial and quasi-perennial lakes to exist. Diatom communities, and surface sediment assemblages, respond strongly to the climatically-controlled salinity gradient found across the wider region (Fritz et al., 1993; Wilson et al., 1996), and have formed the basis for diatom-salinity transfer functions for palaeoclimate inference and their application at several key sites across the NGP (Fritz et al., 1991; Laird et al., 1996b; Laird et al., 2003).

The salinity models explored here are based on models developed by Fritz (1990), Fritz et al. (1991) and Fritz et al. (1993). The 55 site salinity model published by Fritz (1990) and Fritz et al. (1991) includes two sites (Alkaline and East Stump) with poorly preserved diatoms (and counts <100 valves) which were excluded in the 53 site salinity model of Fritz et al. (1993). Both these training sets (hereafter the NGP55 and NGP53 respectively) are examined here incorporating dissolution data. Additionally, an error of incorrect site attribution within earlier models has been corrected.

Many lakes within the NGP53 dataset are endorheic, from 0.1 – 28 m depth (average ~5 m) and range in size from $\sim 10^{-2}$ km² to $\sim 10^2$ km². Salinity ranges from 0.68 to 41.4 g L⁻¹ total dissolved solids (TDS) in NGP53, and to 82.9 g L⁻¹ TDS in NGP55 (mean ~10 g L⁻¹ TDS for both datasets), while pH ranges from 8.25 to 9.6 (median 8.82). Water chemistry is predominantly Na—MgSO₄, with carbonate precipitation at higher salinity. Full details of all NGP lakes may be found in Fritz et al. (1993).

2.2 Devils Lake and Spiritwood Lake

Short cores from two lakes within the NGP also were used to test the dissolution methodology. Devils Lake, North Dakota (48°05' N, 98°56' W) is a large (55 km²) lake that is a terminal basin for a 10 000 km² watershed. Conductivity and salinity have been measured intermittently since 1899 but with a gap between 1923-1948, while annual water level measurements date back to 1901 (Fritz, 1990; Fig. 7). A 1 m piston core was collected in January 1986 from the Creel Bay embayment in 7.5 m water. Maximum water depth in the main bay was 8.5 m in 1985 (434.8 masl). About 100 km to the south, Spiritwood Lake, North Dakota (47°05' N, 98°35' W) is a small, relatively deep dimictic lake (lake area 1.67 km², catchment area 60 km², maximum depth 16 m), occupying a meltwater channel, filled with glaciofluvial sediments following deglaciation (Winters, 1963). This channel is part of the Spiritwood Valley aquifer, an important component of local and regional hydrology (Huxel and Petri, 1963; Huxel and Petri, 1965). In August 1991, core SW1 (82 cm long) was retrieved close to the centre of the lake using a modified piston corer in 13.4 m of water.

Both cores were extruded in the field at 1 cm intervals, and samples were stored at 4°C prior to analysis. Dating control was achieved using ²¹⁰Pb (Appleby, 2001) at both sites, with a constant value subtracted for supported activity at Devils Lake (Fritz, 1990) and individual measurements of both supported and total activity at Spiritwood Lake (corroborated with measurement of ¹³⁷Cs and ²⁴¹Am; P.G. Appleby, University of Liverpool, UK; unpubl. data). A constant rate of supply model for ²¹⁰Pb was applied in both cases (Appleby and Oldfield, 1978) with additional dating control for Devils Lake from known regional pollen markers (Jacobson and Engstrom, 1989; Fritz, 1990).

2.3 Diatom dissolution indices

In order to include dissolution data in numerical modelling, an objective means of quantifying diatom dissolution is needed. Experimental dissolution of fresh valves and surface

sediment assemblages from the NGP in a high pH medium found that groups of morphologically similar taxa (e.g. within the genera *Stephanodiscus* and *Cyclotella*) underwent similar physical patterns of corrosion as dissolution progressed (Ryves, 1994). Such patterns were also observed in deeper core material undergoing natural dissolution (e.g. *Stephanodiscus niagarae* in Figure 1), providing justification for an experimental approach as a means to quantify dissolution in natural settings. Using a combination of both light (LM) and scanning electron microscopy (SEM), dissolution patterns for each major diatom shape were defined, with 2, 3 or 4 dissolution states (“stages”) identified for each taxon (stage 1 being the pristine state and higher stages indicating progressive dissolution). A description of dissolution stages for common taxa is given in Table 2 and illustrated in Fig.1 and Appendix A1.

Under routine counting under LM, all valves (excluding girdle views) were assigned into one of the categories. Two diatom dissolution metrics can be derived from these data (**F** and **DDI**; Ryves et al., 2001; Ryves et al., 2006). The **F** (fractional) index is the proportion of pristine valves (stage 1) to all valves that can be classed (Flower and Likhoshway, 1993; Ryves et al., 2001) and varies from 0 to 1, with **F** = 1 indicating all valves are perfectly preserved to **F** = 0 where all valves appear dissolved under LM. To maximise potential information from more dissolved assemblages, the "diatom dissolution index" (**DDI**, Ryves et al., 2006) was developed, which compares the assemblage to its dissolution endpoint:

$$DDI = \frac{\sum_{s=1}^{s=4} n_s \bullet (S - 1)}{N \bullet (S_{max} - 1)} \quad (1)$$

where **n** is the number of valves in stage **S**, and **N** is the total number of valves classified (i.e. excluding girdle views). **S_{max}** is the highest dissolution stage that valves in the assemblage

(from those present at 5% abundance or more) can reach and varies between 2 and 4. **DDI** also varies from 0 (perfect preservation) to 1 (all valves at highest dissolution stage for assemblage). A comparison between analysts (D.B. Ryves and R.W. Battarbee) counting 15 years apart on the US samples demonstrates that this methodology is robust for both indices (**F**: $r^2 = 0.89$ and **DDI**: $r^2 = 0.93$ for 27 NGP lakes, $p < 0.001$; Fig. 2).

2.4 Diatom analysis

Original slides from the NGP surface sediment diatom training set (Fritz et al., 1993) were re-examined to assess diatom dissolution of species and assemblages. Counts were made under phase-contrast LM at x1000 magnification (mean 291 valves per lake; >300 for 45 lakes). In only two cases were counts <100 valves (Shinbone and Horseshoe). In addition, the uppermost 19 samples (count sum 150–300 valves) from the Devils Lake short core (Fritz, 1990) and 18 samples from the Spiritwood Lake short core (average count 420 valves) were counted using the dissolution methodology. Dissolution assessments of individual species important within any sample (>5% abundance) were based on at least 30 valve views of that species. Analytical effort was focussed on accurately quantifying dissolution within a training set of known species composition, so as to compare the original (unadjusted) model with those incorporating dissolution data. For Alkaline and East Stump, dissolution data were obtained from original countsheets of R.W. Battarbee.

Diatom identification followed standard texts (Patrick and Reimer, 1966; Patrick and Reimer, 1975; Krammer and Lange-Bertalot, 1986; Krammer and Lange-Bertalot, 1988; Krammer and Lange-Bertalot, 1991a; Krammer and Lange-Bertalot, 1991b), more specialist publications dealing with specific saline lake taxa found in the NGP (e.g. Battarbee et al., 1982; Carvalho et al., 1995) and informal saline diatom taxonomy workshop protocols. Initially, raphid and araphid valves from heterovalvar taxa (e.g. *Achnanthes* spp., *Cocconeis* spp. and *Rhoicosphenia curvata*) were enumerated separately during dissolution counting, as they may

show different dissolution behaviour (Ryves et al., 2001). *Chaetoceros* cysts (thought to be *C. elmorei* or *C. muelleri*; Fritz et al., 1993) were classed as stage 2 unless associated with the weakly silicified enclosing cell (considered stage 1), but no further dissolution stages could be reliably identified for this taxon.

2.5 Numerical analyses

Dissolution data obtained directly during routine counting of trainings sets and cores can be used to create new inference models in several ways (Table 1). One approach is to consider the distribution of each dissolution stage of every taxon as a pseudospecies in its own right, which appears logical as dissolution is correlated to salinity within the NGP (Ryves et al., 2006). For this method, the original percentage abundance of each species in every sample (as found by Fritz et al., 1993) was split amongst its constituent “dissolution taxa” (undissolved or corroded under the **F** index, and with up to four stages using the **DDI** methodology). This was calculated using the relative proportion of valves in each dissolution stage found under the recount for that species in that sample. For any sample where taxa were not recorded during the recount (i.e. rare taxa not encountered due to smaller counts or different transects viewed under LM), the average dissolution index for the sample was used. The original distribution of Fritz et al. (1993) across the dataset for all species was thus maintained and provides a direct comparison with the unadjusted NGP models.

