Aquat Ecol (2008) 42:293–305
DOI 10.1007/s10452-008-9186-7

Macroinvertebrate indicators of lake acidification: analysis of monitoring data from UK, Norway and Sweden

Ann Kristin Schartau · S. Jannicke Moe · Leonard Sandin · Ben McFarland · Gunnar G. Raddum

Published online: 23 April 2008

Abstract Although the acid sensitivity of many invertebrate species in lakes is well known, methods for assessment of lake acidification based on macroinvertebrate samples are less developed than for rivers. This article analyses a number of existing metrics developed for assessment of river acidification, and evaluates their performance for assessment of lake acidification. Moreover, new species-based indicators of lake acidification were developed and tested. The selected dataset contains 668 samples on littoral macroinvertebrates from 427 lakes with almost 60% of the samples from Sweden and the rest from UK and Norway. Flexible, non-parametric regression models were used for explorative analyses of the pressure–response relationships. The metrics have been assessed according to their response to pH, the degree of non-linearity of the response and the influence of humic compounds. Acid-sensitive metrics often showed a threshold in response to pH between 5.8 and 6.5. Highly acid-tolerant metrics were typically dominant across the whole pH range. Humic level had a positive effect for most acid-sensitive metrics. Generally, most metrics showed a more non-linear response pattern for the humic lakes than for clear lakes. The significant relationship between these macroinvertebrate metrics and acidification shows that there is a potential for developing further the assessment systems for ecological quality of lakes based on these metrics, although the metrics explained a low % of the variation (<30%). In order to improve the predictive power of the biotic metrics across the acidified part of Europe, further harmonization and standardisation of sampling effort and taxa identification are needed.

A. K. Schartau (✉)
Norwegian Institute for Nature Research,
Gaustadalléen 21, 0349 Oslo, Norway
e-mail: ann.k.schartau@nina.no

S. J. Moe
Norwegian Institute for Water Research,
Gaustadalléen 21, 0349 Oslo, Norway

L. Sandin
Department of Environmental Assessment, Swedish University of Agricultural Sciences, P.O. Box 7050, Uppsala 750 07, Sweden

B. McFarland
Environment Agency, South West Region, Manley House, Kestrel Way, Exeter EX2 7LQ, UK

G. G. Raddum
Department of Biology, LFI-Unifob, University of Bergen, Thormølensgt. 49, 5006 Bergen, Norway

Keywords Acidification indices · Generalised additive regression models (GAM) · Humic content · Littoral macroinvertebrates · pH · Water Framework Directive

Introduction

Anthropogenic acidification has been a severe threat to freshwater ecosystems for the last 40–50 years.

Atmospheric deposition of acid components has significantly impacted lakes and rivers in North America and Europe (Schindler et al. 1989; Skjelkvåle et al. 2005), resulting in marked decreases in pH and alkalinity and associated increases in SO_4 concentrations and in levels of toxic aluminium. Impoverished water quality from acidification has been followed by reduced species richness of aquatic organisms and a general shift in community structure in impacted lakes from acid sensitive taxa to more acid tolerant taxa (Larsen et al. 1996; Sandin et al. 2004). As a result of international agreements and actions to protect and restore natural resources threatened by acidification, the emission of sulphur (S) has been reduced by about two thirds since 1980 in Europe (EMEP 2005) followed by improvements in surface water quality (Stoddard et al. 1999; Skjelkvåle et al. 2005). In contrast to the documented rise in pH levels of surface water, evidence of widespread biotic recovery is generally lacking (Skjelkvåle et al. 2001; Gunn and Sandøy 2003). Large areas, especially in southern Norway, are permanently acidified or will suffer from acidification for years to come. Thus, acidification is still considered as one of the foremost problems affecting the biodiversity and functioning of inland surface waters in northern Europe (e.g. Brodin 1995; EMEP 2005). Furthermore, due to a rapid industrialisation of some developing countries, acidification is already an increasing problem (e.g. China) (Larssen et al. 2006).

Increased concentration of specific forms of inorganic aluminium (Al) following freshwater acidification is a major water quality problem for the biota (e.g. Driscoll et al. 1980). High levels of dissolved organic carbon (DOC) generally reduce the negative effects of aluminium on fish (Witters et al. 1990; Laudon et al. 2005). However, little is known about the interaction between DOC and effects of acidification regarding macroinvertebrates.

Benthic invertebrates are a diverse and generally abundant group with a wide range of environmental tolerances and preferences which can act as indicators of environmental quality (Rosenberg and Resh 1993). The rather straightforward relationship between acid conditions and the presence/absence of certain benthic macroinvertebrate species has therefore been used to assess the effects of acid stress on stream ecosystems (e.g. Henrikson and Medin 1986; Raddum et al. 1999; Davy-Bowker et al. 2005; Furse et al. 2006). Generally, there has been less focus on acid stress effects on macroinvertebrates in lakes. Until the present date, benthic macroinvertebrates have not been consistently included in national monitoring systems for lakes and there is a general lack of data on littoral macroinvertebrates (Moss et al. 2003; Solimini et al. 2006).

The implementation of the EC Water Framework Directive (European Commission 2000) requires development of cost-efficient ecological assessment systems for freshwaters in Europe. Thus it is vital to establish simple biological metrics that can be used to classify ecological status in lakes and rivers. The purpose of this study was to test the performance of various macroinvertebrate metrics in relation to acidification pressures in humic and non-humic lakes in northern Europe. Using compiled multi-national datasets may have several advantages and disadvantages (see Moe et al. this issue). One of the main challenges is the heterogeneity in data caused by differences in taxonomic resolution. Further, preliminary work has indicated that many of the individual indicator taxa that are included in the most commonly used acidification indices are applicable only for restricted parts of northern Europe. Two different approaches were followed to overcome these problems:

(1) Existing macroinvertebrate metrics, developed for assessment of river acidification and general degradation, were tested by aggregating data to family level before analysis;
(2) New indicators of lake acidification were developed by combining several indicator species with similar response to acidification into species-based sensitivity groups. The latter approach enabled us to analyse data at a more detailed taxonomic resolution than the former.

To our knowledge, this is the first analysis of dose–response relationship between acidification and macroinvertebrates in lakes with a broad dataset representing northern Europe.

Material and methods

Datasets

The selected dataset for analyses was extracted from the REBECCA database on lake macroinvertebrates

(Moe et al., this issue). The database currently contains ca. 2,050 samples from 765 lakes in the Northern intercalibration region (UK: 33, Norway: 57, Sweden: 675). Only samples from the littoral zone (0–1 m) of lakes with pH 4–7 and low alkalinity (alk 0.05–0.2 meq/l and low Ca (1–4 mg/l) were included. Sites considered to be treated by liming to reduce acidification, were excluded. Consequently the number of lakes included in the analyses is restricted to 10, 29 and 388 for UK, Norway and Sweden, respectively.

