Article (refereed) - postprint

Dudley, Bernard; Dunbar, Michael; Penning, Ellis; Kolada, Agnieszka; Hellsten, Seppo; Oggioni, Alessandro; Bertrin, Vincent; Ecke, Frauke; Søndergaard, Martin. 2013. **Measurements of uncertainty in macrophyte metrics used to assess European lake water quality**. *Hydrobiologia*, 704 (1). 179-191. 10.1007/s10750-012-1338-z

Measurements of uncertainty in macrophyte metrics used to assess European lake water quality

Authors: Bernard Dudley[1], Michael Dunbar[2], Ellis Penning[3], Agnieszka Kolada[4], Seppo Hellsten[5], Alessandro Oggioni[6], Vincent Bertrin[7], Frauke Ecke[8, 9], Martin Søndergaard[10]

1. Centre for Ecology & Hydrology, Bush Estate, Penicuik, EH26 0QB, United Kingdom.

2. Centre for Ecology & Hydrology, Wallingford, Oxon, OX10 8BB, United Kingdom.

3. Deltares, P.O. Box 177, 2600 MH, Delft, The Netherlands

4. Institute of Environmental Protection – National Research Institute, Poland

5. Finnish Environment Institute - SYKE, P.O.Box 413, FIN-90014 University of Oulu, Finland

6. Institute for Electromagnetic Sensing of the Environment CNR – IREA, Via Bassini, 15, 20133 Milano, Italy

7. Irstea, UR REBX. 50 avenue de Verdun. FR-33612 CESTAS Cedex, France

8. Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, P.O. Box 7050, SE-750 07Uppsala, Sweden

9. Division of Geosciences and Environmental Engineering, Landscape Ecology Group, Luleå University of Ecology, SE-971 87 Luleå, Sweden

10. Institute of Bioscience, Aarhus University, Vejlsoevej 25, 8600 Silkeborg, Denmark

Abstract

Uncertainty is an important factor in ecological assessment, and has important implications for the ecological classification and management of lakes. However, our knowledge of the effects of uncertainty in the assessment of different ecological indicators is limited. Here, we used data from a standardized campaign of aquatic plant surveys, in 28 lakes from 10 European countries, to assess variation in macrophyte metrics across a set of nested spatial scales: countries, lakes, sampling stations, replicate transects, and replicate samples at two depth zones. Metrics investigated in each transect included taxa richness, maximum depth of colonisation and two indicators of trophic status: Ellenberg's N and a metric based on phosphorus trophic status. Metrics were found to have a slightly stronger relationship to pressures when they were calculated on abundance data compared to presence/absence data. Eutrophication metrics based on helophytes were found not to be useful in assessing the effects of nutrient pressure. These metrics were also found to vary with the depth of sampling, with shallower taxa representing higher trophic status. This study demonstrates the complex spatial variability in macrophyte communities, the effect of this variability on the metrics, and the implications to water managers, especially in relation to survey design.

Introduction

Managers of water bodies in Europe are required to assess the water quality of lakes under the terms of European legislation adopted in 2000: the Water Framework Directive 2000/60/EC (WFD). This assessment must be conducted in terms of *biological quality elements* (BQEs), which include *macrophytes* (aquatic plants). BQEs are intended to describe subsets of the biological community, which have inherently highly complex and variable distributions, causing uncertainty in their use as biological indicators (eg Capers et al., 2010). Consequently, creating reliable assessment methods for these BQEs has been a major challenge for the monitoring authorities across Europe since the adoption of the WFD (Poikane et al., 2011). Generally, assessment methods condense the taxonomic and distributional information gained from macrophyte surveys into *metrics*, which are usually designed to reflect water quality in terms of the water's biota (eg Penning et al., 2008a; Birk and Willby 2010).

The WFD requires that estimates of confidence and precision be included in the assessment of the status of lake biological quality elements. Understanding the effect of sampling variation and other sources of uncertainty on the ecological status class assessment and underlying metrics is essential in providing these estimates. Sources of uncertainty include natural spatial and temporal variation, sampling methodology and modelling of reference conditions (Clarke & Hering, 2006). For macrophyte status assessment the sampling methodology is an important source of uncertainty. Standardised, objective, and repeatable monitoring methods are essential in monitoring programs with aims to detect anthropogenic impact on lake ecosystems. Results of macrophyte surveys are

extremely sensitive to both vertical and horizontal variability of macrophyte communities (Jensen 1977, Janauer 2002, Hurford 2010). In addition to spatial variability there are potential errors related to recognition and identification of individual species and also especially to abundance estimations of vegetation.

Previous work on running waters in the EU STAR (Standardisation of River Classifications) project showed that inter-surveyor differences were low and the influences of temporal variation (years and seasons) and shading slightly stronger (Clarke & Hering, 2006). The strongest variation was due to habitat modifications, but several metrics were of sufficient precision in terms of sampling uncertainty to be useful for estimating the ecological status of rivers (Staniszewski et al., 2006). However, the probability of misclassification of a site was found to be largely associated with classification methodology (Szoszkiewicz et al., 2007; 2009).

Assessment of variability in macrophyte assemblages across a range of habitats cannot be adequately performed without accounting for the known natural and anthropomorphic determinants of those communities. For example, it is known that community structure is strongly related to alkalinity due to carbon uptake chemistry; high alkalinity waters are suitable for species that can utilise bicarbonate ion as a carbon source instead of, or as well as, carbon dioxide (Vestergaard & Sand-Jensen, 2000). Additionally, alkalinity affects nutrient availability, by both reducing the decomposition inhibiting effects of acidification, and by providing bicarbonate ions which can compete with orthophosphate in bonding with cations (Smolders, 2006). The consequence of this is that alkalinity, eutrophication and aquatic macrophyte richness are closely linked, with

eutrophication and alkalinity having a positive association (eg Penning et al., 2008b), and species richness having a hump-back distribution in relation to both nutrient availability and alkalinity (e.g. Murphy, 2002). Similarly to trophic status, water colour related to humic substances also affects the number of available habitats, which are also determined by lake area and length of the shoreline (Rørslett, 1991; Murphy, 2002). Additionally, macrophyte richness can be determined by altitude (Jones et al., 2003), latitude (Heino & Toivonen, 2008) and available routes for dispersal and the regional species pool (Heegaard, 2004).