A second method involved using the **F** values for each species in each sample as an indicator of taxon-specific valve loss within that sample, using the empirical relationship observed from dissolution experiments for mixed assemblages (Ryves et al., 2001). A dissolution correction factor (F^*_{ij}) can be applied to each occurrence (y_{ij}) of species j in any sample i , and then new abundances ($F[y_{ij}]$) calculated for each species in every sample as follows:

$$F[y_{ij}] = \frac{y_{ij} \cdot F^*_{ij}}{\sum_{j=1}^{j=i} y_{ij} \cdot F^*_{ij}}$$

(Eq. 2)

The dissolution correction factor, F^*_{ij} , for any species' occurrence is derived from re-arranging the relationship between **F** and valve loss in Ryves et al. (2001):

$$F^*_{ij} = 10^{0.979 - 1.075 F_{ij}}$$

(Eq. 3)

where F_{ij} is the **F** index for species j in sample i . Values of F^*_{ij} range from ~9.5 when **F** = 0, and are capped at 1 where preservation is excellent (**F** ≥ 0.91). To avoid introducing excessive noise from species found at low abundance, the original dataset was screened and taxa occurring at <1% maximum abundance deleted. Taxa that were not recorded in this reduced dataset during the dissolution recount were assigned the assemblage **F** value for that sample. These models can be directly compared to unadjusted NGP models screened with a 1% cut-off (e.g. Fritz et al., 1993).

Transfer function development was carried out within the C2 program (version 1.4.3; Juggins, 2003). Diatom assemblage data were analysed as percentages but were otherwise untransformed. Salinity (expressed as g L⁻¹ TDS) was log-transformed to approximate normality and reduce the influence of large values. C2 includes all weighted averaging (WA) variants (WA with both classical and inverse deshrinking, and with tolerance downweighting of taxa), as well as WA-PLS, analogue matching (MAT) and maximum likelihood modelling (ML). Inverse deshrinking tends to improve overall model performance (lower errors), but may lead to less accuracy at extreme ends of the environmental gradient. Models were internally validated using the bootstrap method in C2 (1000 cycles; Birks et al., 1990). Apparent and predicted root-mean square errors (RMSE and RMSEP) values are reported as log₁₀ TDS. Salinity optima for species (and dissolution pseudospecies) were derived using WA.

User-defined weighting for individual samples and species within C2 is only available by deletion (zero weight; Table 1) but can be implemented within CANOCO to derive a simple WA transfer function with classical deshrinking (ter Braak and Šmilauer, 2002), although there is no straightforward method of assessing predictive error. Given these restrictions, weighting of samples and species using either directly observed dissolution data, or indirectly according to relative resistance to dissolution was not tested in the current study. Previous work using variable weighting based on species' relative resistance to dissolution derived from experiments has suggested that only minor improvements can be expected (Ryves, 1994).

Model validation against the Devils Lake salinity record was based on a short core covering the period of historical salinity measurements (Fritz 1990). Salinity (or conductivity) has been measured sporadically at Spiritwood Lake since the mid-1970s (M. Sauer, North Dakota Department of Health, pers. comm.), with isolated values from 1965 - 72 (G. Van Eeckhout, North Dakota Game and Fish Department, pers. comm., Tarapchak, 1973; values converted to salinity). Although no measured data exist for earlier periods, there is some anecdotal evidence for lake level change from written accounts over the last century (Johnson, 1950). Goodness-of-fit of reconstructions also were estimated statistically by comparing fossil and training set squared residual distances from a single-variable CCA salinity axis (Birks et al., 1990). If fossil samples lie in the upper 10% or 5% of the distribution of such distances in the surface sediments (log transformed for normality), they are deemed to have a "poor" or "very poor" fit respectively to the reconstructed variable (Birks et al., 1990; Laird et al., 1996a).

3. Results

3.1 Dissolution taxonomy and methodology

There is excellent agreement between taxonomists for both **F** and **DDI** with the slopes for both indices indistinguishable from 1 (Fig. 2). Given the time interval separating these counts, this

suggests that objective criteria for assessing dissolution exist and are robust. Four dissolution stages were recognised for common *Cyclotella*, larger *Amphora* and Naviculoid taxa; three for many other pennates (e.g. *Anomoeoneis*, *Gomphonema*, and other smaller Naviculoid taxa), centrics (*Stephanodiscus*) and Surirellaceae; and two for fine and/or small taxa such as many *Nitzschia*, *Cocconeis* and *Achnanthes* species (Table 2, Fig. 2, Appendix 1). Descriptions for each dissolution stage (Table 2) are based on their appearance under LM at x1000 so that the method can be used during routine counting. SEM imagery, however, shows that substantial dissolution can occur before valves appear dissolved under LM (Appendix 1), with areolae becoming enlarged (cf. Fig 1F) and silica surfaces developing a “spongy” appearance in valves otherwise classed as pristine under LM.

3.2 Dissolution pseudospecies within the NGP

As diatom dissolution is correlated with salinity, salinity optima calculated for successive dissolution stages should increase within an individual species and is the logic by which dissolution data may improve diatom-salinity models and inferences. Figure 3 shows the WA salinity optima for all dissolution stages for the 39 species in NGP53 that have a maximum abundance of at least 10% and occur in 5 lakes or more, with 4 species split into raphid and araphid forms. For 31 taxa (72%), less dissolved valves have lower salinity optima than more dissolved, and these often plot in numeric order where 3 or 4 stages are recognised. Dissolution salinity optima can span the species optimum by 0.5 log-salinity units or more, equivalent to a factor of 3. For example, NA054A (*Navicula veneta*; Fig. 3) has a salinity optimum of 4.94 g L⁻¹ TDS, representing the weighted average of its three component dissolution stages, with optima of 2.74 (stage 1 = undissolved), 4.95 (stage 2) and 14.71 (stage 3) g L⁻¹ TDS respectively; *Cyclotella quillensis* (CY017A; optimum 12.53 g L⁻¹ TDS) has 4 dissolution stages (spanning 8.17 - 14.68 g L⁻¹).

Figure 3 suggests there are only minor differences in the salinity range of raphid and

araphid valves, in contrast to results from dissolution experiments which found sometimes large differences in robustness between these valve types (Ryves et al., 2001). Such differences may be obscured by other taphonomic effects during transport and sedimentation of these periphytic taxa from their littoral habitats to offshore, deeper water sediments. In the light of this, heterovalvar species were not separated into raphid and araphid taxa within NGP dataset for subsequent transfer function development.

Many of these dissolution taxa exhibit a unimodal response to salinity as does the species from which they derive, an assumption of WA methods. Figures 4 and 5 show the distribution with respect to salinity of *Cyclotella quillensis* (for both NGP53 and NGP55) and *Navicula cincta* for all observations and separated according to dissolution stage (4 stages for *C. quillensis* and 3 for *N. cincta*). The estimated WA salinity optima and the modelled response under logistic regression (predicted unimodal fit) for NGP53 are in close agreement and suggest true unimodal response for these dissolution taxa, even for pseudospecies that do not show large increases in salinity optima for successive dissolution stages (e.g. *Navicula cincta*, Fig. 5). These results suggest that the distribution of some taxa, and so their “aggregate” WA optima, are strongly influenced by dissolution. In particular, the species optimum for *C. quillensis* is largely defined by badly preserved samples as most valves in the NGP are dissolved (stage 1 optimum of 8.17 g L⁻¹ TDS; Fig. 4). Including East Stump (77% *C. quillensis*, all stage 3 or 4) in the training set increases the species salinity optimum from 12.53 to 16.22 g L⁻¹ TDS, through its effect on optima for stages 3 and 4 while leaving stage 1 and 2 optima unchanged (Fig. 4). This clearly has implications for ecological inferences made from surface sediment training sets that are affected by such taphonomic effects, where the surface sedimentary abundances may not be an accurate reflection of the living community.

3.3 Dissolution-adjusted salinity models

Table 3 summarises the results of modelling experiments incorporating dissolution

information (Table 1) on the NGP training sets using a variety of indirect and direct approaches (models E – K) compared to unadjusted models (A – D). For dissolution-adjusted models (Table 3, E – K), WA methods outperformed WA-PLS, MAT and ML methods in all cases. For each training set, the “best” apparent and predicted models are shown (as judged by lowest RMSE and RMSEP respectively). In all cases, the best apparent models (lowest RMSE) were produced under WA with tolerance downweighting and inverse deshrinking ($WA_{tol_{inv}}$), while lowest predicted errors (minimising RMSEP) were found using simple WA models (but always under inverse deshrinking for models ignoring dissolution, and classical deshrinking for those incorporating dissolution information).

The unadjusted NGP models (Table 3, A – D), with all taxa and a cut-off of 1%, appear robust with low RMSE (around 0.17 log-salinity units, LSU). These errors are likely underestimated, however, as the bootstrapping analysis shows, with RMSEP of around 0.3 LSU. Models display apparent systematic underestimation at higher salinity, which correlates to some extent with dissolution state of the training set sample (Fig. 6a), although when residuals are plotted against predicted values (Racca and Prairie, 2004) this trend disappears. Models tend to improve as the number of taxa increases, in agreement with other studies (Wilson et al., 1996; Birks, 1998). Within its salinity range, all NGP salinity transfer functions compare favourably with a 219 lake training set from British Columbia (Wilson et al., 1996; salinity range 0.02-620 g L⁻¹ TDS, no. taxa 204, $r^2 = 0.87$, $RMSE_{boot} = 0.371$ LSU), and others (e.g. Gasse et al., 1995; Reed, 1998a; Davies et al., 2002; Yang et al., 2003).

