Hydrobiologia **178**: 209–223, 1989.

Inferring pH from diatoms: a comparison of old and new calibration methods

Cajo J. F. ter Braak[1,2] & Herman van Dam[2]
[1] *Agricultural Mathematics Group, Wageningen, Box 100, 6700 AC Wageningen, The Netherlands*;
[2] *Research Institute for Nature Management, Box 46, 3956 ZR Leersum, The Netherlands*

Received 5 August 1987; in revised form 7 January 1988; accepted 20 March 1988

Key words: calibration, diatoms, Gaussian model, indicator value, pH reconstruction, weighted averaging

Abstract

Two new methods for inferring pH from diatoms are presented. Both are based on the observation that the relationships between diatom taxa and pH are often unimodal. The first method is maximum likelihood calibration based on Gaussian logit response curves of taxa against pH. The second is weighted averaging. In a lake with a particular pH, taxa with an optimum close to the lake pH will be most abundant, so an intuitively reasonable estimate of the lake pH is to take a weighted average of the pH optima of the species present.

Optima and tolerances of diatom taxa were estimated from contemporary pH and proportional diatom counts in littoral zone samples from 97 pristine soft water lakes and pools in Western Europe. The optima showed a strong relation with Hustedt's pH preference groups. The two new methods were then compared with existing calibration methods on the basis of differences between inferred and observed pH in a test set of 62 additional samples taken between 1918 and 1983. The methods were ranked in order of performance as follows (between brackets the standard error of inferred pH in pH units); maximum likelihood (0.63) > weighted averaging (0.71) = multiple regression using pH groups (0.71) = the Gasse & Tekaia method (0.71) > Renberg & Hellberg's Index B (0.83) ≫ multiple regression using taxa (2.2). The standard errors are larger than those usually obtained from surface sediment samples. The relatively large standard may be due to seasonal variation and to the effects of other factors such as humus content. The maximum likelihood method is statistically rigorous and can in principle be extended to allow for additional environmental factors. It is computer intensive however. The weighted averaging approach is a good approximation to the maximum likelihood method and is recommended as a practical and robust alternative.

Introduction

Diatom assemblages preserved in lake sediments contain a record of past water pH that is valuable in studying the history of acid rain (Battarbee, 1984; Battarbee *et al.*, 1986; Battarbee & Charles, 1986; Davis, 1987). In water bodies where sediments are absent or often disturbed, e.g. fast flowing rivers or shallow pools, pH-decreases have been observed by comparison of diatom assemblages in old and recent periphyton and plankton tow samples (Berge, 1976; Van Dam & Kooyman-van Blokland, 1978; Van Dam *et al.*, 1981). The Index B method (Renberg & Hellberg, 1982) has proven to be a powerful tool to infer pH. However, like several of its predecessors, this

index method is only a heuristic formula to infer pH quantitatively from the pH preference groups defined by Hustedt (1939); the method does not make full use of contemporary data on the pH distributions of individual taxa. Furthermore, the method does not work well below pH 4.5, when only acidobiontic and acidophilous diatoms are present, e.g. in the extremely acid moorland pools of Northwestern Europe (Van Dam *et al.*, 1981; Van Dam & Beljaars, 1981). More recent approaches using multiple regression of pH on diatom taxa (Davis & Anderson, 1985) seem to assume linear relationships between pH and diatom abundance (cf. Howe & Webb, 1983), an assumption that is inconsistent with the observation that diatom taxa have a preferred pH value or pH optimum. An additional problem is that the estimates of the coefficients of the regression equation are often instable due to high correlations among the abundances of ecologically similar diatom taxa (Montgomery & Peck, 1982: 291). Thus neither the index B nor the multiple regression approach seems fully satisfactory on theoretical grounds.

It has long been observed (Shelford, 1911; Hesse, 1924; Gause, 1930; Thienemann, 1932) that the relationship between a species and an environmental variable is often unimodal, i.e. each species thrives best at a particular value (optimum) and cannot survive when the value is either too low or too high (cf. Sheldford's 'law' of tolerance in Odum, 1959). Our new approach to calibration originates from this ecologically realistic idea of how diatom taxa respond to pH. For simplicity we work with a class of symmetric unimodal curves, namely, the Gaussian response model (Gauch & Whittaker, 1972; Stoermer & Ladewski, 1976) in which each taxon has a pH optimum and a pH tolerance, and we adapt the model to make it suitable for proportional diatom counts.

Inferring pH from diatom taxa is a statistical calibration problem (Ter Braak, 1987a). We use the maximum likelihood approach to calibration. First, the relationship between the abundance of a diatom taxon and pH is specified by a response curve, such as Fig. 1, fitted to present-day data on diatom abundances and pH by non-linear regression. Together the fitted curves for all of the taxa determine what diatom composition is expected at any given pH value. Second, we determine the pH value that is most likely to give rise to the observed diatom composition. It depends on the amount of available data how flexible the class of curves can be that can still usefully be fitted: lack-of-fit must be balanced against estimation variance. In other words, the class of curves should be a compromise between ecological realism and simplicity. We consider the class of Gaussian response curves to be a good compromise.

The maximum likelihood approach with the Gaussian model is computer intensive, but can be approximated, fortunately, by the much simpler method of weighted averaging (Ter Braak, 1987a). Its intuitive rationale is as follows. In a lake with a particular pH, taxa with an optimum close to the lake pH will be most abundant, so a simple estimate of the lake pH is to take a (weighted) average of the pH optima of the taxa present. Taxa with a small pH tolerance (good pH indicators) can be given more weight in the weighted average than taxa with a large pH tolerance; indifferent taxa are given zero weight.

Calibration by weighted averaging is not new. Ellenberg (1948) proposed it to infer pH and other environmental variables from vascular plants and Pantle & Buck (1955) to infer organic pollution (saprobity) in fresh waters from bacteria and algae, including diatoms. The optima are termed 'Zeigerwerte' (indicator values) by Ellenberg (1979) and saprobity indices by Sládeček (1986).

In this paper pH optima and pH tolerances of diatom taxa are estimated from a training set of 99 littoral zone samples taken in 1982 and 1983. The new calibration methods are compared with existing methods; (a) theoretically, (b) on the basis of differences between inferred and observed pH in the training set, and (c) in a test set comprising 62 additional samples taken between 1918 and 1983. The prediction error of pH in the test set includes all error components (lack-of-fit, sampling variation, etc.) and is thus an appropriate benchmark to compare methods.

Data collection

The 99 samples for the training set were taken in 1982 and 1983 from 97 approximately pristine soft water lakes and pools in Denmark, Western Germany, Belgium and The Netherlands (Van Dam & Beljaars, 1984). Selected physical and chemical variables are given in Table 1. Periphytic, bottom-living or plankton-tow (mesh width 40 μm) diatoms were collected in the littoral zone at each site in the same way as has been done by previous investigators at the same site (see next paragraph). When old samples were absent, usually plankton tow diatoms were gathered. The pH was measured simultaneously on the spot with a Methrohm E 488 pH meter or a WTW 91 pH meter. Electrodes were calibrated with buffer solutions of pH 7 and pH 4, regularly checked with a buffer solution of pH 3 and discarded when the reading differed by more than 0.1 unit from 3.0.

Samples for the test set included a few recent samples (after 1970) with accompanying pH readings, obtained as described above. The majority of the test set samples were obtained from plankton tow and periphyton collections by W. Adam, N. Foged, J. Heimans, F. Hustedt, E.G. Jørgensen, and P. van Oije between 1918 and 1955, at a subset of the localities of the training set. Most of the pH readings were done with colorimetric methods simultaneously with diatom sampling. When such pH measurements were lacking they were taken from published or unpublished reports, if the time span between diatom sampling and measurement of pH was less than five years.