Aquatic macrophyte metrics are often used to describe plant community responses to environmental pressure (e.g. Penning et al. 2008b). In Europe, water managers are required to assess water quality in terms of macrophyte community, and metrics are a tool used to summarise the effects of human pressure. Eutrophication pressure is usually described by using total phosphorus content of water as a proxy. This has lead to development of large numbers of phosphorus sensitive indicators used in assessment and especially in the Europe-wide WFD common intercalibration exercise (Poikane et al., 2011; Birk et al., 2012). In this study we have used some of the very few known metrics that can be applied to aquatic macrophyte data collected across Europe.

This study aimed to assess the relative importance of different sources of variation in the sampling data on uncertainty in the available metrics. The general aim of this study was to assess uncertainty in various macrophyte metrics, which might be used in assessing status of this BQE in lakes. This has been achieved by using several sources of information, but primarily a dataset collected as part of the EU WISER (Water bodies in Europe: Integrative Systems to assess Ecological status

and Recovery) project using a common standardized sampling method from 28 lakes in 10 European countries.

In our study, we focussed on four research questions. Firstly, we assessed how qualitative versus quantitative (presence-absence versus abundance) data affect metric results and their uncertainty. Secondly, we analysed how the inclusion or exclusion of helophytic taxa affect the results of the metric. Thirdly, we assessed the uncertainty related to surveying only the 0-1 m depth zone compared to surveying the whole depth range of potentially colonized area. Finally, we evaluated the variability of the different metrics between lakes, within a lake, and between transects. All four questions are relevant to macrophyte assessment methodologies in Europe (Penning et al., 2008a; 2008b; Kolada et al., 2011).

A further, practical aim of this work was to give recommendations on appropriate sampling design and analysis methods that are most likely to reduce uncertainty in the assessment of the status of lake macrophytes.

Methods

Data collection

Customised field survey

A sampling campaign was conducted in the summer of 2009 when 28 lowland clear-water (non humic) lakes from 10 countries representing broad geographical and trophic gradients were selected for survey (Table 1). Lakes selected were between 0.3 and 7.2 km^2 in surface area, below 250 m altitude and had a mean

depth between 3.8 and 18 m. Lakes were selected to represent a range of eutrophication pressures and a range of geographical types. Within each lake, six stations evenly distributed along a shoreline were identified (the first assigned arbitrarily, and the other five at regular intervals around the shore). Within each station three parallel transects were surveyed by boat, each being 5 m from its neighbour and each starting at the shore and proceeding towards the centre of the lake (Fig. 1). Transects followed straight lines as closely as was practicable. Each transect was divided into depth zones of 1 m depth intervals down to the limit of macrophyte colonisation and in each depth zone five randomly selected macrophyte sampling sites were used. At each sampling site a single sample was gathered from a rake dragged along the bottom for approximately 2 m, and supplemented by observation through a bathyscope, where this was possible. In each sample all truly aquatic species, and a pre-defined selection of emergent taxa, were identified and recorded. Identifications were performed by experienced field surveyors in all cases and uncertain specimens were referred to taxonomic experts.

Additional to the aquatic macrophyte survey data, surface water samples were collected from a central station of each lake at least twice during the growing season and alkalinity and total phosphorus (TP) were measured from these samples. Averages of these measurements, from each lake, were used in later statistical analyses.

Data collected during the field campaign were compiled into a database format, maintaining the hierarchical structure of the data in a form analogous to the sampling design. In this format, each observation of a macrophyte taxon in a

sample was given a separate record. Data were extracted from the database at various levels. These levels were depth-zone within transect, transect, station, and whole lake. At the lowest level, a taxon was present if it was recorded at least once. Abundance at each level was measured as relative point frequency, which for each taxon was the number of observations of the taxon, divided by the total number of observations of all taxa at that level.

Data analyses

Exploratory multivariate analyses

A multivariate clustering analysis using group averaging was performed for quick exploration of the data available within the dataset in order to assess whether unexpected subsets of data could be distinguished, using the taxonomic composition of the samples, that were linked more to country or location than to the environmental variables of interest (TP and alkalinity). Species abundance data were averaged per lake and analyses were performed using the statistical software programme PRIMER6. A similarity matrix was calculated using the Bray-Curtis Similarity index on the non-transformed abundance data. Using this similarity matrix, a dendrogram was plotted using group average to visualize specific subgroups of data.

Calculation of metrics

Taxon-specific trophic rank scores, also known as Intercalibration Metric scores and referred to in this report as *ICM-LM scores* (Intercalibration Common Metric for Lake Macrophytes), were supplied by Nigel Willby, of the University of Stirling, United Kingdom. These scores have been used in the Water Framework Directive Intercalibration Exercise for lake macrophyte BQEs as a means of

comparing lake macrophyte status across Member States, where sampling methods and metric derivations are not consistent. For submerged aquatic plant taxa, scores were derived using methods similar to those used by the United Kingdom Technical Advisory Group (UKTAG, 2009), and Birk and Willby (2010). In general, scores were calculated by rescaling the median of the logarithms of the concentration of total phosphorus concentrations in lakes across Europe in which each taxon has been recorded as present. These scores are available in Kolada et al. (2011). Scores are combined across sites to form a metric, either as a simple mean, or by using some measure of abundance to weight the mean (see equation 1, below). The site metric is intended to be representative of the nutrient status of the water.

The ICM-LM scores were only available for real hydrophytes, as this metric was designed to be used with submerged taxa. *Ellenberg's Nitrogen* values for soil fertility (scores from 1 to 9; Ellenberg et al., 1991) were compiled for all taxa in the dataset in order to test the use of a metric with and without helophytes. We supplemented the original values with British values where original values were missing (Hill et al., 1999). Even with these supplements, there were 13 taxa, notably charophytes, for which no Ellenberg score was available, but for which an ICM-LM score was available. Scores for these taxa were inferred using a regression relationship between the ICM-LM and Ellenberg scores for all species with both values (Ellenberg = 0.22 + 0.79 × ICM-LM, n = 51, R^2 = 0.64). These modified taxon-specific Ellenberg scores were then used to calculate an average Ellenberg-N metric per lake. Taxa for which neither Ellenberg nor ICM-LM scores were available were excluded from further analyses (details given in Results section).