All transfer functions making use of dissolution data produce models that appear to perform better, though these improvements are minor when validated under bootstrapping compared to the unadjusted models (Table 3, E – K). Under model E (NGP53), deleting four samples that are badly preserved ($F < 0.15$; Bitter (South Dakota), George, Hazelden and Rabbit) removes only one lake with a residual > 0.3 LSU (Hazelden; Fig. 6). Samples from Mission Bay, Boucher, Porter and Eckelson all have residuals > 0.4 LSU but are better

preserved.

By definition models built using dissolution pseudospecies contain many more taxa than the unadjusted training set, with 365 taxa where species are separated into pristine and dissolved categories (Table 3, F and G), and 536 taxa where all dissolution stages are considered (Table 3, H and I). These two approaches produce models with the lowest apparent RMSE values generated (~ 0.14 LSU for 53 sites), a marked improvement on unadjusted models ($\sim 15\%$ lower). RMSEP values are not better, however (~ 0.3 LSU). Recalculating species' abundances using the **F** index (Table 3, J) produced models with similar RMSE to unadjusted models, and marginally lower RMSEP, though still ~ 0.28 LSU. Examination of the residuals of model J against predicted salinity of the training set lakes (Racca and Prairie, 2004) suggested that Eckelson Lake was an outlier (~ 0.55 LSU; cf. Fig. 6b), which is not the case with unadjusted NGP53 models (Fig. 6a). Deleting this site improved the model further (Model K; Fig. 6b, Table 3), with both RMSE and RMSEP $\sim 10\%$ lower than any of the unadjusted models, and appears to reduce bias at high salinity (Fig 6b).

3.4 Application of salinity models at Devils Lake and Spiritwood Lake

Summary diatom diagrams are shown for Devils Lake (Fig. 7, adapted from Fritz 1990) and Spiritwood Lake (Fig. 8) including all taxa found at $>5\%$ in any sample. The proportion of pristine and dissolved valves for each of the taxa found in each sample are indicated, together with sample **F** index, observed lake level and salinity (Devils Lake) and sedimentation rate data (Spiritwood Lake).

3.4.1 Devils Lake diatom stratigraphy

The recent diatom record at Devils Lake is discussed in detail in Fritz (1990) and only a brief summary is given here (Fig. 7). Over the last ~ 100 years, Devils Lake shows dramatic shifts from an assemblage dominated by freshwater planktonic *Stephanodiscus* taxa (*S. minutulus* and

S. niagarae) to one dominated by saline planktonic species, especially *Cyclotella quillensis* and *Chaetoceros* cysts, in keeping with the observed (and inverse) trends in lake level and salinity. Diatom dissolution is significant ($F \sim 0.4$; Ryves et al. 2006) for all samples before ~1960, even when water levels were high in the early 20th century.

3.4.2 Spiritwood Lake diatom stratigraphy

The diatom record at Spiritwood Lake (Fig. 8) is dominated by planktonic taxa throughout, especially *Stephanodiscus minutulus* and *S. niagarae*, which tend to co-vary inversely to each other (and may reflect Si:P availability and water column turbulence). The flora is indicative of freshwater ($< 0.5 \text{ g L}^{-1}$ TDS) to subsaline ($0.5 - 3 \text{ g L}^{-1}$ TDS) conditions throughout much of the recent past. A period of elevated salinity is suggested in the 1940s and 1950s, when *Cyclotella meneghiniana* (and the appearance of *C. quillensis*) and *Chaetoceros elmorei/muelleri* cysts make up over 30% of the assemblage. Throughout the 1980s, the epiphytic *Cocconeis placentula* becomes important, declining in the upper samples with the rise of the planktonic freshwater *Fragilaria capucina* var. *mesolepta*, in concert with lake management practises to reduce total phosphorus (TP) concentrations. The surface sediment is dominated (~80%) by very well preserved *F. capucina*, likely reflecting a spring bloom when the lake was cored (April 1991).

The **F** index fluctuates between 0.2 and 0.6 throughout the record before ~1970, when there is a transition to higher values indicating better preservation (reaching ~0.9 in the surface sample). Sediment accumulation rates reflect the long term rise in agricultural intensification throughout the region since European immigration in the 19th century, falling in the 1950s and 1960s, but increasing to the present.

3.4.3 Salinity inference at Devils Lake and Spiritwood Lake

Both original models (i.e. ignoring dissolution) and dissolution-adjusted models were applied to the Devils Lake and Spiritwood Lake short cores and compared with measured salinity

at both sites (Fig. 9). Results are shown for comparable 55 and 53 lake datasets with lowest RMSEP (models A, C, F and G in Table 3). At each site, the general patterns of salinity inference are similar, and at Devils Lake (Fig. 9a), all reconstructions are significantly correlated with observed salinity ($p < 0.01$, $n = 15$). Sample-specific errors are substantially lower than predicted (RMSEP), in all cases being ~ 0.1 LSU ($\sim 20\%$ of the back-transformed value), and hence profiles of inferred salinity show significant differences both within and between models at each site.

The effect of deleting Alkaline and especially East Stump (the latter dominated by dissolved *C. quillensis*) is clear at Devils Lake as DI-salinity values are significantly lower under the NGP53 models. At salinity above $\sim 5 \text{ g L}^{-1}$ TDS at Devils Lake, dissolution models tend to infer higher values relative to their unadjusted counterparts, particularly in samples before ~ 1950 , where most valves are dissolved, although highest salinity values in the 1930s-40s are likely underestimated under all models (but to which there are no observed salinity data for comparison). Conversely, dissolution models appear to overestimate salinity in the late 19th century compared to unadjusted models. DI-salinity tends to be inversely related to sample dissolution index, and this relationship is emphasised by dissolution models.

Inferred salinity at Spiritwood Lake (Fig. 9b) over the last ~ 150 years is relatively low compared to Devils Lake, with values between $1\text{--}3 \text{ g L}^{-1}$ TDS (depending on model) for much of this time, except during a period of elevated salinity in the 1940s-50s. At Spiritwood Lake, salinity inference under the 55 site models tends to be higher than the 53 site models in all cases, again except for low values from the mid-1970s, where all models infer very similar values. Dissolution models differ little from unadjusted models at this site, except during the 1940s-50s. The NGP55 dissolution model (G, Table 3) infers the highest salinity of all models in the mid-1950s (almost 5 g L^{-1} TDS), suggesting that salinity was above the saline threshold (3 g L^{-1} TDS) from the 1940s, in contrast to other models (including the NGP53 dissolution model). All DI-inferred values agree well with salinity measurements from the mid-1970s when salinity was low, but observed salinity in 1970 and especially 1965 agree better with the

dissolution models, although too few observations exist for a proper validation. As at Devils Lake, the **F** index varies inversely to DI-salinity, even at these low inferred values.

4. Discussion

4.1 Developing a dissolution taxonomy

Repeatable dissolution patterns exist for many valve shapes observable under LM across different aquatic systems (Flower, 1993; Barker et al., 1994; Mackay et al., 1998), and the strong agreement between different taxonomists (using independently derived dissolution schemes) suggests that these can be objectively defined. While previous research has used the dissolution state of one, two or three common taxa to categorise sample preservation (e.g. Barker, 1992; Flower and Likhoshway, 1993; Mackay et al., 1998; Warnock et al., 2007), basing an assessment upon all valve views counted allows samples to be compared across very different ecological and limnological (or oceanographic) conditions and reduces possible errors from low abundance or reworking that may affect any single taxon. Evaluating dissolution under LM may not be as accurate or detailed as under SEM (e.g. Warnock et al., 2007) but can be easily incorporated into routine counting strategies, without the need for additional sample preparation and analysis.

Dissolution stages for taxa not included here can be easily established by LM examination of material across the full range of dissolution state (ideally obtained from dissolution experiments, where other factors, such as breakage, can be controlled). Examination under SEM can support this and demonstrates that extensive valve corrosion occurs before effects are noticeable under LM. This may help explain the exponential decline of valve number with increasing dissolution (as recorded by **F** or **DDI**; Ryves et al., 2001), as weakened valves may easily break into uncountable fragments leaving a much reduced residual assemblage identifiable during progressive dissolution. Previous work has noted the strong relationship between dissolution and fragmentation in both surface sediments (Ryves et al., 2006) and core studies in both freshwaters (Straub, 1993) and marine sediments (Warnock et al., 2007).