Table 1. Selected physical and chemical variables of the 99 samples from the training set (DOC = dissolved organic carbon).

	Minimum	Median	Maximum
Area (ha)	0.002	0.63	15.9
Depth (m)	0.1	1.0	52.0
Altitude (m)	4	40	1109
pH	3.3	4.2	7.3
SO_4^{2-} (meq l^{-1})	0.04	0.26	2.33
DOC (mg l^{-1})	0.04	9.7	55

The diatom samples were boiled gently for one hour in hydrogen peroxide. Before boiling plant leaves and stems of the periphyton were cut in fragments (c. 1 cm long) and mixed. If necessary, iron was removed with hydrochloric acid. After boiling, cooling and settlement the diatom valves were washed three times with demineralized water and embedded in Clearax mount. In each slide 400 valves were identified and counted with the help of a Zeiss research microscope (magnification × 1250). The identification literature is listed by Van Dam (1984). Most of the samples were counted by Mr. C. N. Beljaars. A minor fraction of the samples was counted by either Mr. G. Suurmond or the second author, who supervised the counting procedure to obtain a consistent taxonomy (Davis & Smol, 1986). In case of doubt slides were sent for consultation to experts in the taxonomy of particular taxa.

Methods

Notation

Let x denote the variable to be calibrated (here always pH), x_i the value of x in sample i and y_{ik} the abundance of taxon k in sample i ($y_{ik} \geqslant 0$). We reserve the subscript i for samples (i = 1,...,n) and the subscript k for taxa (k = 1,...,m). Here, as in much diatom research, the abundance is the fraction of valves of taxon k with respect to all (here 400) valves examined in sample i. Thus our data are *compositional data*, which have a constant-sum constraint ($\Sigma_{k=1}^{m}\ y_{ik} = 1$). The estimator of x_i will be denoted by $\tilde{x}_i$.

Maximum likelihood approach using Gaussian models

Maximum likelihood calibration requires the pH response curves of taxa. These curves are not yet known, so they must be estimated by regression from contemporary data on taxon abundances and pH, using a class of response curves that combines reasonable ecological realism with tractability.

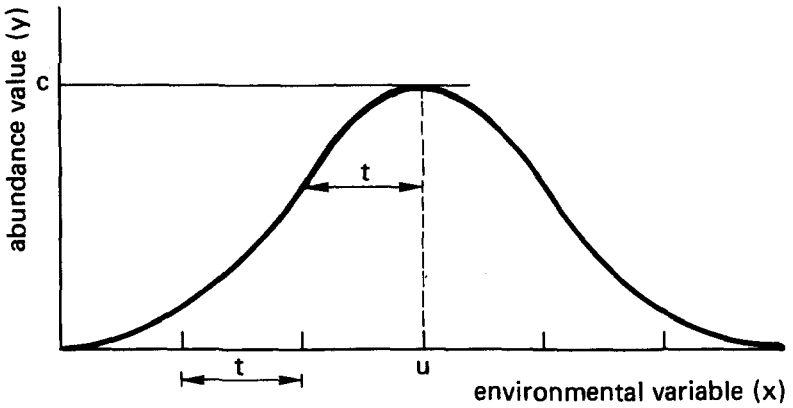

Fig. 1. The Gaussian response curve for the abundance value (y) of a taxon against an environmental variable (x) (u = optimum or mode; t = tolerance; c = maximum).

One suitable class of response curves is provided by the Gaussian model (Fig. 1; Gauch & Whittaker, 1972):

$$f_k(x) = c_k\, e^{-1/2(x-u_k)^2/t_k^2} \qquad (1)$$

where $f_k(x)$ is the expected value of the abundance of taxon k as a function of the calibration variable x, and u_k, t_k and c_k are taxon parameters – u_k is the optimum (the value of x at the peak), t_k the tolerance (a measure of curve width) and c_k the height of the peak. This model can not be exact for compositional data because of the constraint that the abundances of all taxa sum to 1. If we take Eq. (1) as a model for absolute abundances then we can easily obtain a model for proportions; we just divide the absolute abundance (for any given x) by their sum. In this way we obtain a multinomial logit model (e.g. McCullagh & Nelder, 1983: 106) of the form:

$$y_k(x) = \frac{c_k\, e^{-1/2(x-u_k)^2/t_k^2}}{\sum_{l=1}^{m} c_l\, e^{-1/2(x-u_l)^2/t_l^2}} \qquad (2)$$

where $y_k(x)$ is the expected proportional abundance of taxon k as a function of x. However, the parameters of the taxa in (2) are more difficult to understand than those of the Gaussian model (1), because there is indeterminacy in the parameters (Ter Braak, 1987c: 68). Because of this difficulty, we also fitted the Gaussian logit model (Ter Braak & Looman, 1986)

$$y_k(x) = \frac{c_k\, e^{-1/2(x-u_k)^2/t_k^2}}{1 + c_k\, e^{-1/2(x-u_k)^2/t_k^2}} \qquad (3)$$

which acknowledges that each of the fractions lies between 0 and 1, but does not guarantee that their total is precisely 1 for each sample. Ter Braak & Looman (1986) used the Gaussian logit model for presence-absence data, but it can also be used as quasi-likelihood model (McCullagh & Nelder, 1983) for proportions.

The multinomial logit model (2) and the Gaussian logit model (3) were fitted to the training data set by loglinear regression (with multinomial error structure, see McCullagh & Nelder 1983: 106, 142) and logit regression (with binomial error structure), respectively, by using GENSTAT (Alvey *et al.*, 1982) as explained in Ter Braak & Looman (1986). This was done for all 26 taxa that occurred in at least 10 samples of the training data set. With the multinomial logit model two runs were made, one in which the remaining taxa were considered as one extra taxon and a second in which the abundances of each of the 26 taxa were expressed as a fraction of the total number of valves of the 26 taxa in each sample. We present the results of the former analysis only, as it turned out to give a smaller prediction error. The

Gaussian logit is easier to fit than the multinomial logit model because the fit can proceed for each species in turn, whereas in the multinomial logit model all species must be fitted simultaneously. In the multinomial logit model the hypothesis of equal tolerances was tested by a quasi-likelihood F-test (Jørgensen, 1983) at $\alpha = 0.05$. For each taxon the Gaussian logit model was tested against the (simpler) linear logit model to determine whether the optimum was significant (again by a quasi-likelihood F-test, $\alpha = 0.05$; Ter Braak & Looman, 1986). If the optimum of a taxon was not significant, the linear logit model (Ter Braak & Looman, 1986) was tested against the null model that a taxon showed no relationship with x ($\alpha = 0.05$). In this way the simplest acceptable curve was determined for each taxon and this curve was used in the calibration.

The response curves, together with the assumed error structure, form a statistical model of diatom composition in relation to pH. From the model we can calculate the probability that a particular diatom composition will be observed at a given pH value. By determining this probability for a range of pH values, we can select the value that gives the highest probability of observing the diatom composition. This pH value is the maximum likelihood estimate.