All littoral samples of macroinvertebrates were taken by kick-sampling (EN 27828 1994). The sampling effort and the season sampled differed somewhat among countries, and it was not possible to standardise the abundances. However, metrics used in our study are based on relative abundances.

Water samples for chemical analyses were taken at 0.2–0.5 m depth at the middle of the lake, or from the outlet river. The analyses were performed at the Norwegian Institute for Water Research (Norway), SWEDAC laboratory, SLU (Sweden), National Laboratory service (England and Wales) and SEPA regional laboratories (Scottland) using national accreditation scheme (EN ISO/IEC 17025 2005).

The computer programme ASTERICS (version 3.0) developed in the EU-funded projects AQEM (www.aqem.de) and STAR (www.eu-star.at) was used to calculate a number of existing biotic metrics based on benthic macroinvertebrate taxa lists. These metrics are developed for biological assessment of rivers and we wanted to evaluate their performance for assessment of lake acidification. In co-operation with the Acidification Working Group in the Northern Geographic Intercalibration Group (N-GIG) we selected the following metrics developed for assessment of acidification and general degradation of ecological quality: No. of Gastropoda families, No. of Ephemeroptera families, Proportion of Ephemeroptera, Proportion of Diptera, Proportion of predators, family-level Acid Water Indicator Community (AWIC F) (Davy-Bowker et al. 2005) and Medin acidification index (Henrikson and Medin 1986). Except for proportions of Diptera and predators, these metrics are expected to decrease with decreasing pH. All metrics, except Medin index, are included in the Multimetric Index for Lake Acidification (MILA), suggested in the new WFD-compatible quality criteria for Sweden (Johnson and Goedkoop 2007). MILA was also included in our analyses. Data were aggregated to family level before the metrics were calculated. This helped to reduce the variation among countries, which probably resulted from different standards for sampling effort and taxonomic resolution.

We developed species-based indicators by selecting littoral species, and in some few cases genera, associated with different categories of acid sensitivity. These species were identified by examining literature that assessed commonly used acidification indices: Medin index (Henrikson and Medin 1986), Raddum index I (Raddum and Fjellheim 1984; Fjellheim and Raddum 1990), NIVA index (Bækken and Kjellberg 2004) and species-level AWIC (AWIC sp) (J. Davy-Bowker, pers. comm.). We have assigned indicator species at three levels of aggregation. The 1st level indicators consist of selected indicator species that are used by at least two existing indices, and where the sensitivity of the species is generally agreed (see Appendix 1). The 1st level indicators have been used as basis for developing the 2nd and 3rd level indicators, which represent selected groups of indicator species with similar response to acidification (see Table 1). The 2nd level indicators represent mainly order level, and the 3rd level indicators representing EPT species (Ephemeroptera, Plecoptera, Trichoptera) versus non-EPT species. Whereas some species are applicable as indicators only for restricted parts of the Northern intercalibration region, due to limited geographical distribution and/or different sensitivity across the region, the species-based sensitivity groups, represented by more than one species, are expected to be present and showing similar sensitivity to acidification throughout the northern region. This approach also enabled analysis of data of a more detailed taxonomic resolution than most of the existing metrics did. We consider this important, as closely related species/genera may show differences in sensitivity to acidification.

Biological and chemical samples were not always taken simultaneously. It was therefore necessary to group and average the samples within time intervals, before attempting to link the biology and chemistry. The biological metrics were grouped and averaged within season before they were linked with chemistry. The following seasons were defined: spring (April–June), summer (July–August), autumn (September–November) and winter (December–March).

Table 1 Evaluation of littoral macroinvertebrate indicators displaying different sensitivity to acidification. Indicator taxa (1st level: species, genera) identified by the literature and used in different acidification indices are combined into new indicators. 2nd level refers to the first set of species-based sensitivity groups (Sp. sens. groups), whereas 3rd level refers to further combinations of these indicators into EPT-based (=Ephemeroptera + Plecoptera + Trichoptera) and non-EPT-based indicators. A few of the selected species (Bivalvia: *Anodonta* sp. and Coleoptera: *Esolus parallelepipedus*) could not be tested due to lack of observations in the REBECCA database

Sensitivity to acidification	Taxon (1st level)	Sp. sens. groups (2nd level)	Sp. sens. groups (3rd level)
Highly sensitive	*Alainites muticus, Baetis scambus/fuscatus, B. macani, Nigrobaetis digitatus, Procloeon bifidum, Ephemera danica, Caenis luctuosa, C. robusta, C. rivulorum*	Ephemeroptera	Highly sensitive EPT
	Dinocras cephalotes	Plecoptera	
	Glossosoma intermedium, Wormaldia subnigra, Philopotamus montanus, Chimarra marginata, Cheumatopsyche lepida, Ceraclea annulicornis	Tricoptera	
	Potamopyrgus antipodarum, Radix balthica (Lymnaea peregra), Galba truncatula, Bathyomphalus contortus	Gastropoda	Highly sensitive non-EPT
	Glossiphonia complanata, Hemiclipsis marginata, Theromyzon tessulatum	Hirudinea	
	Gammarus pulex, G. lacustris	Crustacea	
Sensitive	*Siphlonurus aestivalis, S. alternatus, Seratella ignita (Ephemerella ignita), Ephemerella aroni (E. aurivillii), Ephemera vulgata, Caenis horaria*	Ephemeroptera	Sensitive EPT
	Capnia atra, C. pygmaea	Plecoptera	
	Agapetus fuscipes gp., A. ochripes, Hydropsyche silfvenii (Ceratopsyche silfvenii), Hydroptila sp., Ithytrichia sp., Mystacides azurea, Oecetis testacea, Trianodes bicolor	Trichoptera	
	Physa fontinalis, Gyraulus albus, G. acronicus, Ancylus fluviatilis	Gastropoda	Sensitive non-EPT
	Helobdella stagnalis, Erpobdella testacea	Hirudinea	
	Dixa sp.	Diptera	
Tolerant	*Nigrobaetis niger, Heptagenia sulphurea*	Ephemeroptera	Tolerant EPT
	Amphinemoura borealis, Nemoura avicularis, Siphonoperla burmeisteri	Plectoptera	
	Rhyacophila fasciata, Oxyethira sp., Molanna angustata, Goera pilosa, Micrasema setiferum	Trichopetra	
Highly tolerant	*Kageronia fuscogrisea, Leptophlebia vespertina, L. marginata*	Ephemeroptera	Highly tolerant EPT
	Taeniopteryx nebulosa, Brachyptera risi, Amphinemoura standfussi, A. sulcicollis, Nemurella picteti, Nemoura cineria, Protonemoura meyeri, Leuctra nigra, L. hippopus, L. digitata, Isoperla difformis	Plecoptera	
	Rhyacophila nubila, Cyrnus flavidus, C. trimaculatus, Neureclipsis bimaculata, Plectrocnemia conspersa, Holocentropus dubius, Polycentropus flavomaculatus, P. irroratus, Athripsodes aterrimus	Trichoptera	

The resulting number of (averaged) samples is given in Table 2.