4.2 Incorporating dissolution data into models

Dissolution experiments allow taxa to be ranked according to their susceptibility to dissolution, and offer an indirect means of incorporating dissolution data into models by downweighting resistant species (and/or samples, if weighted by the rank of constituent species; Ryves 1994). Although many studies have used the presence or proportion of robust and/or susceptible taxa as indicative of dissolution (e.g. Johnson, 1974; Shemesh et al., 1989; Pichon et al., 1992) weighting species and samples according to dissolution resistance is an *ad hoc* approach to the more fundamental problem of actual sample and species preservation within training sets and fossil data. While modest improvements to a version of the NGP salinity model were reported in Ryves (1994), other approaches offer more appropriate solutions. Firstly, it is logistically challenging to include all taxa in dissolution experiments, especially in larger, more diverse training sets (especially as living rather than sedimentary material should be used; Ryves et al., 2001). Secondly, the approach assumes that species' distribution is always a reflection of taphonomy and that dissolution state for any sample can be inferred from the taxa present (for example, downweighting samples even when assemblages consist of well-preserved, but resistant taxa, due to other ecological and environmental conditions). Thirdly, the weightings themselves are arbitrary if based on relative (not absolute) resistance.

Weighting samples inversely according to dissolution state (downweighting or deleting, i.e. weight = 0) offers a direct, if blunt, approach (Model E, Table 3). 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. Downweighting samples according to dissolution index will tend to dampen the response the model at higher salinities as there is a bias towards poor dissolution at higher salinities within the NGP 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* (important at Devils Lake in samples pre-1970) which has its greatest abundances in generally badly preserved samples (e.g. East Stump).

Salinity itself may interact with resistance to dissolution in complex ways. The greater separation of dissolution stages at lower salinity (Fig. 3), and the general trend that dissolution increases with salinity independent of pH and depth (Ryves et al., 2006), despite silica solubility itself declining, may also in part reflect processes of biogenic silica formation. Marine diatoms have been shown to produce silica of lower density (and so higher solubility, and lower resistance to dissolution) at higher salinity (Vrieling et al., 2007), which raises the possibility that dissolution behaviour is not just taxon-specific but salinity-dependent too. This might have implications for ranking schemes and correction factors using data for individual species derived from a range of lakes of different salinity.

Nonetheless, results from this study suggest that approaches making use of individual dissolution state for each taxon within each sample (either by dissolution taxa or individual correction factors; models F-K in Table 3) can improve both models and reconstructions, although more complex models (e.g. with more “species”, or abundances adjusted by **F** scores) may be produced. Even the correction method using **F** index scores (models J and K) is an oversimplification, as aggregate, assemblage-wide relationships between species loss and **F** index (from Ryves et al. 2001) are applied. Previous work has suggested that species exhibit a range of population-dissolution relationships (Ryves 1994), and while it may be unrealistic to evaluate these for all species in a training set or core, where only a few taxa dominate, it may be worth exploring these further (e.g. in Lake Baikal and the African Great Lakes; Battarbee et al., 2005; Mackay et al., 2005). In fact, recent evaluations of diverse diatom training sets have shown that models based on only a small fraction of the total dataset are as robust as the full model (e.g. only 15% of taxa for the SWAP pH dataset; Racca et al., 2003). Using dissolution data to recalculate species abundance for key species, and basing new models on these, may be a fruitful approach for other, species-rich datasets, requiring evaluation of only a few

correction factors.

4.3 Salinity inferences at Devils Lake and Spiritwood Lake

Incorporating dissolution data into models does lead to substantial differences in inferred salinity at both sites, particularly during more saline periods. As samples become more dissolved (driven by increasing salinity and lower lake level; Barker, 1992; Reed, 1998b; Ryves et al., 2006), models making use of information on the salinity distribution of more dissolved stages within the NGP training set infer elevated salinity compared to models ignoring this (Fig. 9). All salinity models can differentiate between high and low salinity events but as inferred salinity co-varies with sample dissolution state, errors are strongly correlated with sample dissolution index within the fossil assemblages (Ryves et al., 2006). As preservation is generally good in the freshwater lakes within the training set, all models produce very similar results in this part of the gradient. Given the strong effect of salinity on preservation, however, in situations where salinity is perhaps not the main factor controlling dissolution, this can lead to error.

While it is unclear why preservation was poor during high lake level at the end of the 19th century at Devils Lake (leading to overestimation of inferred salinity under dissolution models), low lake level during the Dust Bowl likely had important taphonomic impacts. In this case, while dissolution models infer higher salinity than comparable unadjusted models (Fig. 9a), salinity is underestimated in these very poorly preserved samples. Dissolution does destroy information and while inferences may be improved (in the sense of being nearer measured values), beyond some threshold reconstructions cannot be improved. This will depend on the proportion of robust taxa (i.e. identifiable when heavily dissolved) in the original fossil sample, the optima of which will determine the inferred salinity as dissolution progresses to an end state. Ultimately, even dissolution models will produce underestimates at high salinity as the distribution of taxa (and dissolution pseudospecies) is truncated by progressive dissolution, with optima biased to the salinity at which some preservation occurs. For this reason, even though valve counts for Alkaline

and East Stump are less than 50, inclusion of these samples does allow higher salinity values to be inferred, and in some sense provides analogues with poorly preserved samples found in fossil datasets. Inclusion of these sites increased the range of theoretical salinity inference for dissolution models using pseudospecies (models G and I in Table 3) from $\sim 150 \text{ g L}^{-1}$ TDS to $> 300 \text{ g L}^{-1}$ TDS, as estimated by running a dummy sample containing a single taxon with the greatest salinity optimum for each model.

At Spiritwood Lake (Fig. 9b), where dissolution is perhaps more simply linked to salinity without additional taphonomic effects from major lake level change for example, including Alkaline and East Stump does seem to improve the inferences under either 53 site model (whether or not dissolution data are incorporated). Measured salinity in 1965 is close to values inferred under the 55 site models, and about twice that under either 53 site model (thus beyond model error), and fits best with the 55 site dissolution model (G in Table 3). Unfortunately, no measurements of salinity are known before 1965. Lake level was known to be low in 1950 relative to the late 19th century (Johnson, 1950), although this could support all reconstructions. However, if dissolution model G is accepted as being the most accurate available, it would suggest that the lake crossed the saline threshold (3 g L^{-1} TDS) by the 1940s and remained so throughout the 1950s, and would therefore be more responsive to regional aridity than other models imply (though still lagged compared to Devils Lake).

Palaeolimnological studies within the NGP have shown that lakes only 100-200 km apart may have distinctly different salinity histories, even to major climate events that are considered regionally synchronous (Fritz et al., 2000; Laird et al., 2003). Compared with Devils Lake, Spiritwood Lake (Fig. 9) appears to show a damped and lagged response to the Dust Bowl drought of the 1930s-40s, with maximum salinity values occurring in the 1950s, and slightly elevated salinity in the 1970s, perhaps in response to regional drought in the 1960s (cf. Devils Lake, Fig. 7).

Groundwater (via connection to the Spiritwood Valley aquifer) likely plays an

important role in modulating the chemical response to drought. Anecdotal evidence over the 20th century suggests Spiritwood Lake has never become strongly saline over this time, and it was known in pre-European Sioux folklore as a permanent source of drinking water (Johnson, 1950). Large inter-annual lake level changes of several metres (perhaps up to 30% lake volume) have been recorded in the 20th century (Johnson, 1950), however, suggesting that the lake is sensitive to fluctuations in the annual water balance, but the diatom record implies that the inverse relationship between volume and salinity commonly assumed for closed basins does not hold here. Spiritwood Lake diatom assemblages, and hence DI-salinity, since the early 1980s may also be affected by a programme of lake management, aimed at reducing the levels of phosphorus in upper, photic waters by pumping out hypolimnetic waters enriched in P (and dissolved solids) during lake stratification (M. Sauer, pers. comm.), further decoupling this link in the contemporary lake system.

4.4 Dissolution profiles and taphonomic pathways for taxa

As well as using dissolution data for adjusting inference models, dissolution profiles for individual taxa can aid interpretation of sedimentary records by providing insights into, and provoking hypotheses about, sedimentation processes and the means by which environmental signals are transferred to and preserved in sedimentary archives. Valves of different species making up the sedimentary assemblage may have travelled along different taphonomic pathways (e.g. re-working of deposits under major lake level change at Devils Lake, or different patterns of seasonality at Spiritwood Lake), which may be traced by comparison of individual species' *F* values. In addition to analogue matching (Birks et al., 1990) and goodness-of-fit measures (Kingston et al., 1992; Laird et al., 1996a), dissolution indices also provide another means of flagging samples whose inferred values should be treated with caution (Ryves et al., 2006).

Although none of the Devils Lake samples were flagged as outliers by conventional goodness-of-fit measures methods, several assemblages contained significant percentages of

both freshwater (e.g. *Stephanodiscus niagarae*; optimum 1.24 g L^{-1} TDS and *S. minutulus*; 2.06 g L^{-1} TDS) and saline taxa (e.g. *Cyclotella quillensis*; 12.53 g L^{-1} TDS). Although major seasonal salinity changes can occur in such closed lakes, inspection of the distributions of these taxa in the NGP dataset shows there is very little overlap. Species' **F** values within the Devils Lake sequence differ markedly however, with the larger, more robust *C. quillensis* valves much less well preserved, contrary to expectation (Barker et al., 1994; Ryves et al., 2001; Fig. 7). This suggests that the source of some of these *C. quillensis* valves in the upper part of the record may be older sediments laid down during higher salinity periods reworked during subsequent lake filling. Reconstructions in this section may be overestimated as a result.