Weighted averaging approach

A taxon with a particular pH optimum will be most abundant in lakes with pH values close to its optimum. So an intuitively reasonable estimate of the optimum is to take a weighted average of the pH values of lakes in which the taxon is present. The weighted averaging estimate of the optimum u_k is

$$\hat{u}_k = \frac{\sum_{i=1}^{n} y_{ik} x_i}{\sum_{i=1}^{n} y_{ik}} \qquad (4)$$

Ter Braak (1985) proved that the non-linear regression estimator of the optimum of the Gaussian (logit) model is the same as the weighted average if the distribution of x (pH) is uniform over the whole range of occurrence of the species. Ter Braak & Looman (1986) compared the two estimators by simulation for various cases in which the distribution of x was not uniform; they found the estimators about equally efficient for rare species with narrow ecological amplitude, but warned that in other cases the weighted average could be biased. Despite this limitation, the weighted average is a popular statistic for characterizing the environmental preferences of diatoms (e.g. Salden 1978; Charles, 1985).

Once the optima have been estimated, they can be used to infer the lake pH from its diatom composition. In a lake with a particular pH, taxa with an optimum close to the lake pH will be most abundant. So an intuitively reasonable estimate of the lake pH is the weighted average

$$\tilde{x}_i = \frac{\sum_{k=1}^{m} y_{ik} u_k}{\sum_{k=1}^{m} y_{ik}} \qquad (5)$$

Ter Braak (1985) proved that the maximum likelihood estimate of x_i in the Gaussian (logit) model boils down to the simple weighted average if the taxa conform to a species packing model (Whittaker *et al.*, 1973). If the pH tolerance varies strongly among taxa, one can weight taxa in Eq. (5) inversely with their squared tolerance (Ter Braak & Barendregt, 1986)

$$\tilde{x}_i = \frac{\sum_{k=1}^{m} y_{ik} u_k / t_k^2}{\sum_{k=1}^{m} y_{ik} / t_k^2} \qquad (6)$$

as also suggested by Zelinka & Marvan (1961) and Goff and Cottam (1967). (This weighting is done implicitly in the maximum likelihood approach.) In similar vein Sládeček (1986) lists 'saprobic indices' and 'indicative weights' which take the place of the optima u_k and $1/t_k^2$, respectively, in inferring saprobity from diatom assemblages.

A simple weighted averaging estimate of t_k is

$$\hat{t}_k = \{\textstyle\sum_i y_{ik}(x_i - \hat{u}_k)^2 / \sum_i y_{ik}\}^{1/2} . \qquad (7)$$

Multiple regression approach

In the multiple regression approach (e.g. Charles, 1985; Davis & Anderson, 1985; Flower, 1986),

pH is inferred by a linear function of the abundances of taxa or of pH preference groups, the coefficients of which are estimated by multiple regression of pH on the abundances. The method is known in the statistical literature as inverse regression; see Brown (1982) for a review. Inverse regression has been extensively used in climatic calibration from pollen data (Webb & Clark, 1977; Bartlein & Webb, 1985).

Three multiple regressions were calculated for the training set. These were based on (1) the fractional abundance of Hustedt's (1939) pH groups, (2) the abundance of the 26 most common taxa, (3) a step-wise multiple regression (forward selection out of the 26 most common taxa until no significant improvement ($P = 0.05$) could be made).

This multiple regression approach superficially seems to assume linear relationships between diatom abundances and pH. However, there is an interesting link with the weighted averaging approach. For compositional data, the weighted averaging estimate reduces to a linear combination of abundances, because the denominator in Eq. (5) is constant for such data. The multiple regression estimate is a linear combination of abundances also, but commonly includes an intercept. For compositional data, a regression model with an intercept can be rewritten as a model without an intercept but with one extra taxon in which the abundances of all not-yet-included taxa are summed, i.e.

$$\tilde{x}_i = b_O + \Sigma_{k=1}^{m-1} b_k y_{ik} = \Sigma_{k=1}^{m} a_k y_{ik} \qquad (8)$$

with $a_k = b_O + b_k$ $(k = 1,\ldots,m-1)$, $a_m = b_O$ and $y_{im} = 1 - \Sigma_{k=1}^{m-1} y_{ik}$.
The coefficients a_k play the same rôle as the optima u_k in Eq. (5). The methods thus use a similar type of calibration function; they differ only in the way in which its coefficients are estimated. In practice the estimated regression coefficients can however be corrupted by high correlations or multicollinearity among species (Montgomery & Peck, 1982: 291).

Approach of Gasse & Tekaia (1983)
We also considered the approach of Gasse and Tekaia (1983). Here pH is first divided into classes and then a correspondence analysis is applied to a taxon-by-class data table, each cell of which contains the total abundance of a taxon in samples with a pH that falls in the corresponding class. We used four classes, with the pH values 4, 5 and 6 as cutpoints. The final calibration function is obtained by a multiple regression of pH on the axes of the correspondence analysis.

Despite its complexity, this method is closely related to weighted averaging. The first axis' scores assigned by correspondence analysis to the pH classes can be interpreted as an optimal quantification of pH (Gifi, 1981). Both weighted averaging (with equal tolerances) and the Gasse & Tekaia method are special cases of canonical correspondence analysis (Ter Braak, 1986, 1987b,c).

Approach of Renberg & Hellberg (1982); extra regressions
The log Index B approach of Renberg & Hellberg (1982) was used both with its original regression coefficients as with coefficients newly calculated from the training set by simple linear regression. In the weighted averaging approach, averages are taken twice, so that the range of the inferred pH is reduced. To undo this, a simple linear regression was calculated of pH on $\tilde{x}_i$ (Eqs. 5–6) in the training set. Such an extra regression was also calculated for the maximum likelihood approach to make the methods more comparable. The extra regression minimizes the standard error (SE) of prediction in the training set as is done in the multiple regression approach and the approaches of Gasse & Tekaia (1983) and Renberg & Hellberg (1982).

Results

The distribution of pH in the training set is shown in the bottom line of Fig. 2; the mean and standard deviation of pH were 4.5 ± 0.88 in the training set and 5.2 ± 1.03 in the test set. In the

Table 2. pH-optima ($\tilde{u}_k$) and tolerances ($\tilde{t}_k$) for the most common taxa estimated from the training set. – = $\tilde{u}_k$ and $\tilde{t}_k$ poorly determined, because the taxon has only weak relationships with pH or has its optimum outside the sampled pH-range (3.3–7.3), n = number of samples in training set where taxon is present, c = type of curve (u: unimodal, + : increasing, –: decreasing, o: no significant relation), ML = maximum likelihood, WA = weighted averaging, MR = multiple regression (Eq. 8). See Table 5 for abbreviations of pH-groups.