Statistical analyses

All metrics were tested against pH as a measure of acidification. Acid Neutralising Capacity (ANC) is usually included as a potential predictive variable in models that evaluate biological effects of anthropogenic acidification (Driscoll et al. 1991). However, the quantity and quality of the ANC-data were inadequate for our analyses.

Table 2 Number of macroinvertebrate samples available for analyses, after temporal grouping for linking with chemistry

Country	Clear	Humic	All
UK	96	0	96
Norway	151	33	184
Sweden	96	292	388
All	343	325	668

For the species-based indicators, biological response to acidification was tested using proportional abundance or presence/absence data. Proportional abundances of EPT indicators (i.e. Ephemeroptera, Plecoptera, Trichoptera) were calculated on the basis of total number of EPT individuals, whereas other metrics were calculated on the basis of total numbers of invertebrates.

The lakes in the dataset were defined as either clear (colour <30 mg l^{-1} Pt or TOC < 5 mg l^{-1}) or humic (colour >30 mg l^{-1} Pt or TOC >5 mg l^{-1}). These two groups represent the Intercalibration lake types L-N2 and L-N5 (clear) and L-N3 and L-N6 (humic), respectively (European Commission 2004). The data were analysed with humic level as a categorical co-variable, so that estimated response curves and significance probabilities (*P*-values) could be obtained for each data category.

An important purpose of the analyses is to describe the relationships between pressure and response metrics, and in particular explore non-linearities and thresholds in the response curves, because such responses can make a metric useful in determining classification boundaries. We have, therefore, primarily used flexible non-parametric regression methods, which allow for estimation of non-linear response curves without any assumptions about parametric forms. For the existing biotic metrics we have used an ordinary generalised additive regression model (GAM). The estimated number of degrees of freedom (d.f.) indicates the non-linearity of the response (see Table 3). Not all the metrics are normally distributed (as the method assumes), and therefore the *P*-values for these metrics must be interpreted with caution. The species-based indicators show consistently skewed distributions with a high proportion of zeros, which means that regression based on the means may not be suitable. Therefore, we have used three different regression methods, to explore how different aspects of the distribution respond to the pressure gradient. First, the raw metric values (proportions) are analysed with an additive regression model. Second, the metrics values are transformed to presence/absence and analysed with logistic regression. Finally, the proportions are analysed by quantile regression, representing the response of the upper extreme (90%) of the distribution.

All statistical analyses are performed in R 2.3.1 (R Development Core Team 2006). For GAM analyses we have used the package "mgcv" (Wood 2006), and for quantile analyses we have used the package "quantreg" (Koenker 2006).

Results

Relationship between pH and existing biotic metrics

All biotic indices were significantly related to pH for clear and/or humic lake types (Fig. 1, Table 3). All indices showed significant response for humic lakes (deviance explained: 3.2–23.8%, *P*: <0.001 to 0.03), whereas two of them, i.e. proportion of Ephemeroptera and MILA, did not show any clear relationship with pH for clear lakes. There was a small, albeit significant, increase in proportion of Diptera and proportion of predators with decreasing pH for both clear and humic lakes (deviance explained: 3.2–13.3%, *P*: <0.001 to 0.03). The remaining four indices were more promising, with a general decrease in No. of Gastropoda, No. of Ephemeroptera families, the AWIC F index and the Medin index, with decreasing pH (deviance explained: 9.7–23.8%, all $P < 0.001$). Typical for all of these indices were, however, that low index values were found across the whole range of pH values. For example, the number of Ephemeroptera families found in a littoral sample increased with an increase in pH, but also with a pH near 7 there were still a number of sites that did not contain any ephemeropterans. On the other hand, if pH exceeded 6.0 a relatively high proportion of the lakes contained four or more Ephemeroptera families, whereas few lakes with pH < 6.0 had more than three and no lakes had more than four Ephemeroptera families. The same was generally true also for lakes with two or more Gastropoda families, and a Medin index above five.

For four out of the eight indices tested, there was a difference in the response between clear-water lakes and humic lakes, namely proportion on Ephemeroptera, No. of Ephemeroptera families, AWIC F and MILA (Fig. 1, Table 3; all $P < 0.001$). All these indices had a higher index value for humic than for clear-water lakes throughout the pH gradient (Fig. 1). In these cases, the type of response differed too (Table 3). For the clear-water lakes, the response was generally linear (AWIC F, Medin and MILA; d.f. = 1) or near-linear (proportion of Ephemeroptera; d.f. = 1.7). For the humic lakes, most biotic indices showed a more complex response pattern (d.f. = 1.8–8.4): either with hump-shaped curve or a more-than-linear increase. However, No. of Ephemeroptera families showed a more complex response pattern in clear-water lakes (d.f. = 6.6) than

Table 3 Test results from analysis of (A) Existing metrics and (B) species-based sensitivity groups versus pH. Existing metrics are analysed by generalised additive regression models. Species-based sensitivity groups are analysed by generalised additive logistic regression models