In earlier core sections, problems of dating, resuspension and sediment reworking may affect the diatom record during the low water stand of the 1920s to 1940s, when salinities were likely highest (Fritz, 1990; Fig. 7). Diatom-inferred salinity is driven by *Cyclotella quillensis* and *Chaetoceros elmorei/muelleri* over this period, which have salinity optima $< 15 \text{ g L}^{-1}$ TDS for all models (considered as composite species or dissolution pseudospecies). Given the poor preservation of samples in this section, it may be that taxa with higher salinity optima (e.g. *Cyclotella choctawhatcheana*, optimum 21.5 g L^{-1} TDS) have been preferentially removed from the record, although dissolved central areas of this taxon are distinctive. At Devils Lake, sample preservation itself may best track changing salinity ($r^2 = 0.77$ for **F** against $\log(\text{salinity})$, $n = 15$, $p < 0.001$; Fig. 7) perhaps because this is amplified by lake level change crossing critical thresholds where the additional impacts of resuspension and breakage are important. Under such conditions, recovery of the true death assemblage may be impossible, and salinity reconstructions can only reflect what remains in the assemblage, regardless of techniques to adjust or rebuild salinity models (Tables 1, 3). Errors in reconstructed values may then be directly related to preservation state (Ryves et al., 2006).

In the Spiritwood Lake core, **F** index scores for *S. minutulus* and *S. niagarae* are correlated ($r \sim +0.45$, $n = 16$ with valve counts > 10 for each, $p < 0.05$) but *S. minutulus* valves are

on average better preserved than *S. niagarae* ($F = 0.48$ and 0.21 respectively) again contrary to what might be expected from gross valve morphology (Barker et al., 1994). Such patterns may relate to sedimentation processes associated with each species. *S. minutulus* blooms predominantly in the spring/summer (Lund, 1950; Holland, 1969; Interlandi et al., 1999), associated with low Si/P ratios following strong overturn after ice-out (and replenishment of P from the anoxic hypolimnion to the upper water column). By early summer, cells will sediment into anoxic hypolimnetic waters, providing good conditions for preservation where bioturbation and mixing are minimised. In contrast, *S. niagarae* generally has greatest abundance during autumn (Holland, 1969; Baker and Baker, 1981), linked to turbulent conditions and higher Si/P ratio, when the major stratification period is over, with sedimentation into oxic bottom waters.

In a diatom record from the Aral Sea (Austin et al., 2007), basin morphometry presents the unusual situation for a relatively shallow lake where low diatom-inferred salinity (and high lake level) is inferred at times when the assemblage is dominated by periphytic taxa (e.g. the benthic *Amphora pediculus* and epiphytic *Cocconeis neodiminuta*). Source-area habitat modelling has demonstrated this possibility (Stone and Fritz, 2004), and Austin et al., (2007) argue that here it reflects expansion of littoral areas in the eastern Aral basin flooded during higher lake levels (and vice versa). Although the case is not explicitly made, diatom preservation (assessed by Austin et al. using the F index) may support this argument as sample F scores are lower in these sections and rise when diatom-inferred salinity rises (and levels fall), dropping again at higher salinity (see Fig. 3 in Austin et al., 2007). This is consistent with a source of diatom valves some distance from the coring site at high lake level (and so dissolution during transport), with better preservation as littoral habitats move nearer as lake level falls, until higher salinity again reduces preservation. It might be possible to test this hypothesis using habitat-specific dissolution data: if correct, at this site one would expect periphytic forms to be, on average, less well preserved than pelagic taxa at higher lake levels (plankton sedimenting more locally, and above the coring site itself) and vice versa as levels drop. Dissolution indices can highlight

differences in habitat or seasonality and shed light on questions of representativity in pelagic and benthic sedimentation in a variety of situations.

5. Conclusions

Information is inevitably lost along the taphonomic pathway from living community (biocoenosis) to the death assemblage (thanatocoenosis) preserved in sedimentary archives, weakening the link between fossil data and their contemporary environment. Incorporating preservation information from the thanatocoenosis (directly or indirectly) may help reduce the disjunction between living community and sedimentary assemblage, strengthen the link between contemporary environment and sedimentary signal, and improve our inferences of past environments. Dissolution information is most effectively applied in conjunction with (rather than as a substitute for) ecological and sedimentological knowledge, for example in determining whether a fossil occurrence is reworked, as well as the basis for adjusting sedimentary abundances. The quantitative treatment of microfossil preservation may extend the range over which valid and useful palaeoenvironmental inferences can be made.

Diatom dissolution indices allow diatom preservation to be quantified and samples to be compared objectively rather than on an ad hoc basis. Sample dissolution indices can be used diagnostically (as a data screening process and to flag samples) and can be used to adjust inferential models, by supplying an objective means of weighting samples, for instance. Salinity transfer functions incorporating dissolution data perform as well as or better than orthodox models even under internal validation, but (within the NGP) all suffer from the bias (underestimation) 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; Ryves et al., 2006) in the dataset can be expected to improve the transfer function at higher salinities as species parameters will be more accurate, while including even badly dissolved samples in dissolution models can enhance

inferences of past environmental change.

6. Acknowledgements

Part of this work was carried out within NERC grant GR9/02033 to DBR and RWB, and additionally during a NERC studentship to DBR (GT4/90/ALS/28), and during the 1st LIMPACS Salinisation Workshop, Mildura, Australia, September 29th – October 4th 2004. DBR's attendance at this workshop was supported by a grant from the Carlsberg Foundation, Denmark. We also thank Steven M. Robinson, North Dakota Water Science Center, U. S. Geological Survey; Mike T. Sauer, North Dakota Department of Health, and Gene R. Van Eeckhout, North Dakota Game and Fish Department, for supplying conductivity data for Spiritwood Lake.

7. References

- Appleby, P.G., 2001. Chronostratigraphic techniques in recent sediments. In: Last, W.M., Smol, J.P. (Eds.), *Tracking Environmental Change Using Lake Sediments - Volume 1: Basin Analysis, Coring, and Chronological Techniques*. Developments in Paleoenvironmental Research. Springer, New York, pp. 171-203.
- Appleby, P.G., Oldfield, F., 1978. The calculation of ^{210}Pb dates assuming a constant rate of supply of unsupported ^{210}Pb to the sediment. *Catena* 5, 1-8.
- Austin, P., Mackay, A., Palagushkina, O., Leng, M., 2007. A high-resolution diatom-inferred palaeoconductivity and lake level record of the Aral Sea for the last 1600 yr. *Quaternary Research* 67, 383-393.
- Baker, K.K., Baker, A.L., 1981. Seasonal succession of the phytoplankton in the Upper Mississippi River. *Hydrobiologia* 83, 295-301.
- Barker, P., 1992. Differential diatom dissolution in Late Quaternary sediments from Lake Manyara, Tanzania: an experimental approach. *Journal of Paleolimnology* 7, 235-251.

- Barker, P., Fontes, J.-C., Gasse, F., Druart, J.-C., 1994. Experimental dissolution of diatom silica in concentrated salt solutions and implications for paleoenvironmental reconstruction. *Limnology and Oceanography* 39, 99-110.
- Battarbee, R.W., 2000. Palaeolimnological approaches to climate change, with special regard to the biological record. *Quaternary Science Reviews* 19, 107-124.
- Battarbee, R.W., Keister, C.M., Bradbury, J.P., 1982. The Frustular Morphology and Taxonomic Relationships of *Cyclotella quillensis* Bailey, 7th Diatom-Symposium, pp. 173-184.
- Battarbee, R.W., Mackay, A.W., Jewson, D., Ryves, D.B., Sturm, M., 2005. Differential dissolution of Lake Baikal diatoms: correction factors and implications for palaeoclimatic reconstruction. *Global and Planetary Change* 46, 75-86.
- Berger, W.H., 1968. Planktonic foraminifera: selective solution and paleoclimatic interpretation. *Deep-Sea Research* 15, 31-43.
- Birks, H.H., Birks, H.J.B., 2006. Multi-proxy studies in palaeolimnology. *Vegetation History and Archaeobotany* 15, 235-251.
- Birks, H.J.B., 1998. D.G. Frey & E.S. Deevey Review #1: Numerical tools in palaeolimnology - progress, potentialities, and problems. *Journal of Paleolimnology* 20, 307-332.
- Birks, H.J.B., Birks, H.H., 1980. *Quaternary palaeoecology*. Edward Arnold, London.
- Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C., ter Braak, C.J.F., 1990. Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society, London: Series B* 327, 263-278.
- Cameron, N.G., 1995. The representation of diatom communities in a small acid lake. *Journal of paleolimnology* 14, 185-223.
- Carvalho, L.R., Cox, E.J., Fritz, S.C., Juggins, S., Sims, P.A., Gasse, F., Battarbee, R.W., 1995. Standardizing the Taxonomy of Saline Lake *Cyclotella* spp. *Diatom Research* 10, 229-240.