Taxon			ML		WA		WA Charles (1985)	MR		Hustedt pH-group
	n	c	$\tilde{u}_k$	$\tilde{t}_k$	$\tilde{u}_k$	$\tilde{t}_k$	$\tilde{u}_k$	26 spp. a_k	7 spp. a_k	
Achnanthes minutissima Kütz.	20	+	–	–	6.8	0.6	7.2	8.6	8.8	cir
Brachysira vitrea f. *lanceolata* (Mayer) v. Dam	24	u	6.5	0.7	5.9	0.7	–	8.1	7.9	cir
B. brebissonii Ross	12	u	5.3	0.9	4.8	0.7	5.9	5.4	–	acp
Eunotia bilunaris (Ehr.) Nörpel[a]	66	o	–	–	4.3	0.8	5.0	4.5	–	cir
E. bilunaris var. *excisa* nov. comb.[b]	17	–	–	–	3.8	0.5	–	3.2	–	acp
E. exigua (Bréb.) Rabh.	84	–	1.6	1.7	4.1	0.6	5.2	4.0	–	acb
E. incisa Ehr.	49	u	5.3	0.6	5.0	0.6	5.0	5.1	5.5	acp
E. naegelii Mig.	12	u	4.2	0.5	4.2	0.4	4.8	4.1	–	acp
E. paludosa Grun.	25	u	3.8	0.4	3.8	0.3	–	3.7	–	acb
E. rhomboidea Hust. (asymmetric valves)	50	u	5.1	0.6	4.9	0.5	–	4.8	–	acp
E. rhomboidea Hust. (symmetric valves)	40	u	5.5	0.8	5.0	0.7	5.1[c]	3.8	–	acp
Fragilaria construens var. *venter* (Ehr.) Grun.	14	u	5.7	0.4	5.6	0.4	7.3	6.9	6.9	alp
F. virescens Ralfs	18	u	6.0	0.6	5.6	0.6	6.8	6.7	8.2	cir
Frustulia rhomboides (Ehr.) De Toni	16	o	–	–	4.7	0.9	5.5	12.5	–	acp
F. rhomboides var. saxonica (Rabh.) De Toni	77	–	4.2	1.0	4.2	0.6	5.2	4.0	–	acb
Gomphonema parvulum (Kütz.) Grun.	13	u	5.9	1.0	5.1	0.8	–	0.2	–	cir
Navicula leptostriata E. Jørgensen[d]	10	u	6.5	0.7	5.9	0.7	6.0	42.7	58.3	acp
N. mediocris Krasske	22	o	–	–	4.2	0.9	5.4	−0.3	–	acp
N. subtilissima Cleve	19	–	2.9	1.3	4.0	0.5	5.1	3.7	–	acb
Pinnularia abaujensis (Plant.) Ross	11	u	4.9	0.5	4.7	0.5	4.5	6.0	–	cir
P. biceps Greg.	18	u	4.8	0.6	4.6	0.5	5.6	6.9	–	cir
P. microstauron (Ehr.) Cleve	15	u	5.7	0.9	5.0	0.8	5.2	30.8	–	cir
P. irrorata (Grun.) Hust.	10	u	5.2	0.7	4.9	0.6	–	−1.0	–	acp
P. subcapitata Greg.	12	–	–	–	3.5	0.3	–	2.4	2.3	acp
Tabellaria flocculosa (Roth) Kütz.	45	u	5.4	0.7	5.0	0.6	6.3	4.8	–	acp
T. quadriseptata Knudson	35	o	5.7	1.9	4.6	0.9	5.7	3.6	–	acb
Rest category	–	–	–	–	–	–	–	4.7	4.1	–

[a] Synonym: *E. lunaris* (Ehr.) Grun. [b] syn.: *E. lunaris* var. *excisa* Grun., [c] as *E. tenella* (Grun.) Cleve, [d] syn.: *N. heimansii* v. Dam et Kooyman.

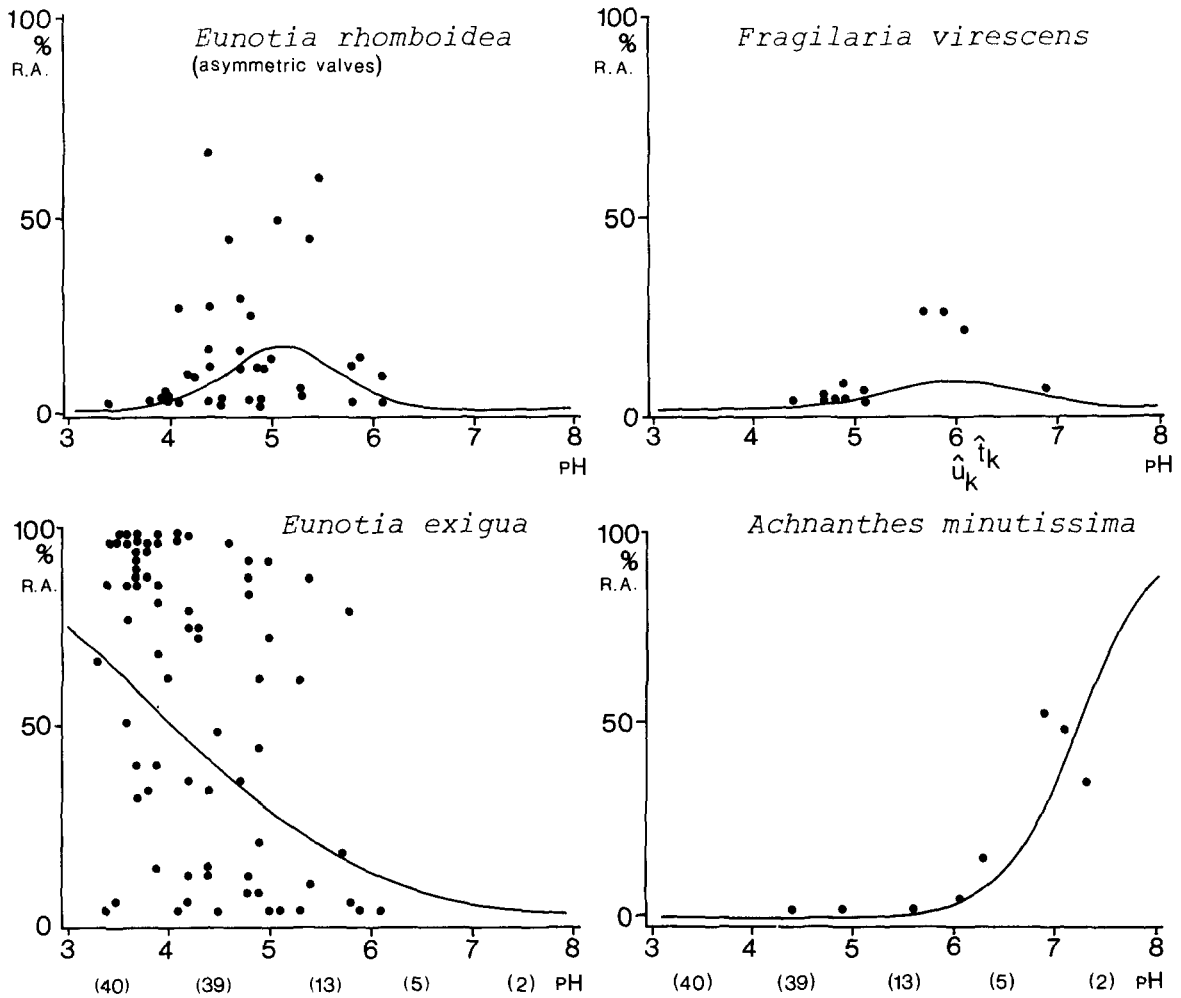

Fig. 2. Relative abundance (R.A.) of selected diatom taxa in 99 samples of the training set against observed pH. Zero observations are not entered. The number of observation in each pH interval is indicated between parentheses at the bottom. Curves are fitted by Gaussian logit regression.

Gaussian logit regressions 16 of the 26 taxa had a significant optimum, five taxa showed a decreasing response curve, one taxon an increasing curve and four species were unrelated to pH (Table 2).

The example curves in Fig. 2 show considerable scatter around the fitted curves. For taxa with a significant optimum, the optima as estimated by weighted averaging are up to 0.6 pH units lower than those estimated by Gaussian logit regression. This downward bias can be explained by the overrepresentation of acid samples (bottom line Fig. 2) (Ter Braak & Looman, 1986). The optima are strongly related to the assignments of taxa to Hustedt's (1939) pH preference groups (Table 2).