For each transect, where at least one taxon for which a score was known, ICM-LM and Ellenberg metrics were calculated from scores both as simple averages of the scores of the taxa found (unweighted), or as weighted averages,

$$M_w = \frac{\sum S_i \times A_i}{\sum A_i} \quad \text{.....eq. 1}$$

where M_w is the weighted metric, S_i is the score for taxon *i*, and A_i is the abundance of taxon *i*, for all taxa in the sample. In many cases, these metrics have been calculated for subsets of the macrophyte community, such as 'submerged only', or 'helophytes'. Because not all transects contained any taxa with scores, it was not always possible to calculate a metric for a transect. For some of the analyses described below this limitation applied to entire lakes, for example for lakes where no taxa were found below 1 m depth, it was not possible to examine the effect of depth zone.

Species richness was calculated as the number of macrophyte taxa identified at an individual sampling location. *Maximum depth of colonisation* (C_{max}) was determined as the greatest depth in which rooted plants were found at each transect.

Uncertainty assessments

The WISER lake macrophyte data were used to examine variability associated with the four levels of the nested sampling scheme: transects, stations, lakes and countries. We assessed this for each of the response metrics described above. Values for each metric were calculated for each transect, the derivation of metric values for progressively higher nesting levels in the sampling design (sampling

stations, lakes, countries), being derived implicitly as part of the model fitting process. We fitted linear mixed effects models (nlme package; Pinheiro and Bates, 2000) to the dataset, using the R environment for statistical computing (R Development Core Team, 2011). The levels of the sampling hierarchy were specified as nested random effects, with the lowest level, variation between transects, forming the residual. As a measure of uncertainty, we used standard deviation, as provided by the nlme package, to calculate absolute and relative variance at each of the nested levels of the overall sampling design.

We used lake level alkalinity and TP concentration as covariates in the analyses for two reasons. Firstly these variables define strong gradients in the dataset (Table 2). Secondly, the response of the macrophyte metrics to lake-level TP, accounting for variations in alkalinity, and for uncertainty in surveying lakes, is of considerable interest in itself (Penning et al., 2008a; 2008b). As these covariates were measured at the lake level they explain variance between lakes and countries, but not within lakes. As alkalinity information for Étang des Aulnes (France) was not available, data from this lake were excluded from further analyses. For the analyses, values for TP and alkalinity were log transformed. Metric response values were un-transformed except for species richness, which was square-root transformed prior to model fitting.

Models were fitted using Residual Maximum Likelihood Estimation (REML) to produce unbiased estimates of random effect variances, but any comparison of models differing in their fixed effects was undertaken using Akaike's Information Criterion and models fitted by standard Maximum Likelihood.

To assess how qualitative versus quantitative (presence-absence versus abundance) data affect metric results and their uncertainty, we compared variance components from models using metrics calculated from presence/absence data with those using metrics calculated from scores weighted by point frequency. To assess how the inclusion or exclusion of helophytic taxa affect the results of metrics, we compared the results of models run using all species, with those run using only submerged and floating plants. To assess the uncertainty related to surveying only the 0-1 m depth zone compared to surveying the whole depth range of potentially colonized area, each transect was divided into two depth zones, above and below 1 m of water depth. Metrics were calculated at the transect level for both depth zones and the variance components, as well as the response to the covariates from both models were compared.

This study does not address the effects of probability of misclassification of water bodies in status classes as common status boundaries have not yet been defined for the metrics used in this study.

Results

There were 124 plant taxa recorded from the 28 lakes surveyed. 15 taxa, including filamentous algae, an undefined moss, woody species (for example *Alnus* sp. and *Salix* sp.) and some taxa recorded at genus or higher level, were excluded from all analyses. The remaining 109 taxa included 101 recorded at species level and 8 at genus level (*Callitriche*, *Chara*, *Fontinalis*, *Mentha*, *Nitella*, *Nymphaea*, *Sparganium* and *Utricularia*). There were 10 taxa for which neither an Ellenberg-N, nor an ICM-LM score was available, and were therefore not used in the calculation of metrics. None of these 10 taxa occurred in more than 2 lakes.

Exploratory multivariate analyses

Three main groups of lakes were distinguished from the similarity analyses (Fig. 2): 1. mainly higher alkalinity central European lakes (France, Estonia, Poland, Germany, Denmark), with two more eutrophic, moderate alkaline lakes from the Northern GIG (United Kingdom, Finland), 2. a small group of higher altitude lakes (all Italian lakes and two Norwegian lakes), and 3. Nordic moderate and low alkaline lakes (Finland, Sweden, Norway, United Kingdom). Similarity between lakes was never more than 60%. There was one outlying lake, Skirösjön in Sweden. Only one submerged species was recorded from this lake so it was considered to fall outside the three defined groups.

Qualitative versus quantitative macrophyte data

Unweighted and abundance weighted LCM-LM metrics for lakes were highly correlated (Table 2, Fig. 3). This correlation was highest at the lake scale, and progressively lower as one moves to the finer scales of station and transect within the lake. Compared to unweighted ICM-LM, the abundance weighted ICM-LM gave a steeper (0.46 vs 0.38) but slightly less precise (standard error of 0.30 vs 0.28) association with TP, while the response to alkalinity was similar (Table 2). The unweighted ICM-LM shows greater variance at the lake scale than weighted ICM.

Helophytic taxa

There was a weak relationship between Ellenberg metrics calculated on weighted averages of submerged taxa only and the same metric calculated on helophytes only. When calculated at the lake scale, metrics based on Ellenberg scores for

helophytes only had a much smaller range than their counterparts based on scores for submerged species (Fig. 4). Residual correlations between Ellenberg scores calculated for submerged taxa only and for helophytic taxa only were relatively weak (Table 3). This is likely because there is no overlap in the taxa used to calculate the alternative metric formulations. Furthermore, the helophyte metrics had weaker relationships with both pressure variables total phosphorus (TP; results not shown) and Alkalinity (Table 3).

0-1 m depth zone versus whole depth range

ICM-LM metrics calculated from the scores of plants found in depths greater than 1 m were lower than corresponding metrics calculated from shallower water (less than 1 m) plants (intercept of model at 4.51 vs 5.05), indicating that, on average, species found in the shallow zone have higher trophic status (Table 4).