	Clear lakes			Humic lakes			Humic versus clear lakes		
	d.f.	Deviance explained (%)	*P*-value	d.f.	Deviance explained (%)	*P*-value	Humic effect	Deviance explained (%)	*P*-value
(A) Existing metrics									
Ephemeroptera %	1.7	0.7	0.500	3.4	7.0	<0.001	7.6	9.5	<0.001
Diptera %	4.2	4.9	0.020	2.5	3.2	0.030	−1.3	3.9	0.200
Ephemer. no. families	6.6	9.7	<0.001	1.8	23.0	<0.001	8.9	25.1	<0.001
Gastropoda no. families	1.7	11.4	<0.001	8.4	19.3	<0.001	−0.2	15.3	0.900
AWIC F	1.0	12.7	<0.001	3.4	22.7	<0.001	11	29.7	<0.001
Predators %	1.6	3.8	0.006	6.1	13.3	<0.001	−0.4	5.7	0.700
Henrikson-Medin	1.0	18.4	<0.001	3.8	23.8	<0.001	1	20.5	0.300
MILA	1.0	0.1	0.600	1.9	15.4	<0.001	13	23.3	<0.001
(B) Species-based sensitivity groups									
2nd level indicators									
Highly sens. Ephemer.	1.0	1.5	0.400	1.0	6.5	0.002	2.3	6.4	0.020
Highly sens. Gammarus	2.5	13.2	0.009	1.0	6.7	0.004	−4.3	15.8	<0.001
Sensitive Ephemer.	2.4	8.2	0.100	2.7	12.8	<0.001	6.7	16.1	<0.001
Sensitive Hirudinea	1.4	1.1	0.400	1.0	2.8	0.003	3	2.7	0.002
Highly toler. Ephemer.	5.8	20.1	<0.001	2.0	1.0	0.200	1.8	3.2	0.070
Highly toler. Trichoptera	1.0	3.3	<0.001	1.0	2.5	<0.001	4.4	5.0	<0.001
3rd level indicators									
Highly sensitive EPT	1.0	1.9	0.300	1.0	7.0	<0.001	2.5	7.0	0.010
Highly sens. non-EPT	1.0	8.2	<0.001	1.1	5.8	0.003	−2.5	8.6	0.010
Sensitive EPT	2.1	2.6	0.090	2.5	10.7	<0.001	5.3	9.5	<0.001
Sensitive non-EPT	1.0	2.4	0.020	1.0	3.6	<0.001	1.9	3.4	0.060
Tolerant EPT	1.3	0.3	0.400	1.4	3.7	0.005	0.6	2.2	0.600
Highly tolerant EPT	1.9	3.3	0.008	1.2	2.2	0.003	4	4.2	<0.001

Explanation to headers under "Clear lakes" and "Humic lakes": Estimated d.f. (degrees of freedom) indicates the degree on non-linearity for the response curve: 1 = approximately linear, higher d.f. = more non-linear. Deviance explained: the proportion of variation that can be explained by the model (metric = *f*(pH)). *P*-value indicates whether the estimated response curve is significantly different from a horizontal line. Explanation to headers "Clear versus Humic lakes": Deviance explained: the proportion of variation that can be explained by the model (metric = *f*(pH) + humic). *P*-value indicates whether the effect of humic level (high vs. low) on the response curve is significant. Humic effect: estimated difference in response value for humic versus clear lakes (note that the different metrics may have different scales, cf. Fig. 1)

in humic lakes (d.f. = 1.8). The unexplained variation within the dataset with the measured pH values is still quite large, which indicates that there are still factors other than pH and humic content that affect the biotic metrics.

Relationship between pH and species-based sensitivity groups

Initial analysis of presence/absence data of all macroinvertebrate samples (all lake types) showed that eight out of eighteen species-based sensitivity indicators of 2nd level (Table 1) displayed a significant response to pH. These were highly sensitive *Gammarus*, highly sensitive Ephemeroptera, sensitive Gastropoda, sensitive Hirudinea, sensitive Ephemeroptera, sensitive Plecoptera, highly tolerant Ephemeroptera and highly tolerant Trichoptera (all $P < 0.05$). Only six of these indicators were selected for further analyses (Fig. 2), because of scarcity of data and similar response to pH for several of the indicators. Further combinations of indicators resulted in six 3rd level indicators (Fig. 3). In all,

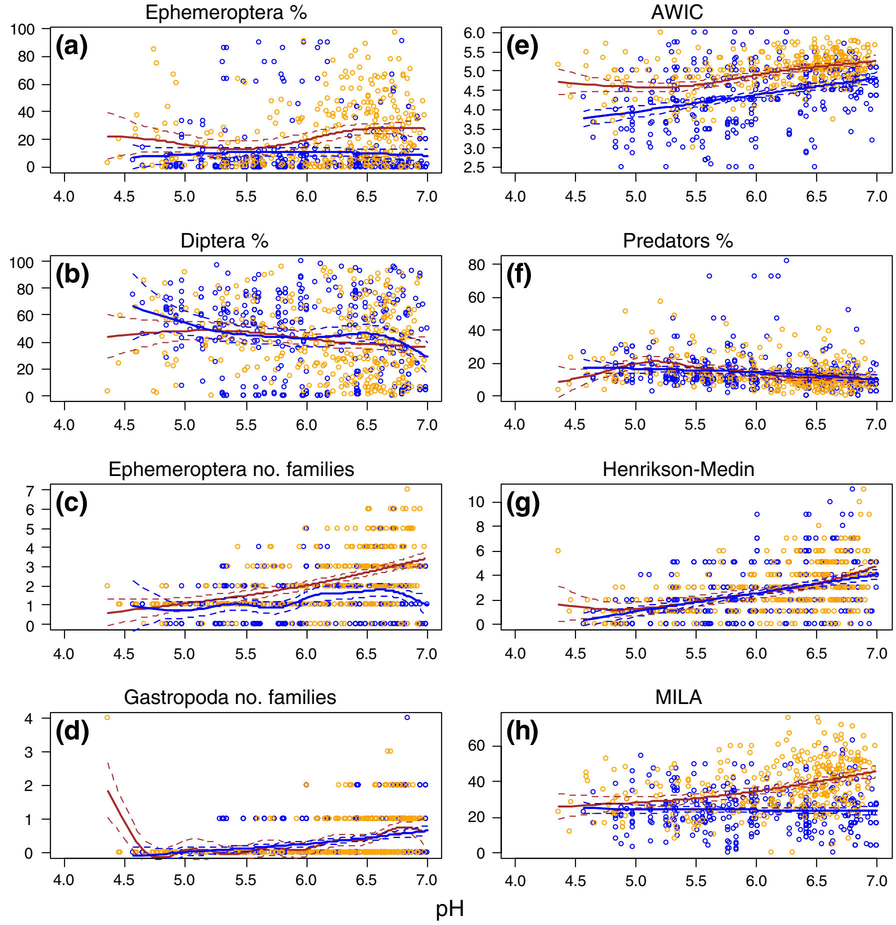

Fig. 1 Relationships between existing macroinvertebrate metrics and pH for clear (blue) and humic (orange) lakes. The curves are estimated by a generalised additive regression model (GAM). Stipled lines are ±2 standard errors. Humic level can be interpreted as having a statistically significant effect when the two curves have non-overlapping confidence intervals

twelve species-based indicators were selected (Table 3). All indicators, except highly tolerant Ephemeroptera, responded significantly to pH if analysed on data from humic lakes (Figs. 2, 3, Table 3; deviance explained: 2.2–12.8%, P: <0.001 to 0.005), whereas only six indicators showed significant response for clear lakes. These were highly sensitive *Gammarus*, highly tolerant Ephemeroptera, highly tolerant Trichoptera, highly sensitive non-EPT species, sensitive non-EPT species and highly tolerant EPT species (Figs. 2, 3, Table 3; deviance explained: 2.2–20.1%, P: <0.001 to 0.02).