The estimated tolerances are nearly all over 0.5 pH units. This means that for nearly all taxa the range of occurrence is greater than 2 pH units. The tolerances as estimated by Gaussian logit regression and weighted averaging are remarkably close for taxa with a significant optimum. For the other taxa, both methods give qualitatively similar results: taxa with a decreasing (increasing) response curve have a low (high) weighted averaging estimate of the optimum and the taxa that are not related to pH have a large weighted averaging estimate of the tolerance. ($\geqslant 0.8$).

If the coefficients of the multiple regression on 26 taxa are interpreted as optima (Eq. 8), three

taxa have an absurdly high coefficient, and two taxa are estimated to have their optimum at a negative pH value; yet for eight taxa the coefficient is within 0.2 pH unit of the optimum as estimated by weighted averaging. Most of the coefficients obtained with step-wise regression cannot be considered optima; the coefficient of *Navicula leptostriata* (selected as the third taxon) is nonsensical, and was already so at its entry in the model.

The correlations between observed and inferred pH ranged from 0.65 to 0.86 in the training set and 0.33 to 0.79 in the test set. The corresponding SE's of prediction are shown in Table 3. In the training set, multiple regression on taxa had the lowest SE closely followed by weighted averaging using taxa. In the test set, however, multiple regression on taxa performed by far the worst due to several outliers in the abundance values. Although the outliers disappeared when the multiple regression was carried out on logarithmically transformed abundances (SE = 0.43 in training set), multiple regression still performed worst in the test set (SE = 0.90 with and 0.83 without selection). In the test set the variants of maximum likelihood performed best (Fig. 3c), followed by weighted averaging using taxa, multiple regression on pH groups and the Gasse & Takaia method. The Renberg & Hellberg (1982) approach (Fig. 3b, d) gave higher SE's and could not be used for 13 samples in the training set and for one sample in the test set where only acidobiontic and acidophilous taxa were present.

In the multinomial logit model the hypothesis of equal tolerances could be rejected ($P \leqslant 0.01$). Despite this, the model with unequal tolerances, did not noticebly improve the prediction of pH compared to the model with equal tolerances. Similarly, the tolerance-weighted version of weighted averaging (Eq. 6) appeared to have no advantage over straight weighted averaging (Eq. 5).

The extra regression carried out after calibration by maximum likelihood and weighted averaging (Table 4) improved the SE's (Table 3); its effect is to shrink the values inferred by maximum likelihood and to expand the values inferred by weighted averaging. The regression coefficients for use with Index B differed considerably from those given by Renberg & Hellberg (1982). The weighted averaging estimates of the optima of Hustedt's pH groups are expanded in Table 5 by using the extra regression (Table 4). After expansion, the estimates are remarkably close to those

Table 3. Standard error (SE) of inferred pH for calibration methods without (I) and with (II) an extra regression

Method	Training set		Test set	
	I	II	I	II
Maximum likelihood				
– Gaussian logit	0.66	0.57	0.69	0.63
– Multinomial logit	0.66	0.58	0.75	0.67
– Multinomial logit ($t_k = t$)	0.65	0.55	0.75	0.70
Weighted averaging				
– 26 taxa (Eq. 5)	0.52	0.49	0.74	0.71
– 26 taxa (Eq. 6)	0.51	0.47	0.77	0.74
– pH groups (Eq. 5)	0.68	0.62	0.89	0.75
Multiple regression				
– 26 taxa	–	0.45	–	2.24
– 7 step-wise selected taxa	–	0.51	–	2.74
– pH groups	–	0.61	–	0.71
Renberg & Hellberg (1982)*	0.78	0.68	0.88	0.83
Gasse & Tekaia (1983)	–	0.49	–	0.71

* For 86 of the 99 samples in training set and 61 of the 62 samples in test set.

Table 4. Intercepts (a) and regression coefficients (b) of extra regression in calibration methods (r = correlation coefficient).

Method	a	b	r
Maximum likelihood			
– Gaussian logit	1.47	0.68	0.76
– Multinomial logit	1.43	0.68	0.75
– Multinomial logit ($t_k = t$)	1.48	0.66	0.78
Weighted averaging			
– 26 taxa (Eq. 5)	– 1.49	1.34	0.83
– 26 taxa (Eq. 6)	– 1.32	1.31	0.84
– pH groups (Eq. 5)	– 3.39	1.76	0.71
Renberg & Hellberg (1982)*	5.85	– 0.54	0.65

* Regression of pH on ^{10}log Index B.

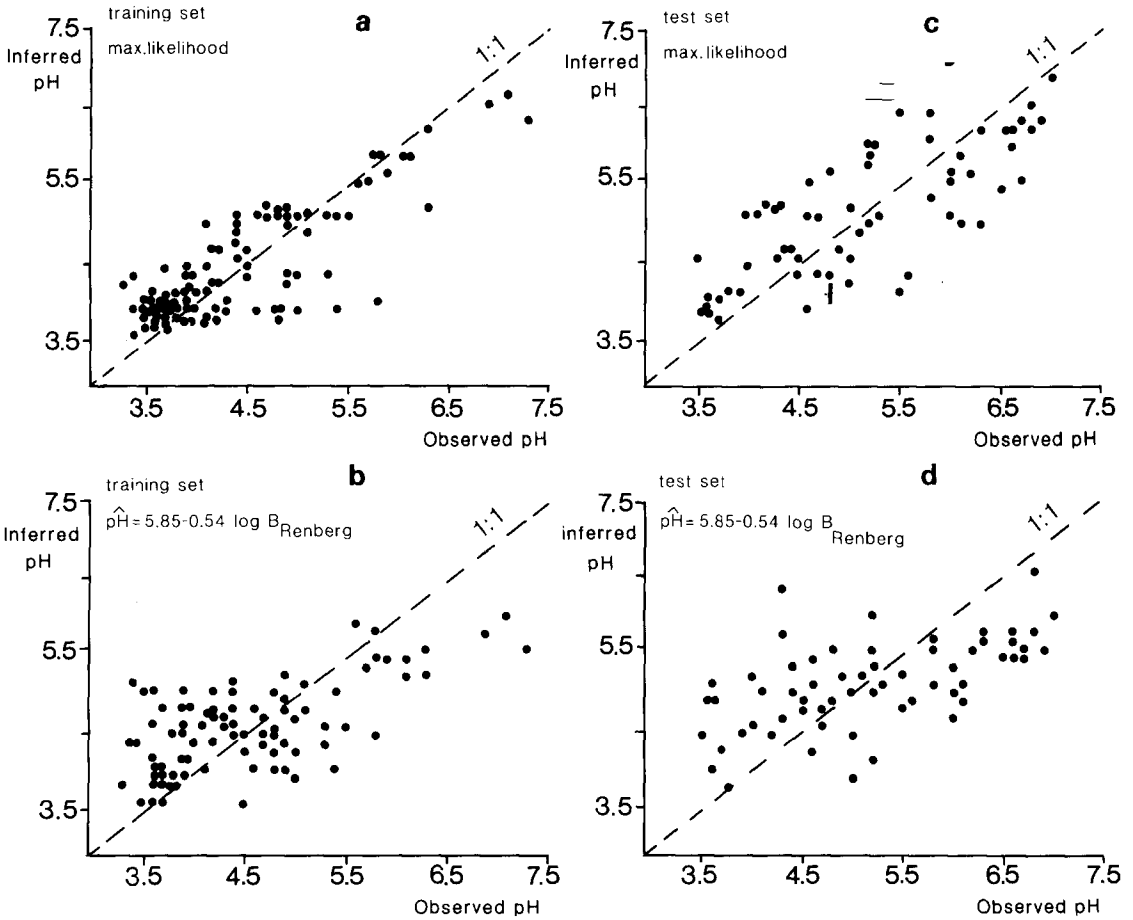

Fig. 3. Inferred pH versus observed pH in training set (a, b) and test set (c, d) for the maximum likelihood method based on Gaussian logit curves (a, c) and the Renberg & Hellberg (1982) method (b, d).

by multiple regression on pH groups; this explains why the SE's of weighted averaging and multiple regression using pH groups are about equal (Table 3).