ICM-LM metrics calculated from deeper (> 1 m) taxa were less variable between stations within lakes but marginally more variable at the station and transect scales (Table 4). There was a fairly high residual correlation between ICM-LM for the different depth zones at the lake (0.79) and station (0.67) scales, but low correlation at the transect scale (0.1; Table 4). Hence variation between depth zones at the finer spatial scale (transect) tends to be averaged out between transects within stations and between stations within lakes.

Variability of metrics between lakes, within a lake, and between transects

For all metrics, the proportion of variance at the transect level was much smaller (generally around half) than at the station level (Table 5). The proportion of variance at the country and lake sampling levels was dependent on the metric

used, and on whether the explanatory driver variables were included in the model. Except for the Richness metric, inclusion of the explanatory variables always reduced the *between-lake* (country and lake) proportion of variance, mostly by reducing the variance at the country level. ICM-LM, compared to Ellenberg, illustrated a slightly higher proportion of variance between lakes, with correspondingly less variance within lakes. Maximum growing depth also behaved similarly to ICM-LM, although the covariates appeared slightly more successful in explaining between-lake variance. The Richness metric followed a different behaviour; introduction of the covariates reduced the variance between lakes but accentuated the variance between countries. Total between-lake variance remained roughly constant (Table 5).

Although inclusion of the explanatory variables, TP and Alkalinity, reduced variance in the models, their relationships to the metrics were not always significant at the traditional ($p < 0.05$) level (Table 6). Alkalinity showed strong relationships with all metrics ($p < 0.01$) except Richness. Relationships between TP and metrics were always in the expected direction (higher TP corresponded to higher ICM-LM and Ellenberg, and lower C_{max} and Richness), but for both ICM-LM and Ellenberg the relationships were relatively imprecise. This general pattern was confirmed through re-fitting models using maximum likelihood (ML) estimation and comparison of Akaike's Information Criteria (AIC) values. The significant relationships between TP and both C_{max} and Richness metrics were notable. We re-fitted the C_{max} and richness TP/Alkalinity models to the subset of data with an ICM-LM score. The C_{max} -TP relationship was robust to this fitting to a smaller subset of the data, but the richness relationship was not. 49 transects had values for richness but not ICM-LM (meaning that plants were recorded from

these transects but these plants had no ICM-LM score, mostly because they were helophytes); these were spread across 12 lakes, the lake with the largest number of transects lost being Glindower See with 15 (only one non-helophyte taxon was recorded from this lake). For C_{max}, it is notable that the strong relationship with TP was entirely dependent on alkalinity also being in the model. Without the inclusion of alkalinity as a co-variate, the C_{max} -TP relationship was very weak (results not shown).

Discussion

This study assesses the complex spatial variability in macrophyte communities, the effect of this variability on various plant metrics, and the implications to water managers, especially in relation to appropriate survey design. Although the study focuses on the assessment of aquatic macrophytes in European lakes, the results have implications for all BQEs in all WFD waterbody types, and indeed for any assessment of the quality of a biological community that uses a metric derived from taxonomic data, anywhere in the world.

Abundance-weighted metrics are preferable to metrics calculated from presence-absence data, but only when all sampling is done using the same methods.

In this study, metrics calculated as abundance-weighted means of taxon-specific scores provided a steeper relationship with the nutrient pressure (TP), and should therefore be considered better indicators of this pressure. It is arguable that abundance-weighted metrics should be used to assess ecological change because

they reflect changes in the abundance of taxa, which cannot be detected by metrics based only on presence or absence of taxa.

The results in this study are contrary to Penning et al. (2008a), who found that there was little evidence of benefit in using metrics calculated using mean scores weighted by abundance. In fact, Penning et al. (2008a) found that relationships became weaker when plant abundance was used to weight metric scores. In that study, data were from multiple sources and collected using disparate sampling and quantification methods. The abundance data in the Penning et al. (2008a) study had to be re-scaled to the lowest resolution abundance scale from within the collated datasets, which had only three possible values. This accounts for the observed relative imprecision associated with abundance weighting.

The use of helophytes in the calculation of metrics appeared to provide little additional information, and metrics based on helophytes do not respond as well to nutrient pressure (TP) as do the submerged species.

The data used in this study provide a stronger basis for these conclusions than has been previously available to answer this question, and these conclusions are consistent with Penning et al. (2008b). Helophytes are less affected by water quality than submerged plants as their environment is not sub-aquatic, so their response to eutrophication is obscured by soil trophic characteristics, exposure, shoreline management and especially water level fluctuation dynamics (e.g. Coops et al., 1994). It is possible that the use of large datasets collated from multiple sources will provide spurious answers to this question, as it is likely that bias in sampling is related to trophic status. In regions with lakes where the

submerged taxa are highly visible, flourishing and diverse, sampling effort will be concentrated on these plants, in contrast to regions where lakes are more eutrophic, so predominately have few submerged taxa, but a flourishing emergent community, where it is likely that sampling effort will concentrate on the helophyte taxa.

It should be noted that this assessment has been made on the basis that eutrophication pressure has been measured in terms of phosphorus, but the impact has been measured in terms of the Ellenberg metric, which is intended to reflect nitrogen status (Ellenberg et al., 1991). Unfortunately, neither a measure of eutrophication based on nitrogen, nor a measure of impact (for helophytes) based on phosphorus, was available for this study.

Surveying only shallow vegetation may result in a worse classification of the macrophyte BQE than surveying the entire depth range of plant colonisation.

Calculation of the ICM-LM metric from shallow (< 1 m) vegetation only resulted in a lower ecological status than calculation based on plants from the entire depth range of colonisation. Also, the ICM-LM metric calculated on macrophyte data from deeper samples (> 1 m) showed a steeper and more precise relationship with TP than the metric calculated from shallow samples. In this study, higher ICM-LM scores were obtained from shallow zone samples than from deeper zone samples. The shallow littoral zone is often affected by incoming ditches and also provides sheltered habitats for species preferring more nutrient rich conditions (Alahuhta et al. 2012). Deeper areas are also typical habitats for oligotrophy-indicating large isoetids in soft water lakes (Murphy, 2002). This has important

implications for the assessment of macrophyte status of lakes. If an assessment method uses only shore-based data (obtained by wading), it is likely to result in an assessment of condition that is worse and less precise than if the method used data from deep water as well (obtained by boat). Overall, assuming eutrophication from excess phosphorus is the stressor of prime interest, including macrophyte data from the full depth range is likely to give a more precise and less biased estimate of status.