The species-based sensitivity groups, displayed as relative abundance, have very skewed distributions. Although the non-parametric version of logistic regression is more flexible than ordinary parametric logistic regression, the logit link function forces the estimated curve to be more or less monotonously increasing or decreasing. However, not all indicators show a monotonous response to pH, as shown by using quantile regression. For instance, a high proportion of highly tolerant Trichoptera is more likely at pH 5.4–6 than at lower or higher pH (Fig. 2).

For most metrics tested, the response between clear-water and humic lakes differed (Figs. 2, 3, Table 3). Highly sensitive Ephemeroptera, sensitive Ephemeroptera, sensitive Hirudinea, highly tolerant Trichoptera, highly sensitive EPT, sensitive EPT, highly tolerant EPT had a higher proportional abundance for humic lakes than for clear-water lakes, along most of the pH gradient (Figs. 2, 3, Table 3; P:<0.001 to 0.002). The only metrics that showed a different pattern were highly sensitive *Gammarus* ($P < 0.001$) and highly sensitive non-EPT (which includes *Gammarus*) ($P = 0.01$). Higher proportions of these metrics were generally found in the clear-water lakes than the humic lakes (Figs. 2, 3). The shape of response curve between

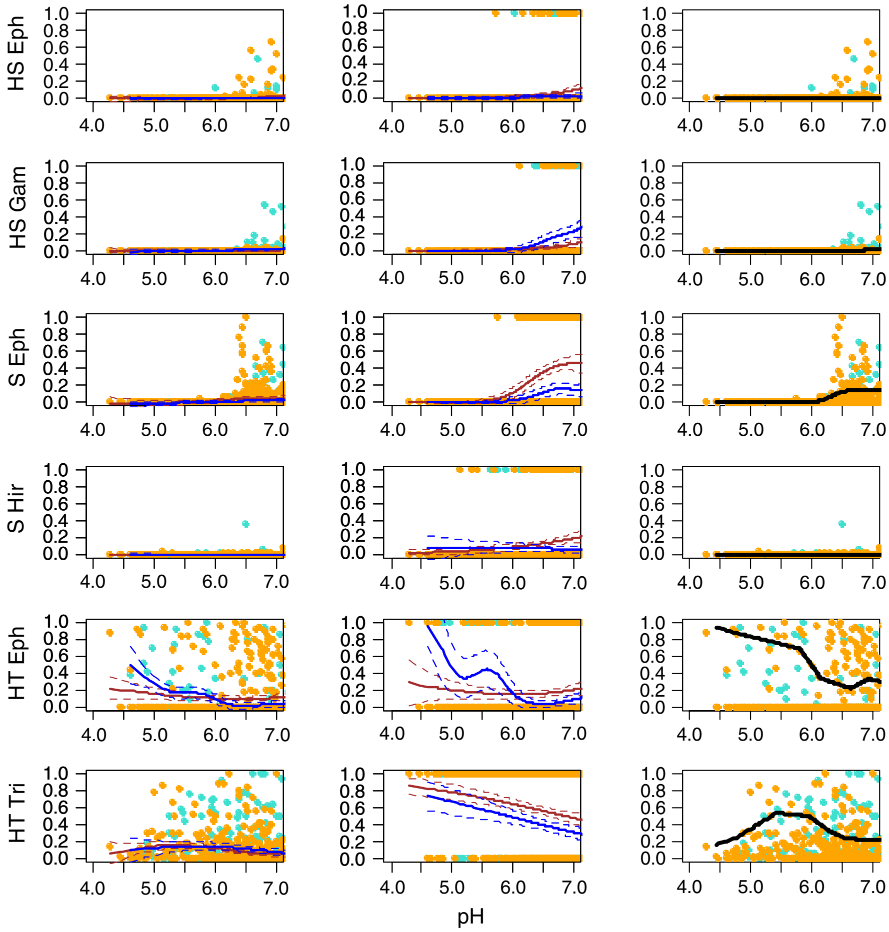

Fig. 2 Relationships between pH and proportional abundance of selected species-based sensitivity groups (2nd level indicators; see Table 1 for definitions). HS: Highly sensitive, S: Sensitive, HT: Highly tolerant, T: Tolerant. The curves are estimated by three regression methods. Left panel: additive regression model. Middle panel: generalised additive logistic regression model (the estimated curve can be interpreted as the probability of presence (*y*-axis) in a location with the given pH (*x*-axis)). Right panel: 90% quantile regression (the estimated curves represent the lower 90% of the metrics values along the pH gradient: at a given pH value, approximately 10% of the samples with this pH value will have a metric value above the estimated curve). Stipled curves represent ±2 standard errors. Black curves are estimated for all data combined. Where there was a significant difference between clear and humic stations (see Table 3), regression curves were also estimated separately for clear (blue) and humic (orange) lakes

clear-water (d.f.: 1.0–5.8) and humic lakes (d.f.: 1.0–2.7) differed for some metrics but there was no general pattern.

Highly sensitive and sensitive indicators show a threshold with regard to pH, below which there were no observations. Generally, the relative abundance of highly sensitive indicators decrease at pH 6.5 with very few records below pH 6.0, whereas the relative abundance of less sensitive indicators are quite high at pH 6.0 and also occasionally found below pH 5.5.

Discussion

In our study, the metrics have been assessed according to three criteria: statistical significance, explanatory degree and non-linearity or threshold response. In addition, the effect of humic level on the responses has been analysed. Threshold-type responses are especially useful for setting classification boundaries. The graphs indicate that several of the species-based indicators categorised as sensitive to acidification show a threshold with regard to pH,