Discussion

The standard error in the training set was much lower than that in the test set. The training set thus gave an overoptimistic impression of the prediction error encountered in practice, especially when multiple regression on taxa was used (Table 3). The standard errors from surficial sediment assemblages reported by Battarbee (1984), Steinberg *et al.* (1984), Charles (1985), Davis & Anderson (1985), Dixit (1986), Dixit & Evans (1986) and Flower (1986) are in the range of 0.3–0.5 pH units and, also being estimated from training sets, presumably underestimate the real prediction error. The standard error in our training set of littoral zone samples is relatively large (c. 0.6). There are several possible reasons:

1. Our data set is heterogeneous in physical and chemical variables (Table 1) and contains different types of diatom samples. Heterogeneity increases the prediction error because pH is not the only factor influencing diatom abundance. The error found by Gasse & Tekaia (1983) in a large (n = 160) and heterogeneous data set (including hard waters) is of the same order of magnitude as ours. Taylor *et al.* (1986) obtained an even larger standard error (c. 0.7).

2. Diatom assemblages and pH values in water bodies have spatial and temporal variability (Jones & Flower, 1986). These variabilities are not covered by litoral zone sampling and our single measurements of pH. Surficial sediment samples contain the valves of diatoms which have been accumulated through a number of years. Multiple pH-readings are therefore even more useful for surficial sediment samples than for littoral zone samples. Using surficial sediment samples, Davis & Anderson (1985) found smaller standard errors in Norwegian lakes, where multiple pH readings were available, than in New England lakes, where only single pH readings were done.
3. Our data set has a large proportion of very acid water bodies, where only a very small number of taxa is able to live.

The standard errors of diatom-inferred pH-values in the test set are presumably overpessimistic because of imprecise pH measurements in the test set:

4. The pH measurements were taken from different authors, sometimes the measurements were done in the field, sometimes in the laboratory; pH measurements were not always done simultaneously with diatom sampling; and sometimes pH measurement was done at other sites in the lakes than the diatom sampling.
5. Most pH measurements in the test set were performed with colorimetric methods, which can give errors of about one pH unit in weakly buffered low alkalinity waters (Haines *et al.*, 1983; Blakar & Digernes, 1984).

Multiple regression on pH groups performed quite well in comparison with the other methods in our study. To obtain more precise estimates of the regression coefficients, the multiple regression was also applied to the training and test set com-

Table 5. Coefficients a_k of calibration function (Eq. 8) using Hustedt's pH groups estimated by weighted averaging and multiple regression for training set and literature data (acb = acidobiontic; acp = acidophilous; cir = circumneutral; alp = alkaliphilous; n = number of samples; SE standard error of prediction).

	acb	acp	cir	alp[1]	n	SE
Weighted averaging:						
I (Eq. 4)	4.1	4.7	5.3	5.8	99	0.68
II[2]	3.8	4.9	5.9	6.8	99	0.62
Multiple regression:						
training set	3.9	4.6	6.3	7.0	99	0.61
training + test set	3.9	4.8	6.2	7.8	161	0.64
95%-conf. interv.	(3.7–4.1)	(4.5–5.1)	(5.8–6.6)	(6.8–8.8)		
Charles (1985: A)	4.1	4.5	7.1	7.9	38	0.30
Davis & Anderson (1985)						
New England Full range	2.6	4.9[3]	7.1	6.1	31	0.44
Low range	2.5	4.9	7.0	4.6[3]	22	0.47
Norway Full range	1.5	5.1[3]	6.7	8.0	36	0.31
Low range	1.5	5.2	5.8[3]	9.8	28	0.28
Flower (1986)	4.1	4.3	6.5	9.3	33	0.36
Walker & Paterson (1986)	0.9	4.3[3]	6.8	6.3	28	0.45
Dixit (1986)	−5.0	4.7	6.2	7.4	28	0.46
Taylor *et al.* (1986)	3.4	5.1	6.4	8.0	60	0.68

[1] Including alkalibiontic taxa.
[2] Taking into account the coefficients of the extra regression (Table 4).
[3] Including unclassified taxa, if such taxa were not excluded in the calculation of fractions.

bined. Table 5 shows the estimated a_k-coefficients (Eq. 8) with 95%-confidence intervals and also summarizes the results of previous studies in terms of a_k-coefficients. These coefficients can tentatively be interpreted as optima of the pH groups. In general the coefficients increase – as they should – when going from the acidobiontic group to the alkaliphilous and alkalibiontic group, except in the New England set of Davis & Anderson (1985). Oehlert (1986) proposed Bayesian estimation of the coefficients. In Bayesian estimation existing information is incorporated in the estimation procedure. On the basis of Table 5 attractive 'prior' values seem to us (3, 5, 7, 8, 9), in which the acidobiontic group is very acid (pH = 3), the circumneutral group is neutral and the alkalibiontic group is very alkalic (pH = 9).

Our results point to the danger of using calibrations based on multiple regression on the untransformed counts of individual taxa. When translated to a_k-coefficients (Eq. 8), the coefficients obtained by Davis & Anderson (1985) and Flower (1986) ranged from -6 to 68 and 4.4 to 29 pH units, respectively, therefore cannot be considered optima. The resulting calibration functions do not allow even the slightest extrapolation beyond the range of diatom compositions found in the training set. For example, if a taxon with an a_k-coefficient of 50 has a relative abundance of 20% in a sample, then the inferred pH will be larger than 10 pH units (at least if the other taxa have positive a_k-coefficients). The pH inferred with such functions will be most erratic because of (1) the high random variability in diatom abundance and (2) the 'no modern analogue' problem (Davis & Anderson, 1985). Our results do not discredit the intelligent use of multiple regression in climate reconstruction from pollen data (Bartlein & Webb, 1985). Their calibration procedure contains checks to safeguard against non-linearity, outliers and extrapolation.

The maximum likelihood method using the Gaussian logit model yielded the smallest prediction error in the test set. Yet, despite its sophistication, it performed only slightly better than the method of weighted averaging, presumably because of the high variability in the relative abundances of the taxa (Fig. 2). The variability cannot be reduced greatly by taking more complex pH response curves than the Gaussian curve and is certainly partly due to the effects of other variables than pH. In the future we may attempt to account for this variability by adding other important variables (like humus content) to the response model. Calibration may then proceed by joint estimation of the variables using the maximum likelihood principle (Ter Braak, 1987a).