Within lakes, metrics were twice as variable between stations as between replicate transects (5 m apart). Variance in metrics between lakes depends on the metric used, and on whether explanatory variables are included.

These results support the use of more sampling stations in macrophyte surveys to improve precision of macrophyte metrics, and show that sampling repeat transects at a station is less important. The results also illustrate that differences in the number of transects for which metrics may be calculated can have a strong influence on the results (Jensen, 1977). In particular, as TP levels increase, taxa richness decreases, but the number of taxa from which metrics such as ICM-LM can be calculated decreases even more rapidly. Increased imprecision of metrics associated with low richness of indicator taxa, and at the most extreme, non-calculability of such indices can have a significant influence on perceived metric performance. Therefore, to maintain the same degree of uncertainty, more sampling is required at either end of the trophic scale, when there is less vegetation to be sampled.

This study highlights the importance of alkalinity in the assessment of aquatic plants, especially when considering total phosphorus as an explanatory variable. TP and alkalinity are fairly well correlated in the dataset; hence it is not surprising that in some cases either variable on its own may show apparent relationships with metrics. In particular, in this dataset, for ICM-LM and Ellenberg, alkalinity was clearly the dominant explanatory variable, and although the partial relationships with TP were in the expected direction, they were less precise. Furthermore, the fact that the precision of the relationship between TP and C_{max} was conditional on alkalinity also being in the model is notable and highlights the inter-relatedness of these variables.

Recommendations for sampling, data analysis and assessment methods

This paper supports the following recommendations:

1. Assessment methods should include samples from the entire depth range of aquatic vegetation, as using only shallow samples can result in a worse assessment of trophic status.
2. Submerged taxa should be used in the assessment of the status of lake macrophyte communities. Helophytic taxa should not be used when assessing the effects of eutrophication pressure as they do not respond in the same way. Helophytes may still be useful in the assessment of hydromorphological pressure and general degradation.
3. Assessment of lake status should use data sampled from multiple stations around a lake.

4. In order to control metric uncertainty, more sampling is required in lakes where macrophytes are scarce or taxa richness is low. At these lakes, scores of individual taxa can have a much larger impact than in lakes with more macrophyte cover or more taxa.
5. Assessment methods should use quantitative data (not just presence/ absence) where possible, but only in cases where all data has been collected using the same methods.
6. Examination of uncertainty in metrics should not be undertaken in the absence of the relationships between metrics and stressors. In the worst case scenario, a metric may illustrate desirable properties of low variance within lakes relative to variance between lakes, but may have undesirably low response to stressors.

Acknowledgements

The authors are indebted to all of the surveyors involved in collecting the data presented here. This paper was written as part of the WISER project, which was funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change), contract No. 226273. The authors are also very grateful to the two reviewers, who gave very valuable comments that resulted in a much improved paper.

References

Alahuhta, J., A. Kanninen & K. M. Vuori, 2012. Response of macrophyte communities and status metrics to natural gradients and land use in boreal

lakes. Aquatic Botany. In press, http://dx.doi.org/10.1016/j.bbr.2011.03.031

Birk, S. & N.Willby, 2010. Towards harmonization of ecological quality classification: establishing common grounds in European macrophyte assessment for rivers. Hydrobiologia 652: 149–163.

Birk, S., W. Bonne, A. Borja, S. Brucet, A. Courrat, S. Poikane, A. G. Solimini, W. van de Bund, N. Zampoukas, D. Hering, 2012. Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators 18, 31-41

Capers, R. S., R. Selsky, & G. J. Bugbee, 2010. The relative importance of local conditions and regional processes in structuring aquatic plant communities. Freshw. Biol. 55, 952-966.

Clarke, R. T. & D. Hering, 2006. Errors and uncertainty in bioassessment methods – major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia 566:433–439.

Coops, H., N. Geilen & G. van der Velde, 1994. Distribution and growth of the helophyte species Phragmites australis and Scirpus lacustris in water depths gradients in relation to wave exposure. Aquatic Botany 48: 273-284.

Ellenberg, H., H. E. Weber, R. Düll, V. Wirth, W. Werner & D. Paulissen, 1991. Zeigerwerte von Pflanzen in Mitteleuropa. Scripta Geobotanica 18: 1-248.

Heegaard, E., 2004. Trends in aquatic macrophyte species turnover in Northern Ireland - Which factors determine the spatial distribution of local species turnover? Global Ecology and Biogeography 13: 397-408.

Heino, J. & H. Toivonen, 2008: Aquatic plant biodiversity at high latitudes: Patterns of richness and rarity in finnish freshwater macrophytes. - Boreal Environ.Res. 13: 1-14.

Hill, M. O., J. O. Mountford, D. B. Roy & R. G. H. Bunce, 1999. Ellenberg's indicator values for British plants. ECOFACT Volume 2 Technical Annex. Huntingdon, Institute of Terrestrial Ecology, 46pp. (ECOFACT, 2a). Available at http://nora.nerc.ac.uk/6411/

Hurford, C., 2010. Observer variation in river macrophyte surveys: the results of multiple-observer sampling trials on the Western Cleddau. Conservation Monitoring in Freshwater Habitats: A Practical Guide and Case Studies (eds C. Hurford, M. Schneider & I. Cowx), pp. 137-146.

Janauer, G. A., 2002. Water Framework Directive, European Standards and the Assesment of Macrophytes in Lakes: A Methodology for Scientific and Practical Application. Verh. Zoo. – Bot. Ges.Osterreich 139:143-147.

Jensen, S., 1977. An objective method for sampling the macrophyte vegetation in lakes. Vegetatio 33:107-118.

Jones, J. I., W. Li & S. C. Maberly, 2003. Area, altitude and aquatic plant diversity. Ecography 26: 411-420.