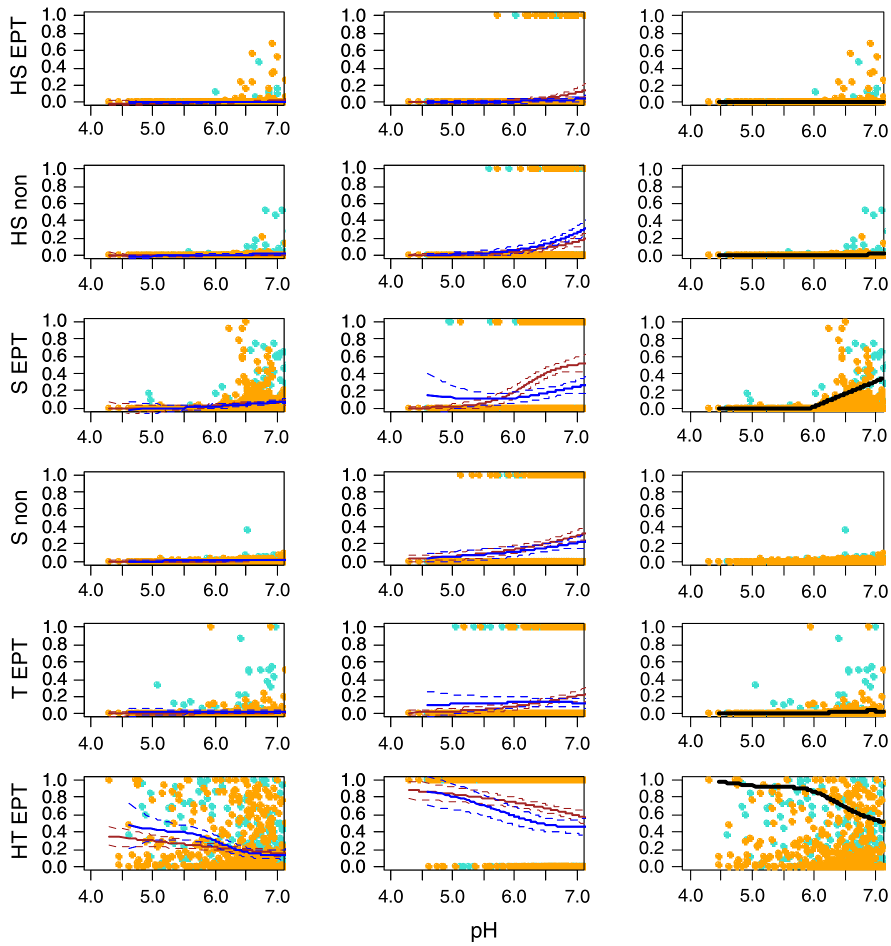

Fig. 3 Relationships between pH and proportional abundance of selected species-based sensitivity groups (3rd level indicators; see Table 1 for definitions). EPT = Ephemeropterea + Plecoptera + Trichoptera. See Fig. 2 for description of the three panels

below which no records should be expected. This type of response is more typical for the new species-based sensitivity groups than for the existing metrics. One reason may be that the existing metrics are often more composite, so that possible threshold-responses of several taxa/metrics merge. Another possible reason is that several of the existing metrics are based on presence/absence of taxa only, whereas, the new metrics are based on relative abundances. Presence/absence based metrics may have a high power to detect improvement of water quality, since the colonisation of the site by one sensitive taxon alone will improve the score (Sandin and Johnson 2000). On the other hand, such metrics may have low power to detect impoverishment of a site, since all individuals of all sensitive taxa must disappear before the site is assessed as affected (Sandin and Johnson 2000). The new metrics developed here allow for using proportional abundances, and therefore will be able to show changes in species composition before the species disappear completely.

The selected metrics generally show a significant response to pH, at least in humic lakes, but the explanatory degree (R^2 or reduction in deviance) is often low (<30%). The low explanatory degree is often related to a high variation in the data, especially at high pH values. One possible reason for why so few of the species-based sensitivity groups show significant response to pH for clear lakes is the high proportion of no-records or low abundances, even in cases where pH is quite high. For these metrics the selected taxa are generally more abundant in humic systems (three of the four indices, which are the basis for the development of species-based sensitivity groups, are developed primarily for humic systems: AWIC sp, Medin index, NIVA index). Moreover, high levels of humic substances may ameliorate the toxicity of metals, primarily of aluminium, which has documented toxic effects on biota in acidified freshwaters (Driscoll et al. 1980). This ameliorating effect may be especially important for species which experience sub-optimal conditions in these low-alkalinity lakes with low calcium concentrations. Calcium-demanding organisms, such as snails and mussels (cf Økland 1980), may experience such sub-optimal conditions even though the lake is not or only slightly acidified.

Humic level had a positive effect for most acid-sensitive metrics analysed, i.e. at a given pH higher proportions of acid sensitive taxa were found for humic lakes than for clear-water lakes. This supports the hypothesis that humus buffers macroinvertebrates against the detrimental effects of low pH (Hageby and Petersen 1988; Gensemer and Playle 1999). Dangles et al. (2004) showed that anthropogenic acidification has more severe effects than natural acidity on macroinvertebrate communities and processes in rivers. On the other hand, in humic lakes natural acidity may interfere with anthropogenic acidification: natural pH for low-alkalinity, humic waters are typical 4.5–7.0. Therefore, we are not able to separate the anthropogenic contribution to low pH values from the contribution by natural acidity in humic lakes. Further work on lake macroinvertebrates should look at the validity of using ANC in combination with DOC, as a measure of anthropogenically induced acidification (cf Lydersen et al. 2004). Hesthagen et al. (this issue) show that at a given value of pH and inorganic Al, critical limits for ANC must be higher in humic waters than in clear waters in order to retain unaffected fish populations.

Dangles et al. (2004) showed that although the macroinvertebrate richness was not different between the circumneutral and the naturally acid streams in northern Sweden, the species composition nevertheless differed. This supports the idea that clear-water and humic lakes belong to separate types, which differ both regarding reference conditions and response to acidification.

Based on the considerations above, the following existing metrics appear most promising for classification of ecological status: AWIC F, and Medin index both for clear and humic lakes, No. of Gastropoda families for clear lakes only and No. of Ephemeroptera families for humic lakes only. Out of the species-based sensitivity groups, the most promising were highly tolerant Ephemeroptera and highly sensitive non-EPT for clear lakes and sensitive Ephemeroptera and sensitive EPT for humic lakes.

Taxa and indices that show a positive response with increasing pH might be good indicators of biological recovery. However, the indicators analysed here sometimes show low values and even "0" also in cases where pH is quite high. Four possible explanations are: (i) There is truly no response from the selected taxa to increases in pH (e.g. sensitivity indicated by the literature is not correct). (ii) The sensitive-scoring taxa are not present due to other factors than acidification. Acid-sensitive lakes, as those included in this study, have low concentrations of Ca; in many cases below critical values for Ca-demanding taxa like snails and mussels (cf Økland 1980). (iii) The lake is actually acidified although this is not indicated by the measured pH, or the biology corresponds to a previously more acid water (i.e. the corresponding water chemistry samples are not representative in time or space). (iv) The sensitive-scoring taxa are present at the site, but have been missed in sampling, sorting or identification of the organisms. This can be caused by taking too small samples for instance (Resh 1979; Doberstein et al. 2000). It is well known that the number of taxa observed increases with the area or number of individuals sampled (Arrhenius 1921; Li et al. 2001). Natural variability will affect the ability to detect changes using biological metrics. Generally, metrics based on the presence of indicator taxa are more robust than enumeration and taxonomic richness measures (Barbour et al. 1992; Goedkoop et al. 2000). However, the effect of sampling effort on taxonomic richness is considered to be less pronounced by using less detailed taxonomic resolution (cf Warwick 1993). In our analysis of taxonomic richness measures the data were aggregated to family level, therefore our results should not be much affected by sampling effort. On the other hand, our analysis of species-based sensitivity groups may be affected by a high degree of under recording of taxa, i.e. for a substantial part of the samples many specimens are identified to genera or higher taxonomic level whereas the metric require species-level data. High number of juvenile specimens which make species level identification error-prone or even impossible, and lack of identification keys for some taxonomic groups, are the main reason for such under recording. Moreover, not all species are included in the analysis, in spite of being indentified to species, because these species are not recognised by the ASTERICS software. Standardisation of the taxonomic resolution (generally to species or genera) across the region as well as amendment of the software calculating the metrics, are expected to improve the relationship between pH and metrics based on lake macroinvertebrates.