The weighted averaging method combines simplicity with a good performance. It can also be used for other variables than pH. It is the natural end-point of a historical development that started with Imbrie & Kipp (1971). To reconstruct past sea-surface temperature from Foraminifera, Imbrie and Kipp (1971) considered applying multiple regression of temperature on the abundances of the taxa. But this method was considered inappropriate as the abundances of species showed multicollinearity. So, they reduced the abundances of the taxa to a few axes by principal components analysis and then regressed temperature on these axes. The resulting equation was used for reconstruction. Roux (1979) produced better estimates of temperature, at least in the training set, by replacing principal components analysis – which assumes linear relationships – by correspondence analysis – which works also with unimodal relationships (Ter Braak, 1985). By rearranging taxa and samples in the data matrix in order of their scores on the first axis of correspondence analysis, he obtained a table with large abundance values near the principal 'diagonal' of the table and small values elsewhere. The table thus showed that relationships were unimodal. Gasse & Tekaia (1983) were concerned about the fact that only part of the information on the relationship of species to x is retained in the first few axes of the correspondence analysis. This is a potential problem in all indirect gradient analysis methods (Ter Braak, 1986). Gasse & Tekaia (1983) therefore developed a direct gradient analysis method (see Methods Section). The method of weighted averaging is a direct method also. The main difference is that, in the

method of Gasse & Tekaia (1983), pH is divided into classes; whereas in the weighted averaging method pH is treated as a quantitative variable (Ter Braak, 1987c).

The weighted averaging method cannot only be used to generate calibration functions that are specific to the region of application (Davis & Smol, 1986), but can possibly also be used with literature data on the pH optima and tolerance of diatoms, provided that comparable taxonomical and chemical methods have been used. The Hustedt system is a good basis for the compilation of a standard list of optima $\{u_k\}$ and 'indicative weights' $\{w_k\}$ for pH like that of Sládeček (1986) for saprobity (with $w_k = 1/t_k^2$, Eq. 6). This approach presupposes that pH optima and tolerances of the particular diatom taxa are the same throughout the world, or at least in large regions. To check this, data of Charles (1985) from the North American Adirondack lakes have been entered in Table 2. On average, the optima reported by Charles (1985) are 0.8 unit higher than our optima. The Spearman rank correlation coefficient is not very high ($r_s = 0.59$). There are several reasons for the disagreement. Charles used surface sediment samples; our samples were from the littoral zone. In surficial sediments of soft water lakes pH values may be higher than in the overlying water (Kelly *et al.*, 1984). The weighted averaging method has been found to be sensitive to the distribution of pH in the data set (Ter Braak & Looman, 1986). Thus the much lower median pH of our samples (4.2) than in Charles' samples (6.5) could also contribute to the observed differences between the pH optima of the taxa in Table 2. The differences point to the need of regional calibration equations.

This paper shows that for littoral zone samples the Index B method of Renberg & Hellberg (1982) can be improved upon. For use in acid waters in Western-Europe, we suggest two alternative methods. If the taxa listed in Table 2 contribute more than half the number of valves, we recommend using weighted averaging (Eq. 5) with the coefficients given in Tables 2 and 4. Otherwise, we propose to use the number of valves in Hustedt's pH groups by way of the formula obtained by multiple regression on all our 161 samples, i.e. by

$$pH = \frac{3.9 \times acb + 4.8 \times acp + 6.2 \times cir + 7.8 \times (alp + 1lb)}{acb + acp + cir + alp + alb} \quad (9)$$

Although the inferred pH of individual samples is not very precise by either method (SE $\approx$ 0.7), trends in pH may still show up.

Acknowledgements

We are much indebted to Dr. P. F. M. Coesel and Mrs. H. Kooyman-van Blokland (Hugo de Vries-Laboratorium, Universiteit van Amsterdam, Amsterdam), Dr. H. J. Dumont and Dr. I. van der Velde (Instituut voor Ekologie, Rijksuniversiteit, Gent), Dr. N. Foged (Odense), Prof. Dr. L. Franzisket and Dr. B. Gries (Westfälisches Museum für Naturkunde, Münster), Prof. Dr. E. G. Jørgensen (Lyngby), Dr. R. Simonsen (Institut für Meeresforschung, Bremerhaven) and Dr. K. Wouters (Koninklijk Belgisch Instituut voor Natuurwetenschappen, Brussels) for making available the samples of the test sest. We would like to thank Drs. I. C. Prentice, B. Engel, H. J. B. Birks and two anonymous referees for comments on the manuscript, Mr. C. N. Beljaars for counting diatoms and Mr. J. C. Buijs for technical assistance. Prof Dr. H. Lange-Bertalot and Dipl.-Biol M. Nörpel gave advice on taxonomy of some diatom taxa. The investigations were supported in part by the European Community (Contract nr. ENV-650-N (N)).

References

Alvey, N. C., N. Galwey & P. Lane, 1982. An introduction to GENSTAT. Academic Press, London, 152 pp.

Bartlein, P. J. & T. III Webb, 1985. Mean July temperature at 6000 yr B.P. in Eastern North America: regression equations from fossil-pollen data. Syllogeus 55: 301–342.

Battarbee, R. W., 1984. Diatom analysis and the acidification of lakes. Phil. Trans. r. Soc., London 305: 451–477.

Battarbee, R. W. & D. F. Charles, 1986. Diatom-based pH-reconstruction of acid lakes in Europe and North America. Wat. Air and Soil Pollut 30: 347–354.

Battarbee, R. W., J. P. Smol & J. Meriläinen, 1986. Diatoms as indicators of pH: an historical review. In J. P. Smol, R. W. Battarbee, R. B. Davis & J. Meriläinen (eds), Diatoms and lake acidity. Junk, Dordrecht. Developments in Hydrobiology 29: 6–14.

Berge, F., 1976. Diatoms and pH in some rivers and lakes in Agder and Telemark (Norway). A comparison between the years 1949 and 1975. Internal Report 18/76. SNSF-Project, Oslo-Ås, 36 pp. (in Norwegian).

Blakar, I. A. & T. Digernes, 1984. Evaluation of acidification based on former colorimetric determination of pH: the effect of indicators on pH in poorly buffered water. Verh. int. Ver. theor. und angew. Limnol. 22: 679–685.

Brown, P. J., 1982. Multivariate calibration. J. Roy. Statist. Soc. Ser. B 44: 287–321.

Charles, D. F., 1985. Relationships between surface sediment diatom assemblages and lakewater characteristics in Adirondack lakes. Ecology 66: 994–1011.

Davis, R. B., 1987. Paleolimnological diatom studies of acidification of lakes by acid rain: an application of Quaternary science. Quater. Sci. Rev. 6: 147–163.

Davis, R. B. & D. S. Anderson, 1985. Methods of pH calibration of sedimentary diatom remains for reconstructing history of pH in lakes. Hydrobiologia 120: 69–87.

Davis, R. B. & J. P. Smol, 1986. The use of sedimentary remains of siliceous algae for inferring past chemistry of lake waters – problems, potentials and research needs. In J. P. Smol, R. W. Battarbee, R. B. Davis & J. Meriläinen (eds), Diatoms and lake acidity. Junk, Dordrecht. Developments in Hydrobiology 29: 291–300.

Dixit, S. S., 1986. Diatom-inferred pH calibration of lakes near Wawa, Ontario. Can. J. Bot. 64: 1129–1133.

Dixit, S. S. & R. D. Evans, 1986. Spatial variability in sedimentary algal microfossils and its bearing on diatom-inferred pH reconstructions. Can. J. Fish. aquat. Sci. 43: 1836–1845.

Ellenberg, H., 1948. Unkrautgesellschaften als Mass für den Säuregrad, die Verdichtung und andere Eigenschaften des Ackerbodens. Ber. Landtech. 4: 130–146.

Ellenberg, H., 1979. Zeigerwerte der Gefässpflanzen Mitteleuropas. Scripta Geobotanica 9, Göttingen, 121 pp.

Flower, R. J., 1986. The relationship between surface sediment diatom assemblages and pH in 33 Galloway Lakes: some regression models for reconstructing pH and their application to sediment cores. Hydrobiologia 143: 93–103.