- Conley, D.J., Schelske, C.L., Stoermer, E.F., 1993. Modification of silica biogeochemistry with eutrophication in aquatic systems. *Marine Ecology Progress Series* 101, 179-192.
- Davies, S.J., Metcalfe, S.E., Caballero, M.E., Juggins, S., 2002. Developing diatom-based transfer functions for Central Mexican lakes. *Hydrobiologia* 467, 199-213.
- Dixit, S., Van Cappellen, P., van Bennekom, A.J., 2001. Processes controlling solubility of biogenic silica and pore water build-up of silicic acid in marine sediments. *Marine Chemistry* 73, 333-352.
- Flower, R.J., 1993. Diatom preservation: experiments and observations on dissolution and breakage in modern and fossil material. *Hydrobiologia* 269/270, 473-484.
- Flower, R.J., Likhoshway, Y.V., 1993. An investigation of diatom preservation in Lake Baikal. In: Grachev, M.A. (editor), *Fifth Workshop on Diatom Algae: Diatom algae as indicators of changes in climate and environment*. Russian Botanical Society Publications, Irkutsk, Irkutsk, pp. 77-78.
- Fritz, S.C., 1990. Twentieth-century salinity and water-level fluctuations in Devil's Lake, North Dakota: test of a diatom-based transfer function. *Limnology and Oceanography* 35, 1771-1781.
- Fritz, S.C., Ito, E., Yu, Z.C., Laird, K.R., Engstrom, D.R., 2000. Hydrologic variation in the northern Great Plains during the last two millennia. *Quaternary Research* 53, 175-184.
- 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. *Canadian Journal of Fisheries Aquatic Sciences* 50, 1844-1856.
- 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.
- Gasse, F., Juggins, S., Khelifa, L.B., 1995. Diatom-based transfer functions for inferring past hydrochemical characteristics of African lakes. *Palaeogeography, Palaeoclimatology,*

- Palaeoecology 117, 31-54.
- Holland, R.E., 1969. Seasonal Fluctuations of Lake Michigan Diatoms. *Limnology and Oceanography* 14, 423-&.
- Hutson, W.H., 1977. Transfer functions under no-analog conditions: experiments with Indian Ocean planktonic foraminifera. *Quaternary Research* 8, 355-367.
- Huxel, C.J., Petri, L.R., 1963. Geology and ground water resources of Stutsman County, North Dakota: Part 2 - Ground water basin data, North Dakota Geological Survey, Grand Forks, North Dakota.
- Huxel, C.J., Petri, L.R., 1965. Geology and ground water resources of Stutsman County, North Dakota: Part 3 - Ground water and its chemical quality, North Dakota Geological Survey, Grand Forks, North Dakota.
- Interlandi, S.J., Kilham, S.S., Theriot, E.C., 1999. Responses of phytoplankton to varied resource availability in large lakes of the Greater Yellowstone Ecosystem. *Limnology and Oceanography* 44, 668-682.
- IPCC, 2007. Climate Change 2007 - The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, 1056 pp.
- Jacobson, H.A., Engstrom, D.R., 1989. Resolving the chronology of recent lake sediments: an example from Devil's Lake, North Dakota. *Journal of Paleolimnology* 2, 81-97.
- Johnson, E.M., 1950. Spiritwood Lake. *Stutsman County Record*.
- Johnson, T.C., 1974. The dissolution of siliceous microfossils in surface sediments of the eastern tropical Pacific. *Deep Sea Research* 21, 851-864.
- Jongman, R.H.G., ter Braak, C.J.F., van Tongeren, O.F.R., 1995. Data analysis in community and landscape ecology. Cambridge University Press, Cambridge, England, 299 pp.
- Juggins, S., 1992. Diatoms in the Thames Estuary, England: ecology, palaeoecology, and salinity transfer function. *Bibliotheca Diatomologica*, 25. J. Cramer, Stuttgart, 216 pp.

- Juggins, S., 2003. C2 User guide. Software for ecological and palaeoecological data analysis and visualisation (C2 version 1.4.3, build 1). University of Newcastle, Newcastle upon Tyne, 69 pp.
- 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 Sciences* 49, 116-127.
- Krammer, K., Lange-Bertalot, H., 1986. Bacillariophyceae. 1: Teil: Naviculaceae. In: Ettl, H., Gärtner, G., Gerloff, J., Heynig, H., Mollenhauer, D. (Eds.), *Süsswasserflora von Mitteleuropa*, Band 2/1. Gustav Fischer Verlag, Stuttgart, New York, pp. 876.
- Krammer, K., Lange-Bertalot, H., 1988. Bacillariophyceae. 2: Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. In: Ettl, H., Gärtner, G., Gerloff, J., Heynig, H., Mollenhauer, D. (Eds.), *Süsswasserflora von Mitteleuropa*, Band 2/2. Gustav Fischer Verlag, Stuttgart, New York, pp. 596.
- Krammer, K., Lange-Bertalot, H., 1991a. Bacillariophyceae. 3: Teil: Centrales, Fragilariaceae, Eunotiaceae. In: Ettl, H., Gärtner, G., Gerloff, J., Heynig, H., Mollenhauer, D. (Eds.), *Süsswasserflora von Mitteleuropa*, Band 2/3. Gustav Fischer Verlag, Stuttgart, Jena, pp. 576.
- Krammer, K., Lange-Bertalot, H., 1991b. Bacillariophyceae. 4: Teil: Achnanthaceae. In: Ettl, H., Gärtner, G., Gerloff, J., Heynig, H., Mollenhauer, D. (Eds.), *Süsswasserflora von Mitteleuropa*, Band 2/4. Gustav Fischer Verlag, Stuttgart, Jena, pp. 437.
- Laird, K.R., Cumming, B.F., Wunsam, S., Rusak, J.A., Oglesby, R.J., Fritz, S.C., Leavitt, P.R., 2003. Lake sediments record large-scale shifts in moisture regimes across the northern prairies of North America during the past two millennia. *PNAS* 100, 2483-2488.
- 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. *Limnology and Oceanography* 41, 890-902.

- Laird, K.R., Fritz, S.C., Maasch, K.A., Cumming, B.F., 1996b. Greater drought intensity and frequency before AD 1200 in the Northern Great Plains, USA. *Nature* 384, 552-554.
- Laird, K.R., Michels, A., Stuart, C.T.L., Wilson, S.E., Last, W.M., Cumming, B.F., 2007. Examination of diatom-based changes from a climatically sensitive prairie lake (Saskatchewan, Canada) at different temporal perspectives. *Quaternary Science Reviews* 26, 3328-3343.
- Lewin, J.C., 1961. The dissolution of silica from diatom walls. *Geochimica et Cosmochimica Acta* 21, 182-198.
- Lund, J.W.G., 1950. Studies on *Asterionella formosa* Hass. II. Nutrient depletion and the spring maximum. *Journal of Ecology* 38, 1-35.
- Mackay, A.W., Battarbee, R.W., Flower, R.J., Jewson, D.H., Ryves, D.B., Sturm, M., 2003. Assessing the potential for developing internal diatom-based transfer functions for Lake Baikal. *Limnology and Oceanography* 48.
- Mackay, A.W., Flower, R.J., Kuzmina, A.E., Granina, L.Z., Rose, N.L., Appleby, P.G., Boyle, J.F., Battarbee, R.W., 1998. Diatom succession & pollution trends in recent sediments from Lake Baikal and their relation to atmospheric pollution and to climate change. *Phil. Trans. R. Soc. Lond. B.* 353, 1011-1055.
- Mackay, A.W., Ryves, D.B., Battarbee, R.W., Flower, R.J., Jewson, D., Rioual, P., Sturm, M., 2005. 1000 years of climate variability in central Asia: assessing the evidence using Lake Baikal (Russia) diatom assemblages and the application of a diatom-inferred model of snow cover on the lake. *Global and Planetary Change* 46, 281-297.
- Patrick, R., Reimer, C., 1966. The Diatoms of the United States, Vol. 1: Fragilariaceae, Eunotiaceae, Achnanthaceae, Naviculaceae. The Academy of Sciences of Philadelphia, 688 pp.
- Patrick, R., Reimer, C., 1975. The Diatoms of the United States, Vol. 2: Entomoneidaceae, Cymbellaceae, Gomphonemaceae, and Epithemiaceae. The Academy of Sciences of

Philadelphia, 213 pp.

Pichon, J.-J., Bareille, G., Labracherie, M., Labeyrie, L.D., Baudrimont, A., Turon, J.-L., 1992.

Quantification of the biogenic silica dissolution in Southern Ocean sediments.