For all highly tolerant indicators, high relative abundance and high frequency of occurrence were found across the whole range of pH values. Thus, the presence of these indicators cannot be used alone to conclude that a sites is affected by acidification. The EU Common Implementation Strategy Guidance on the Intercalibration process (http://forum.europa.eu.int/Public/irc/env/wfd/library) recommends that both sensitive and tolerant indicators are used together for setting boundary values, especially for the good/moderate boundary. Another approach is to combine tolerant and sensitive indicators into a multimetric index. To our Knowledge, The MILA index developed for assessing acid stress in Swedish lakes (Johnson and Goedkoop 2007) is the only multimetric index developed for littoral benthic macroinvertebrates addressing lake acidification. In our study, the MILA index was calculated from macroinvertebrate data identified to family level only. Thus this index was expected to be applicable for a larger geographical region. However, our results from clear-water lakes demonstrated that although several of the metrics included in the MILA index show significant response to acidification, the MILA index itself did not. Generally, a multimetric index should only include metrics showing a quantitative dose-response change across a stressor gradient that is reliable, interpretable and not dispersed or obscured by natural variation (Hering et al. 2004). In addition, the MILA approach does not take different sensitivity among closely related taxa into consideration. In order to reduce the variation in indicator values that are caused by other factors than acidification, the number and type of taxa, included in the taxa list for both tolerant and sensitive indicators, should be carefully considered (cf. Sandin and Johnson 2000).

Conclusions

The strong relationship between lake littoral macroinvertebrate communities and acidification show that there is a potential for developing further systems for assessing the ecological quality of lakes based on these indicators. This could be done by combining those macroinvertebrate metrics showing the strongest response to acidification into a new multimetric index. However, to improve the predictive power of the biotic metrics, there is need for further harmonization and standardisation of sampling effort and taxa identification with respect to taxonomic resolution across the region.

The work on developing a new multimetric index for lake acidification based on macroinvertebrates will be taken further by the Northern Geographical Intercalibration Group, WG Acidification.

Acknowledgement Members of the N-GIG, WG Acidification have given valuable input to the development of new metrics and interpretation of analyses: Ian Fozzard, Willem Goedkoop, Ruth Little. We also thank Scottish Environmental Protection Agency, Environment Agency for England and Wales, LFI-Unifob at the University of Bergen, Norway, and the Swedish Environmental Protection Agency for providing datasets on littoral lake macroinvertebrates and chemistry. This study has been a part of the EU-funded project REBECCA (SSPI-CY-2003-502158) and has also been financially supported by the Norwegian Agency for Pollution control, Norwegian Institute for Nature Research, Norwegian Institute for Water Research and Swedish University of Agricultural Sciences. Ken Irvine (Trinity College of Dublin, Ireland) has been responsible for the REBECCA Lakes Macroinvertebrate group. The article benefited from constructive comments from Ramesh Gulati and an anonymous reviewer.

Appendix 1

Rules for assigning taxa (species/genera) sensitivities for development of species-based sensitivity groups (see Table 1).

1) A taxa should be used by at least two of the indices
2) For a taxa to be included and assigned to a tolerance class (highly sensitive (HS) or sensitive (S), tolerant (T) or highly tolerant (HT)), the indices have to have a positive agreement of two or more). For example, if taxa A is used by three indices and two define it as HS or S but one defines it as T or HT, then the taxa cannot be used (+1). If taxa B is used by four indices, three as HS or S and one as T, then the taxa can be used (+2).
3) If two indices use a taxon which is assigned as HS and S, respectively, then S is chosen using

the precautionary principle. This also applies to HT and T.

4) Point 3 does not apply when a genus is compared with species of the same genus. The default score being assigned to the lower taxonomic level.
5) Only those taxa that meet the above criteria can be used as basis for the development of 2nd and 3rd level indicators.

The Raddum, Medin and NIVA indices allocate taxa scores into four classes, easily translating into HS, S, T and HT. AWIC sp scores taxa from 1 to 9. After consideration the scores were assigned to the following classes 1-5 = HT, 6 & 7 = T, 8 = S, 9 = HS (J. Davy-Bowker, pers. comm.).

References

Arrhenius O (1921) Species and area. J Ecol 9:95–99

Barbour MT, Plafkin JL, Bradley BP et al (1992) Evaluation of EPA's Rapid Bioassessment benthic metrics: metric redundancy and variability among reference stream sites. Environ Toxicol Chem 11:437–449

Brodin Y-W (1995) Acidification of lakes and watercourses in a global perspective. In: Henrikson L, Brodin Y-W (eds) Liming of acidified surface waters. Springer Verlag, Berlin, pp 45–62

Bækken T, Kjellberg G (2004) Classification of acidity and assessment of acidification in running waters based on the presence of macroinvertebrates. Classification system developed for humic rivers and brooks in Eastern Norway. Report 4923, Norwegian Institute for Water Research (NIVA), Oslo, Norway, 13 pp (In Norwegian)

Dangles O, Malmqvist B, Laudon H (2004) Naturally acid freshwater ecosystems are diverse and functional: evidence from boreal streams. Oikos 104:149–155

Davy-Bowker J, Murphy JF, Rutt GP et al (2005) The development and testing of a macroinvertebrate biotic index for detecting the impact of acidity on streams. Arch Hydrobiol 163:383–403

Doberstein CP, Karr JR, Conquest LL (2000) The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. Freshw Biol 44:355–371

Driscoll CT, Baker JP, Bosogni JJ et al (1980) Effect of aluminium speciation on fish in dilute acidified waters. Nature 284:161–164

Driscoll CT, Newton RM, Gubala CO et al (1991) Adirondack Mountains. In: Charles IDF (ed) Acidic deposition and aquatic ecosystems. Regional case studies. Springer-Verlag, New York, pp 133–202