Kolada, A., S. Hellsten, M. Søndergaard, M. Mjelde, B. Dudley, G. van Geest, B. Goldsmith, T. Davidson, H. Bennion, P. Nõges & V.Bertrin, 2011. Report on the most suitable lake macrophyte based assessment methods for

impacts of eutrophication and water level fluctuations. Available at: www.wiser.eu

Murphy, K. J., 2002. Plant communities and plant diversity in softwater lakes of northern Europe. Aquatic Botany 73: 287–324.

Penning, W. E., B. Dudley, M. Mjelde, S. Hellsten, J. Hanganu, A. Kolada, M. van den Berg, S. Poikane, G. Phillips, N. Willby & F. Ecke, 2008a. Using aquatic macrophyte community indices to define the ecological status of European lakes. Aquatic Ecology 42:253–264.

Penning, W.E., M. Mjelde, B. Dudley, S. Hellsten, J. Hanganu, A. Kolada, M. van den Berg, S. Poikane, G. Phillips, N. Willby & F. Ecke, 2008b. Classifying aquatic macrophytes as indicators of eutrophication in European lakes. Aquatic Ecology 42:237–251.

Pinheiro, J. C., & D M. Bates. 2000. Mixed-Effects Models in S and S-PLUS. Springer, New York.

Poikane, S., M. van den Berg, S. Hellsten, C. de Hoyos, J. Ortiz-Casas, K. Pall, R. Portielje, G. Phillips, A. Lyche Solheim, D. Tierney, G. Wolfram & W. van de Bund. 2011. Lake ecological assessment systems and intercalibration for the European Water Framework Directive: aims, achievements and further challenges. Procedia Environmental Sciences 9: 153-168.

Rørslett, B., 1991. Principal determinants of aquatic macrophyte richness in northern European lakes. Aquatic Botany 39: 173-193.

R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Smolders, A. J. P., L. P. M. Lamers, E. C. H. E. T. Lucassen, G. van der Velde & J. G. M. Roelofs, 2006. Internal eutrophication: How it works and what to do about it – a review. Chemistry and Ecology 22(2): 93–111

Staniszewski, R., K. Szoszkiewicz, J. Zbierska, J. Lesny, S. Jusik & R.T. Clarke, 2006. Assessment of sources of uncertainty in macrophyte surveys and the consequences for river classification. Hydrobiologia 566:235-246.

Szoszkiewicz K., S. Jusik, T. Zgola, M. Czechowska & B. Hryc, 2007. Uncertainty of macrophyte-based monitoring for different types of lowland rivers. Belgian Journal of Botany 140(1) Special Issue: 7-16 International Symposium on Aquatic Vascular Plants. Vrije Univ, Brussels, Belgium, January 11-13, 2006.

Szoszkiewicz K., J. Zbierska, & R. Staniszewski, 2009. The variability of macrophyte metrics used in river monitoring. Oceanological and Hydrobiological Studies 38(4)117-126.

UKTAG (2009). UKTAG Lake Assessment Methods: Macrophyte and Phytobenthos. Macrophytes (Lake LEAFPACS). Water Framework Directive - United Kingdom Technical Advisory Group (WFD-UKTAG). Available at: http://www.wfduk.org/bio_assessment/

Vestergaard, O. & K. Sand-Jensen, 2000. Aquatic macrophyte richness in Danish lakes in relation to alkalinity, transparency, and lake area. Canadian Journal of Fisheries and Aquatic Sciences 57: 2022-2031.

Tables

Table 1. List of lakes surveyed for macrophytes in 2009. GIG (WFD Geographic Intercalibration Group) Regions are Central-Baltic (CB), Nordic (N) and Mediterranean (Med). Alkalinity (Alk.) types are Low (< 0.2 meq L-1), Medium (≥ 0.2 and < 1 meq L-1), and High (≥ 1 meq L-1). Provisional (Prov.) status refers to local assessment of the macrophyte biological element of either High (H), Good (G), Moderate (M), Poor (P), or Bad (B). Eutrophication (Eutro.) and Hydromorphological (Hymo.) pressures were subjective expert judgements of surveyors, of low, medium or high pressure. Where information was not available this is denoted with a dash.

Country	Lake Name	GIG Region	Alk. Type	Prov. Status	Eutro. pressure	Hymo. pressure
Germany	Roofensee	CB	High	H/G	Low	Low
	Grienericksee	CB	High	G/M	Medium	Medium
	Glindower See	CB	High	P/B	High	Medium
Denmark	Fussingsø	CB	High	G	Medium	-
	Nordborgsø	CB	High	G	High	-
Estonia	Saadjärv	CB	High	H/G	Low	Low
	Viljandi	CB	High	G/M	Low	Medium
Poland	Kiełpińskie	CB	High	G	Medium	Low
	Rumian	CB	High	M	Medium	Low
	Lidzbarskie	CB	High	P/B	High	Low
United Kingdom	Rostherne Mere	CB	High	P/B	High	Low
	Loweswater	N	Medium	M	Medium	Low
	Grasmere	N	Medium	M	Medium	Low
Finland	Sääksjärvi	N	Medium	G/M	Low	Low
	Vuojärvi	N	Medium	M/P	High	Low
	Iso-Jurvo	N	Low	H/G	Low	Low
Norway	Nøklevann	N	Low	H	Low	Low
	Longumvatnet	N	Medium	G/M	Medium	Low
	Temse	N	Medium	M	Medium	Low
Sweden	Västra Solsjön	N	Low	H	Low	Low
	Fiolen	N	Low	M	Medium	Low
	Skirösjön	N	Medium	P	High	Medium
France	Aulnes (étang des)	Med	High	M	High	Low
	Salagou (lac du)	Med	High	M	Medium	Medium
Italy	Segrino	Med	High	H/G	Low	Medium
	Lago di Monate	Med	Medium	G	Medium	Medium
	Candia	Med	Medium	G/M	Medium	Low
	Alserio	Med	High	M/P	High	Low

Table 2. Parameters for multivariate model of transect-level abundance-weighted and presence-absence ICM-LM metrics as a function of alkalinity and total phosphorus (TP), including their covariance at all levels of the model. Random effects for abundance weighted and presence absence presented as standard deviations, fixed effects at the lake scale are presented as parameter value with standard error in brackets. Note that the unit of observation is the transect, hence there is an implicit weighting at the station and lake scale based on the number of times a taxon was observed.