Quaternary Research 37, 361-378.

Racca, J.M.J., Prairie, Y.T., 2004. Apparent and real bias in numerical transfer functions in

palaeolimnology. Journal of Paleolimnology 31, 117-124.

Racca, J.M.J., Wild, M., Birks, H.J.B., Prairie, Y.T., 2003. Separating wheat from chaff:

Diatom taxon selection using an artificial neural network pruning algorithm. Journal of

Paleolimnology 29, 123-133.

Reed, J.M., 1998a. A diatom-conductivity transfer function for Spanish salt lakes. Journal of

Paleolimnology 19, 399-416.

Reed, J.M., 1998b. Diatom preservation in the recent sediment record of Spanish saline lakes:

implications for palaeoclimate study. Journal of Paleolimnology 19, 129-137.

Round, F.E., 1964. The diatom sequence in lake deposits, some problems of interpretation.

Verhandlungen Internationale Vereinigung für Theoretische und angewandte

Limnologie 15, 1012-1020.

Ryves, D.B., 1994. Diatom dissolution in saline lake sediments: an experimental study in the

Great Plains of North America. Unpublished PhD Thesis, University College London,

London, 306 pp.

Ryves, D.B., Battarbee, R.W., Juggins, S., Fritz, S.C., Anderson, N.J., 2006. Physical and

chemical predictors of diatom dissolution in freshwater and saline lake sediments in

North America and West Greenland. Limnology and Oceanography 51, 1355-1368.

Ryves, D.B., Jewson, D.H., Sturm, M., Battarbee, R.W., Flower, R.J., Mackay, A.W., Granin,

N., 2003. Quantitative and qualitative relationships between planktonic diatom

communities and diatom assemblages in sedimenting material and surface sediments in

Lake Baikal, Siberia. Limnology and Oceanography 48, 1643-1661.

- Ryves, D.B., Juggins, S., Fritz, S.C., Battarbee, R.W., 2001. Experimental diatom dissolution and the quantification of microfossil preservation in sediments. *Palaeogeography, Palaeoclimatology, Palaeoecology* 172, 99-113.
- Shemesh, A., Burckle, L.H., Froelich, P.N., 1989. Dissolution and preservation of Antarctic diatoms and the effect on the sediment thanatocoenoses. *Quaternary Research* 31, 288-308.
- Stoermer, E.F., Smol, J.P., 1999. *The Diatoms: Applications for the Environmental and Earth Sciences*. Cambridge University Press, Cambridge, 469 pp.
- Stone, J.R., Fritz, S.C., 2004. Three-dimensional modeling of lacustrine diatom habitat areas: Improving paleolimnological interpretation of planktic: benthic ratios. *Limnology and Oceanography* 49, 1540-1548.
- Straub, F., 1993. Diatoms and their preservation in the sediments of Lake Neuchatel (Switzerland) as evidence of past hydrological changes. *Hydrobiologia* 269, 167-178.
- Tarapchak, S.J., 1973. Studies on phytoplankton distribution and indicators of trophic state in Minnesota lakes. Unpublished PhD Thesis, University of Minnesota, Minneapolis, 204 pp.
- ter Braak, C.J.F., Juggins, S., 1993. Weighted averaging partial least squares regression (WAPLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia* 269/270, 485-502.
- ter Braak, C.J.F., Šmilauer, P., 2002. *Canoco for Windows*. Centre for Biometry, Wageningen.
- Tréguer, P., Nelson, D.M., Van Bennekom, A.J., DeMaster, D.J., Leynaert, A., Quéguiner, B., 1995. The silica balance in the world ocean: a reestimate. *Science* 268, 375-379.
- Van Cappellen, P., Dixit, S., van Beusekom, J., 2002. Biogenic silica dissolution in the oceans: Reconciling experimental and field-based dissolution rates. *Global Biogeochemical Cycles* 16, Article no. 1075.
- Vrieling, E.G., Sun, Q.Y., Tian, M., Kooyman, P.J., Gieskes, W.W.C., van Santen, R.A.,

- Sommerdijk, N., 2007. Salinity-dependent diatom biosilicification implies an important role of external ionic strength. *Proceedings of the National Academy of Sciences of the United States of America* 104, 10441-10446.
- Warnock, J., Scherer, R., Loubere, P., 2007. A quantitative assessment of diatom dissolution and late quaternary primary productivity in the Eastern Equatorial Pacific. *Deep-Sea Research Part II-Topical Studies in Oceanography* 54, 772-783.
- 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. *Canadian Journal of Fisheries and Aquatic Sciences* 53, 1580-1594.
- Winters, H.A., 1963. Geology and ground water resources of Stutsman County, North Dakota: Part 1 - Geology, North Dakota Geological Survey, Grand Forks, North Dakota.
- Yang, X.D., Kamenik, C., Schmidt, R., Wang, S.M., 2003. Diatom-based conductivity and water-level inference models from eastern Tibetan (Qinghai-Xizang) Plateau lakes. *Journal of Paleolimnology* 30, 1-19.

Figure captions

Figure 1. Dissolution stages for *Stephanodiscus niagarae* Ehrenberg (SEM micrographs of fossil material from Spiritwood Lake, ND). Stage 1 is the “pristine” state and higher stages are progressively more dissolved. **A.** Stage 1 showing the spines and complete valve and mantle. **B.** Stage 2 displays clear signs of dissolution with partial loss of the mantle and loss of finer ornamentation, e.g. spines. Note also the valve starting to fracture along striae, starting to enlarge under dissolution. **C.** Stage 3 is the highest stage recognised for this taxon. All spines are gone and the mantle is removed. Parts of the valve face may be missing. **D.** Details of a valve in stage 1 (A) – areolae are not noticeably dissolved. **E.** Detail of stage 3 – areolae are enlarged and starting to coalesce as dissolution progresses, creating lines of structural weakness along the valve.

Figure 2. Correlation between (A) **F** index and (B) **DDI** values for 27 NGP lakes independently counted by D.B. Ryves (DBR) and R.W. Battarbee (RWB). Dotted line is the 1: 1 correlation.

Figure 3. Salinity optima (weighted average, log TDS, g L⁻¹) of the most abundant taxa in the NGP dataset (ordered according to species optimum of all valves without reference to dissolution state). For each species, salinity optima for each dissolution stage (1 to 4, depending on species) are also plotted (dissolution pseudospecies). A = araphid valve ; R = raphid valve. Species mentioned in the text are indicated.

Figure 4. Species abundance (%) across the NGP dataset for *Cyclotella quillensis* Bailey (CY017A; see Fig. 3). **Upper plot:** Species distribution for all valves. **Lower plots:** Distributions for the 4 dissolution pseudospecies recognised for this species (CY017A1 = stage 1, CY017A2 = stage 2, CY017A3 = stage 3, CY017A4 = stage 4). **û** = weighted

average (WA) optimum, g L^{-1} TDS, indicated by the down arrow in each plot. Adding East Stump to the training set (empty circles) alters the WA optimum for stage 3 and 4, and hence optimum for all occurrences (shown in italics).

Figure 5. Species abundance (%) across the NGP dataset for *Navicula cincta* (Ehrenberg) Ralfs (NA021A; see Fig. 3). **Upper plot:** Species distribution for all valves. **Lower plots:** Distributions for the 3 dissolution pseudospecies recognised for this species (NA021A1 = stage 1, NA021A2 = stage 2, NA021A3 = stage 3). $\hat{\mu}$ = weighted average (WA) optimum, g L^{-1} TDS, indicated by the down arrow in each plot.

Figure 6. A. Plot of residuals ($\log \text{TDS, g L}^{-1}$) against predicted salinity for model D (Table 3) sized according to dissolution index (**DDI**). Sites mentioned in the text are labelled. Solid line is LOESS smoother added through data. Dotted line is LOESS smoother when residuals are plotted against observed salinity. **B.** Plot of residuals ($\log \text{TDS, g L}^{-1}$) against predicted salinity for model H2 (Table 3) with the outlier Eckelson Lake removed. Points are sized according to dissolution index (**DDI**). Sites mentioned in the text are labelled. LOESS smoother added through data.

Figure 7. Summary diatom diagram (%) for Devils Lake short core, North Dakota (after Fritz 1990). Hollow bars indicate proportion of valves dissolved for each species. Thin bars are x5 exaggeration. Also plotted are sample **F** index, measured lake level and observed salinity.

Figure 8. Summary diatom diagram (%) for Spiritwood Lake short core, North Dakota. Hollow bars indicate proportion of valves dissolved for each species. Thin bars are x5 exaggeration. Also plotted are sedimentation rate (flux and accumulation) and sample **F** index.

Figure 9. Diatom-inferred salinity for short cores from (A) Devils Lake and (B) Spiritwood Lake, using comparable 53 and 55 site NGP models both ignoring (models A and C, Table 3) and incorporating dissolution data (models F and G, Table 3). At Devils Lake, measured salinity values are shown by the dotted line. For Spiritwood Lake, shaded diamonds are measured mean summer salinity values (1965, 1970, 1972: data supplied by G.R. Van Eeckhout; 1970: from Tarapchak, 1973, conductivity converted to salinity; 1977-1988: values from M. Sauer, pers. comm.).