EMEP (2005) Transboundary acidification, eutrophication and ground level ozone in Europe. Norwegian Meteorological Institute, EMEP Status report 1/2005

EN 27828 (1994) Water quality – Methods for biological sampling – Guidance on handnet sampling of aquatic benthic macro-invertebrates

EN ISO/IEC 17025 (2005) General requirements for the competence of testing and calibration laboratories

European Commission (2000) Directive 2000/EC/60 of the European Parliament and the Council establishing a framework for community action in the field of water policy

European Commission (2004) Overview of common Intercalibration types, final version 5.1, dated 23 April 2004. http://forum.europa.eu.int/Public/irc/env/wfd/library

Fjellheim A, Raddum GG (1990) Acid precipitation: biological monitoring of streams and lakes. Sci Total Environ 96:57–66

Furse M, Hering D, Moog O et al (2006) The STAR project: context, objectives and approaches. Hydrobiologia 566:3

Gensemer RW, Playle RC (1999) The bioavailability and toxicity of aluminium in aquatic environments. Crit Rev Environ Sci Technol 29:315–450

Goedkoop W, Johnson R, Spånberg E (2000) The importance of sampling effort for the assessment of ecological quality using macroinvertebrates. Verh Internat Verein Limnol 27:326–332

Gunn JM, Sandøy S (2003) Introduction to the Ambio special issue on biological recovery from acidification: Northern Lakes Recovery Study. Ambio 32:162–164

Hageby A, Jr Petersen RC (1988) Effects of low pH and humus on the survivorship, growth and feeding of *Gammarus pulex* (L.) (Amphipoda). Freshw Biol 19:235–247

Henrikson L, Medin M (1986) Biologisk bedömning av försurningspåverkan på Lelångens tillflöden och grundområden 1986. Aquaekologerna, Rapport till Länsstyrelsen i Älvsborgs län (In Swedish)

Hering D, Feld CK, Moog O et al (2004) Cook book for the development of a multimetric index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. Hydrobiologia 566:311–324

Hesthagen T, Fiske P, Skjelkvåle BL (2008) Critical limits for acid neutralizing capacity of brown trout (*Salmo trutta*) in Norwegian lakes differing in organic carbon concentrations. Aquat Ecol (in press) doi: 10.1007/s10452-008-9191-x

Johnson RK, Goedkoop W (2007) Assessment criteria using benthic invertebrates in lakes and streams – User's manual and background documentation. Dept. Environmental Assessment, Swedish University of Agricultural Sciences, Report 2007:4 (In Swedish)

Koenker R (2006) quantreg: Quantile Regression. R package version 3.89. http://www.r-project.org

Larsen J, Birks HJB, Raddum GG et al (1996) Quantitative relationships of invertebrates to pH in Norwegian river systems. Hydrobiología 328:57–74

Larssen T, Lydersen E, Tang D et al (2006) Acid rain in China. Environ Sci Technol 40:418–425

Laudon H, Poléo ABS, Vøllestad LA et al (2005) Survival of brown trout during spring flood in DOC-rich streams in northern Sweden: the effect of present acid deposition and modelled pre-industrial water quality. Environ Pollut 135:121–130

Li J, Herlihy A, Gerth W et al (2001) Variability in stream macroinvertebrates as multiple spatial scales. Freshw Biol 46:87–97

Lydersen E, Larssen T, Fjeld E (2004) The influence of total organic carbon (TOC) on the relationship between acid neutralizing capacity (ANC) and fish status in Norwegian lakes. Sci Total Environ 326:63–69

Moe J, Dudley B, Ptacnik R (2008) REBECCA databases: experiences from compilation and analyses of monitoring data from 5000 lakes in 20 European countries. Aquat Ecol (in press) doi: 10.1007/s10452-008-9190-y

Moss B, Stephen D, Alvarez C et al (2003) The determination of ecological status in shallow lakes–a tested system (ECOFRAME) for implementation of the European Water Framework Direcitve. Aquatic Conserv Mar Freshw Ecosyst 13:507–549

R Development Core Team (2006) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Raddum GG, Fjellheim A (1984) Acidification and early warning organisms in freshwater in western Norway. Verh Int Verein Limnol 22:1973–1980

Raddum GG, Rosseland BO, Bowman J (1999) Workshop on biological assessment and monitoring; evaluation and models. NIVA-report 4091–99, ICP Waters report 50/99, 96 pp

Resh VH (1979) Sampling variability and life history features: basic considerations in the design of aquatic insect studies. J Fish Res Board Canada 36:290–311

Rosenberg DM, Resh VH (1993) Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York, p 488

Sandin L, Johnson RK (2000) The statistical power of selected indicator metrics using macroinvertebrates for assessing acidification and eutrophication of running waters. Hydrobiologia 422:233–243

Sandin L, Dahl J, Johnson RK (2004) Assessing acid stress in Swedish boreal and alpine streams using benthic macroinvertebrates. Hydrobiologia 516:129–148

Solimini AG, Free G, Donohue I et al (2006) Using benthic macroinvertebrates to assess ecological status of lakes. Current knowledge and way forward to support WFD implementation. EUR 22347 EN, Joint Recearch Centre, Inst for Environ and Sustain, Luxembourg, 46 pp

Schindler DW, Kasian SEM, Hesslein RH (1989) Losses of biota from American aquatic communities due to acid rain. Environ Monit Assess 12:269–285

Skjelkvåle BL, Mannio J, Wilander A et al (2001) Recovery from acidification of lakes in Finland, Norway and Sweden 1990–1999. Hydrol Earth Syst Sci 5(3):327–337

Skjelkvåle BL, Stoddard J, Jeffries D et al (2005) Regional scale evidence for improvements in surface water chemistry 1990–2001. Environ Pollut 137:165–176

Stoddard JL, Jeffries DS, Lükewille A et al (1999) Regional trends in aquatic recovery from acidification in North America and Europe. Nature 401:575–578

Warwick RM (1993) Environmental impact studies on marine communities: Pragmatic considerations. Aust J Ecol 18:63–80

Witters HE, Van Puymbroeck S, Vangenechten JHD et al (1990) The effect of humic substances on the toxicity of aluminium to adult rainbow trout, *Oncorhynchus mykiss* (Walbaum). J Fish Biol 37:43–53

Wood SN (2006) Generalized additive models: an introduction with R. Chapman and Hall/CRC

Økland J (1980) Environment and snails (Gastropoda): studies of 1000 lakes in Norway. In Drabløs D, Tollan A (eds) Ecological impact of acid precipitation, Proc Int Conf, Sandefjord, Norway, March 11–14 1980. SNSF Project, pp. 322–323