	Abundance Weighted	**Presence-absence**	**Correlation**	**Number of observations**
Random effects				
Lake	0.84	0.78	0.99	22
Station	0.75	0.72	0.94	113
Transect	0.49	0.50	0.84	634
Fixed effects (lake level)				
Intercept	4.57 (0.92)	4.77 (0.86)		
alkalinity	0.71 (0.22)	0.69 (0.20)		
TP	0.46 (0.30)	0.38 (0.28)		

Table 3. Parameters for multivariate model of Ellenberg metrics based on only helophyte scores and only submerged taxa scores, as a function of alkalinity, including their covariance at all levels of the model. Random effects for the two metrics presented as standard deviations, fixed effects presented as parameter value with standard error in brackets.

	Submerged	**Helophytes**	**Correlation**	**Number of observations**
Random effects				
Lake	0.91	0.46	-0.31	27
Station	0.69	0.85	-0.01	149
Transect	0.51	0.55	0.03	661
Fixed effects				
Intercept	5.34 (0.19)	6.19 (0.13)		
alkalinity	1.15 (0.14)	0.36 (0.10)		

Table 4. Parameters for multivariate model of ICM-LM (abundance weighted for submerged species only) for depth zone <1m and >1m vs alkalinity and TP. Random effects for depth zone presented as standard deviations, fixed effects presented as parameter value with standard error in brackets

	< 1m	> 1m	Correlation	Number of observations
Random effects				
Lake	0.93	0.80	0.79	22
Station	0.76	0.79	0.67	113
Transect	0.54	0.50	0.10	529
Fixed effects				
Intercept	5.05 (1.03)	4.51 (0.90)		
alkalinity	0.72 (0.24)	0.68 (0.21)		
TP	0.33 (0.33)	0.49 (0.29)		

Table 5. Proportions of variance at different levels of the sampling strategy for four different metrics and two formulations of the model: with and without TP/alkalinity

Metric	**Model**	**Country**	**Lake**	**Station**	**Transect**	**Total Between Lake**	**Total Within Lake**
ICM-LM	Null	0.11	0.61	0.19	0.08	0.72	0.28
	TP + Alk	0.00	0.47	0.37	0.16	0.47	0.53
Ellenberg	Null	0.31	0.42	0.18	0.08	0.74	0.26
	TP + Alk	0.00	0.41	0.40	0.19	0.41	0.59
C_{max}	Null	0.39	0.31	0.21	0.08	0.70	0.30
	TP + Alk	0.01	0.38	0.44	0.17	0.39	0.61
Richness	Null	0.18	0.19	0.45	0.18	0.37	0.63
	TP + Alk	0.28	0.10	0.44	0.18	0.38	0.62

Table 6. Significance (p-values) for approximate tests for TP and alkalinity fixed effects for models for each metric in Table 5, and numbers of samples at each level of the model

Metrics	**TP**	**Alkalinity**	**Country**	**Lake**	**Station**	**Transect**
ICM-LM	0.144	0.007	8	22	113	317
Ellenberg	0.115	0.002	8	22	123	360
C_{max}	0.001	<0.001	8	18	100	282
Richness	0.027	0.191	8	22	125	366

Figures

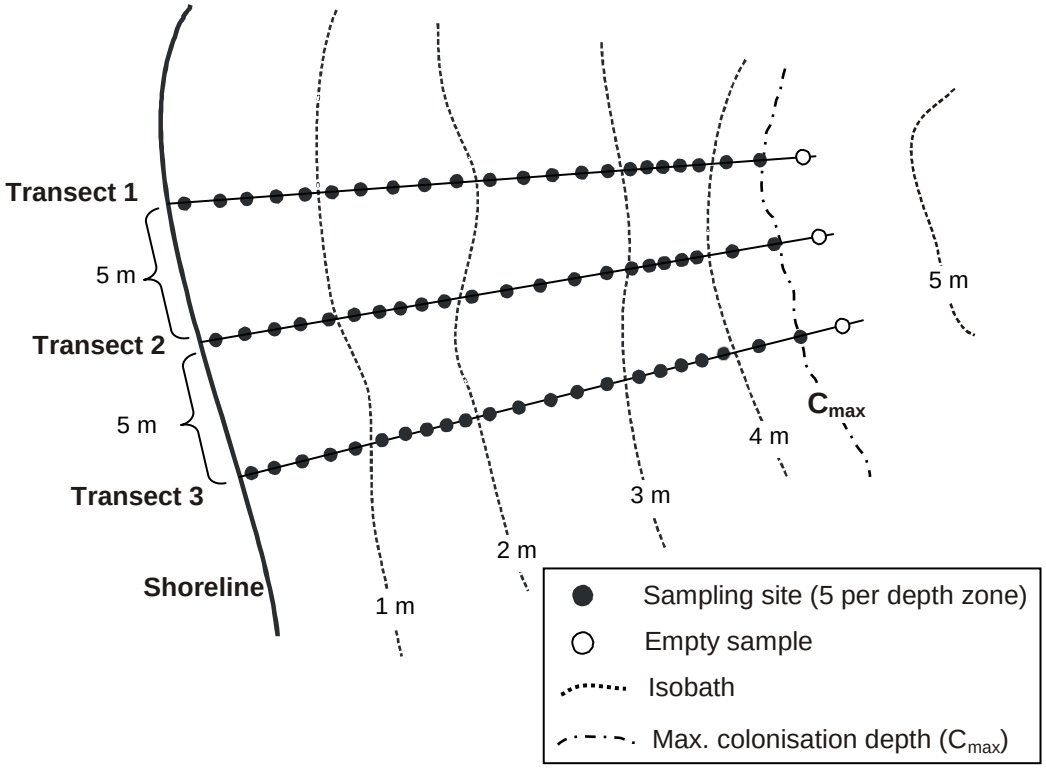

Fig. 1 Idealised sampling design used in the common field sampling protocol, employed in 2009. Three transects at one of six stations are shown.