Appendix 1

Figures A1-A3. Where scale is not provided on figure, see caption for details of scale bar.

[NB All these figures, and others covering 20 other taxa, can be put in the Online Appendices (Multi Media Components). This material can be supplied for review if requested.]

Fig. A1. Dissolution stages of *Amphora libyca* (A-G)

A. Dissolution stage 1 (LM). Scale bar = 18 μm .

B. Dissolution stage 1 (LM) – same valve as A at different focal plane. Scale bar = 18 μm .

C. Dissolution stage 1 (SEM).

D. Dissolution stage 2 (SEM). Dissolution begins at valve apices.

E. Dissolution stage 2 (SEM) - detail of C.

F. Dissolution stage 3 (SEM). Valve dissolves towards raphe canal, which begins to dissolve away from apices.

G. Dissolution stage 4 (SEM). Raphe canal dissolves and/or breaks off, leaving a distinct central area, which is readily identifiable.

Fig. A2. Dissolution stages of *Cyclotella meneghiniana* (A-G)

A. Dissolution stage 1 (LM). Scale bar = 8 μm

B. Dissolution stage 1 (SEM). Detail shows structure of process (internal view).

C. Dissolution stage 2 (SEM), with the outer margin complete.

D. Dissolution stage 3 (SEM), with outer ring absent, but costae intact.

E. Dissolution stage 3 (SEM).

F. Dissolution stage 3 (LM). Scale bar = 4 μm .

G. Dissolution stage 4 (SEM). Only silica disc remains.

Fig. A3. Dissolution stages of *Navicula oblonga* (A-H). Similar patterns are observed with many other pennate taxa with a distinct central area.

- A.** Dissolution stage 1 (SEM).
 - B.** Dissolution stage 2 (SEM). Vela within striae dissolve.
 - C.** Dissolution stage 2 (SEM), detail of B.
 - D.** Dissolution stage 2 (SEM), with detail of dissolution of internal structures within striae.
 - E.** Dissolution stage 3 (SEM). Dissolution proceeds from the apices to the central area.
 - F.** Dissolution stage 4 (LM). Robust central area allows identification often to species for larger taxa. Scale bar = 10 μm .
 - G.** Dissolution stage 4 (SEM).
 - H.** dissolved end of valve distinct for *N. oblonga* but this often breaks up along raphe.
- Identification is thus preferably based upon central areas (F, G) (SEM).

Table 1

Table 1. Approaches to modifying training sets and fossil data according to dissolution information.

Type	Dissolution information	Approach	Fossil data	Test	Reference (NGP)	Model (Table 3)
Original	None	No change	No change	Control	Fritz (1990); Fritz et al. (1991, 1993); this paper	A – D
Indirect adjustment	Species’ relative resistance	Downweight species and/or samples	No change	WA _{class} (CANOCO)	Ryves (1994)	–
Direct adjustment	Dissolution index for species and samples (F, DDI)	Downweight species and/or samples	No change	WA _{class} (CANOCO)	–	–
	Dissolution index for species and samples (F, DDI)	Delete samples (and/or species)	No change	WA, WA-PLS, MAT, ML	This paper	E
	Dissolution index for species and samples (F, DDI)	Create many new dissolution taxa	Fossil dissolution data used	WA, WA-PLS, MAT, ML	This paper	F – I
	Dissolution index for species and samples and dissolution/population relationship (F)	Recalculate species’ abundances using dissolution index	Fossil dissolution data used	WA, WA-PLS, MAT, ML	This paper	J, K

Table 2. Dissolution stages for some major morphological groups

Stage 1 is the pristine condition in all cases.

Amphora (Appendix Fig. A1)**Stage 2** Striae at distal raphe ends dissolve and coalesce. Areolae enlarged throughout valve.**Stage 3** Valve dissolves from distal raphe ends, particularly along dorsal margin, which may quickly disappear.**Stage 4** Central hyaline area of thicker silica remains; raphe, if present, vestigial. No striae remain.*Anomoeoneis***Stage 2** Areolae/ocelli enlarged; valve walls noticeably thinner**Stage 3** Valve dissolves from margins to raphe, and from apices to centre. Valve may split along raphe canal. Identification is by central area.*Campylodiscus***Stage 2** Inter costal area corrodes at margin; central area persists.**Stage 3** Central area tends to dissolve before the marginal ribs, which are progressively exposed. Raphe and ribs remain, usually at expense of intercostal and central areas.*Cocconeis* – (araphid valve)**Stage 2** Areolae enlarged; marginal rim dissolved. Valve may split along apical axis.*Cocconeis* + (raphid valve)**Stage 2** Areolae enlarged; raphe often resists dissolution relative to the surrounding valve face. Valve may split along raphe.*Cyclotella* (Appendix Fig. A2)**Stage 2** Intercostal areas dissolve but marginal rim remains.**Stage 3** Costae remain but marginal rim dissolved.**Stage 4** Costae dissolved or vestigial; central area (with processes enlarged) may survive as a featureless disc.*Cymbella***Stage 2** Distal valve ends dissolve from both margins towards raphe.

Stage 3 Raphe and central area remain; striae dissolved but silica ribs separating striae remain. Both margins dissolved.

Naviculoid pennate - e.g. *Navicula oblonga*, *Mastogloia* spp. (Appendix Fig. A3)

Stage 2 Chambers of striae coalesce, especially at distal raphe ends. Valve outline maintained. Margins remain.

Stage 3 Valve dissolves from distal apices towards central area. Valve margins distinguishable towards centre.

Stage 4 Central area only remains, with vestigial raphe canal.

Nitzschia

Stage 2 Areolae coalesce; striae dissolve; often raphe (and keel punctae) only remain.

Stephanodiscus (Fig. 1)

Stage 2 Areolae enlarged and may coalesce; margins remain. Spines often corroded, irregularly preserved.

Stage 3 Valve dissolves from margin towards central area; margin dissolved. Spines disappear.

Surirella

Stage 2 Inter costal area corrodes at margin; central area and raphe persists.

Stage 3 Marginal raphe and central area remain; intercostal area dissolved.

Synedra

Stage 2 Striae chambers coalesce and dissolve at valve margins; valve apices distinct.

Stage 3 Valve dissolves towards central area from apices.

Table 3

Table 3. NGP model results under various dissolution-dependent strategies. Predictions are based on 1000 bootstrap cycles. Inclusion indicates the cut-off value below which species (as relative abundance, %) or samples (**F** scores) were deleted. RMSE and RMSEP values are log salinity units (g L⁻¹ TDS).

Code	Dataset	No. taxa	No. lakes	Inclusion	Best Apparent			Best predicted		
Unadjusted models		<i>j</i>	<i>i</i>	Value	Model	r ²	RMSE	Model	r ² _{boot}	RMSEP
A	Original (NGP53)	258	53	–	WAtol _{inv}	0.91	0.158	WA _{inv}	0.72	0.291
B	Original (NGP53)	148	53	1%	WAtol _{inv}	0.91	0.158	WA _{inv}	0.74	0.284
C	Original (NGP55)	258	55	–	WAtol _{inv}	0.89	0.178	WA _{inv}	0.70	0.309
D	Original (NGP55)	149	55	1%	WAtol _{inv}	0.89	0.178	WA _{inv}	0.72	0.302
Dissolution-adjusted models										
E	Original – deleted sites	255	49	F < 0.15	WAtol _{inv}	0.91	0.153	WA _{class}	0.71	0.294
F	Dissolution – binary	365	53	–	WAtol _{inv}	0.93	0.141	WA _{class}	0.73	0.286
G	Dissolution – binary	365	55	–	WAtol _{inv}	0.91	0.159	WA _{class}	0.71	0.305
H	Dissolution – all stages	536	53	–	WAtol _{inv}	0.93	0.135	WA _{class}	0.73	0.286
I	Dissolution – all stages	536	55	–	WAtol _{inv}	0.92	0.149	WA _{class}	0.72	0.302
J	F adjusted	148	53	1%	WAtol _{inv}	0.90	0.162	WA _{class}	0.75	0.276
K	F adjusted	148	52	1%	WAtol _{inv}	0.92	0.145	WA _{class}	0.77	0.261

Figure 2

Figure 2 (Ryves et al.)

Figure 3

Figure 3 (Ryves et al.)

Figure 4

Figure 4 (Ryves et al.)

Figure 5

Figure 5 (Ryves et al.)

A

B

Figure 7

Figure 7 (Ryves et al.)

Devils Lake, North Dakota

Figure 8

Figure 8 (Ryves et al.)

Spiritwood Lake, North Dakota

Figure 9 (Ryves et al.)

Supplementary Data: Appendix 1 (3 Figs)

[Click here to download Supplementary Data: Ryves_QSR_FigsA1-A3.ppt](#)