Gasse, F. & F. Tekaia, 1983. Transfer functions for estimating paleoecological conditions (pH) from East African diatoms. Hydrobiologia 103: 85–90.

Gauch, H. G. & R. H. Whittaker, 1972. Coenocline simulation. Ecology 53: 446–451.

Gause, G. F., 1930. Studies on the ecology of the Orthoptera. Ecology 11: 307–325.

Gifi, A., 1981. Nonlinear multivariate analysis. DSWO-press, Leiden, 451 pp.

Goff, F. G. & G. Cottam, 1967. Gradient analysis: the use of species and synthetic indices. Ecology 48: 793–806.

Haines, T. A., J. G. Akielaszek, S. A. Norton & R. B. Davis, 1983. Errors in pH measurement with colorimetric methods in low alkalinity waters. Hydrobiologia 107: 56–61.

Hesse, R., 1924. Tiergeographie auf ökologischer Grundlage. Verlag von Gustav Fisher, Jena, 613 pp.

Howe, S. E. & T. III Webb, 1983. Calibrating pollen data in climatic terms: improving the methods. Quater. Sci. Rev. 2: 17–51.

Hustedt, F., 1939. Systematische und ökologische Untersuchungen über die Diatomeenflora von Java, Bali und Sumatra. Arch. Hydrobiol./Suppl. 16: 274–394.

Imbrie, J. & N. G. Kipp, 1971. A new micropaleontological method for quantitative paleoclimatology: application to a late Pleistocene Caribbean core. In K. K. Turekian (ed.): The late Cenozoic glacial ages. Yale University Press, New Haven: 77–181.

Jones, V. J. & R. J. Flower, 1986. Spatial and temporal variability in periphytic diatom communities: Palaeoecological significance in an acidified lake. In J. P. Smol, R. W. Battarbee, R. B. Davis & J. Meriläinen (eds), Diatoms and lake acidity. Junk, Dordrecht, Developments in Hydrobiology 29: 87–94.

Jørgensen, B., 1983. Maximum likelihood estimation and large-sample inference for generalized linear and nonlinear regression models. Biometrika 70: 19–28.

Kelly, C. A., J. W. M. Rudd, A. Furutani & D. W. Schindler, 1984. Effects of lake acidification on rates of organic matter decomposition in sediments. Limnol. Oceanogr. 29: 687–694.

McCullagh, P. & J. A. Nelder, 1983. Generalized linear models. Chapman and Hall, London, 261 pp.

Montgomery, D. C. & E. A. Peck, 1982. Introduction to linear regression analysis. J. Wiley & Sons, NY, 504 pp.

Oehlert, G. W., 1986. Interval estimates for diatom inferred lake pH histories. Techn. report 469, School of Statistics, University of Minnesota, Minnesota.

Pantle, R. & H. Buck, 1955. Die biologische Ueberwachung der Gewässer und die Darstellung der Ergebnisse. Gas- und Wasserfach 96: 604.

Renberg, I. & T. Hellberg, 1982. The pH history of lakes in southwestern Sweden as calculated from the subfossil diatom flora of the sediments. Ambio 11: 30–33.

Roux, M., 1979. Estimation des paléoclimats d'après l'écologie des foraminifères. Cah. Anal. Données 4: 61–79.

Salden, N., 1978. Beiträge zur Ökologie der Diatomeen (Bacillariophyceae) des Süsswassers. Decheniana, Beih. (Bonn) 22: 1–238.

Shelford, V. E., 1911. Ecological succession: stream fishes and the method of physiographic analysis. Biol. Bull. (Woods Hole) 21: 9–34.

Sládeček, V., 1986. Diatoms as indicators of organic pollution. Acta hydrochim. hydrobiol. 14: 555–566.

Steinberg, C., K. Arzet & D. Krause-Dellin, 1984. Gewässerversauerung in der Bundesrepublik Deutschland im Lichte paläolimnologischer Studien. Naturwissenschaften 71: 631–634.

Stoermer, E. F. & T. B. Ladewski, 1976. Apparent optimal temperatures for the occurrence of some common phytoplankton species in southern Lake Michigan. Great Lakes Research Division, publication 18, the University of Michigan, Ann Arbor.

Taylor, M. C., H. C. Duthie & S. M. Smith, 1986. Surficial diatom assemblages – limnological relationships in Canadian Shield lakes. Poster, presented at the Ninth International Symposium on Living and Fossil Diatoms, Bristol. 24–30 August 1986.

Ter Braak, C. J. F., 1985. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal reponse model. Biometrics 41: 859–873.

Ter Braak, C. J. F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179.

Ter Braak, C. J. F., 1987a. Calibration. In R. H. G. Jongman, C. J. F. ter Braak & O. F. R. van Tongeren (eds), Data analysis in community and landscape ecology. Pudoc, Wageningen: 78–90.

Ter Braak, C. J. F., 1987b. The analysis of vegetation-environment relationships by canonical correspondence analysis. Vegetatio 69: 69–77.

Ter Braak, C. J. F., 1987c. CANOCO – a FORTRAN program for *canon*ical *c*ommunity *o*rdination by [partial]-[detrended][canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). TNO Institute of Applied Computer Science, Wageningen, 95 pp.

Ter Braak, C. J. F. & L. G. Barendregt, 1986. Weighted averaging of species indicator values: its efficiency in environmental calibration. Math. Biosci. 78: 57–72.

Ter Braak, C. J. F. & C. W. N. Looman, 1986. Weighted averaging, logistic regression and the Gaussian response model. Vegetatio 65: 3–11.

Thienemann, A., 1932. Limnologie. In R. Dittler, G. Joos, E. Korschelt, G. Linck, F. Oltmanns & K. Schaum (eds), Handbuch der Naturwissenschaften (2. Aufl.), 6. Band. Fischer, Jena: 434–475.

Van Dam, H., 1984. A guide to the literature for the identification of freshwater diatoms in The Netherlands. Hydrobiol. Bull. 18: 11–16.

Van Dam, H. & K. Beljaars, 1984. Nachweis von Versauerung in West-Europäischen kalkarmen stehenden Gewässern durch Vergleich von alten und rezenten Kieselalgenproben. In J. Wieting, B. Lenhart, C. Steinberg, A. Hamm & R. Schoen (eds), Gewässerversauerung in der Bundesrepublik Deutschland. Umweltbundesamt/ Schmidt, Berlin. Materialien 1: 184–188.

Van Dam, H. & H. Kooyman-van Blokland, 1978. Man-made changes in some Dutch moorland pools, as reflected by historical and recent data about diatoms and macrophytes. Int. Revue ges. Hydrobiol. 63: 587–607.

Van Dam, H., G. Suurmond & C. J. F. ter Braak, 1981. Impact of acidification on diatoms and chemistry of Dutch moorland pools. Hydrobiologia 35: 425–459.

Walker, I. R. & C. G. Paterson, 1986. Associations of diatoms in the surficial sediments of lakes and peat pools in Atlantic Canada. Hydrobiologia 134: 265–272.

Webb, T. III & D. R. Clark, 1977. Calibrating micropaleontological data in climatic terms: a critical review. Ann. N.Y. Acad. Sci. 288: 93–118.

Whittaker, R. H., S. A. Levin & R. B. Root, 1973. Niche, habitat and ecotope. Am. Nat. 107: 321–338.

Zelinka, M. & P. Marvan, 1961. Zur Präzisierung der biologischen Klassifikation der Reinheit fliessender Gewässer. Arch. Hydrobiol. 57: 389–407.