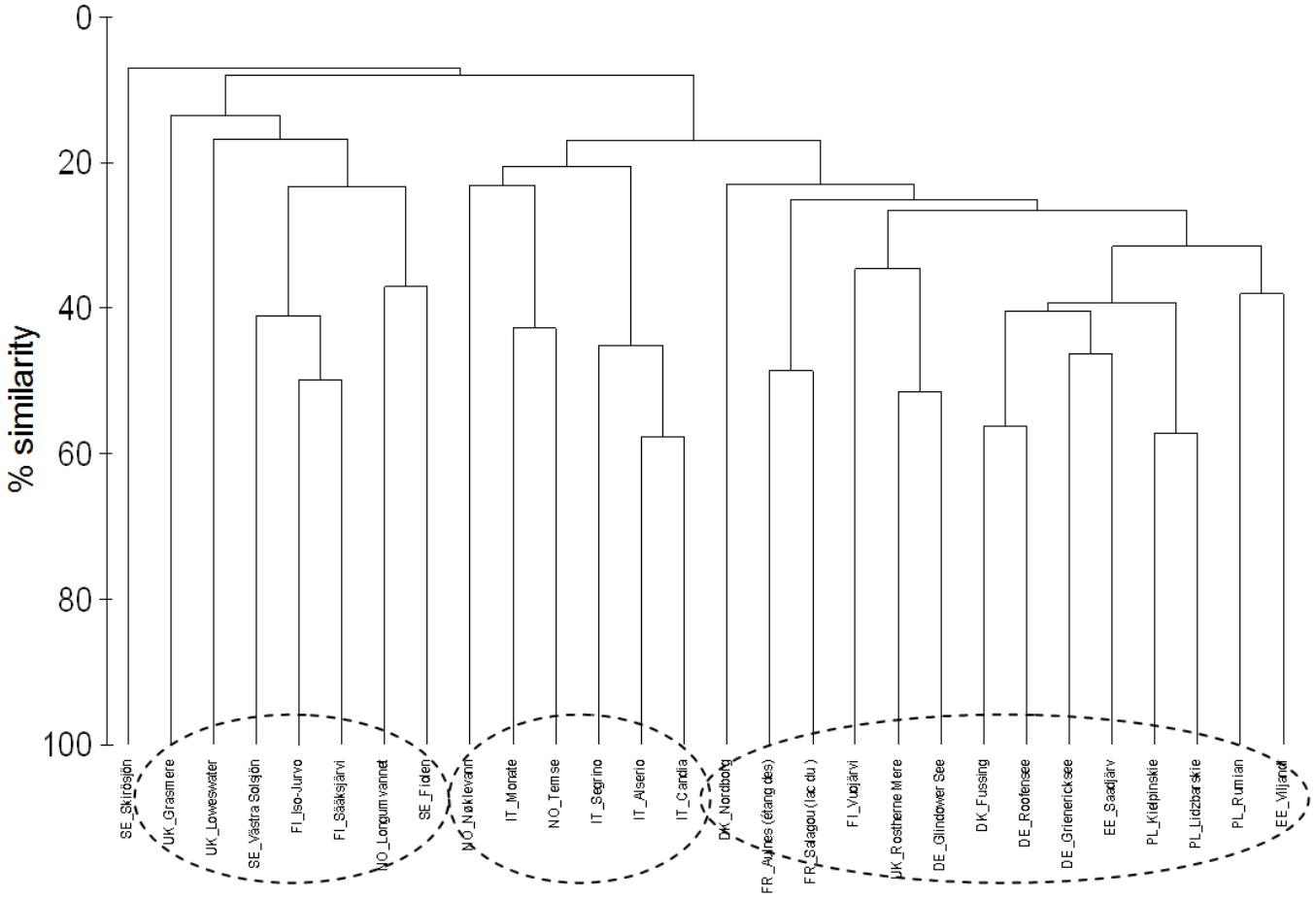

Fig. 2 Hierarchical clustering of the lakes (country code followed by lake name), based on the abundance of aquatic plants, showing three main groups outlined in dashed ovals

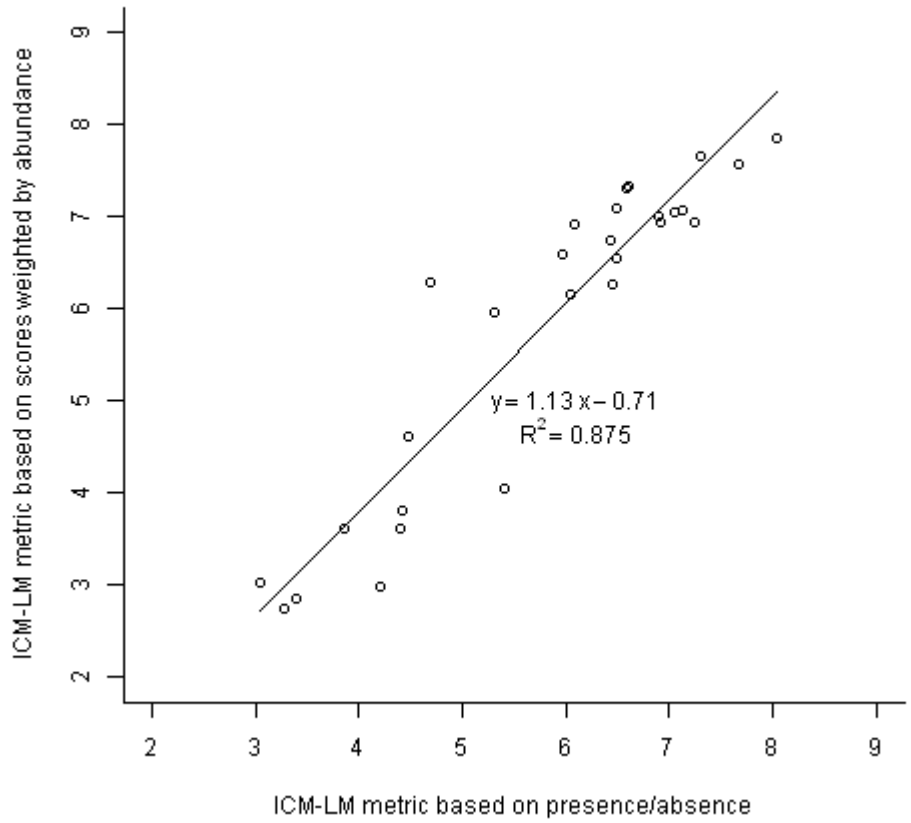

Fig. 3 Comparison of ICM-LM unweighted metrics (based on presence/absence data) and metrics weighted by abundance, calculated at the lake level for submerged and helophyte taxa

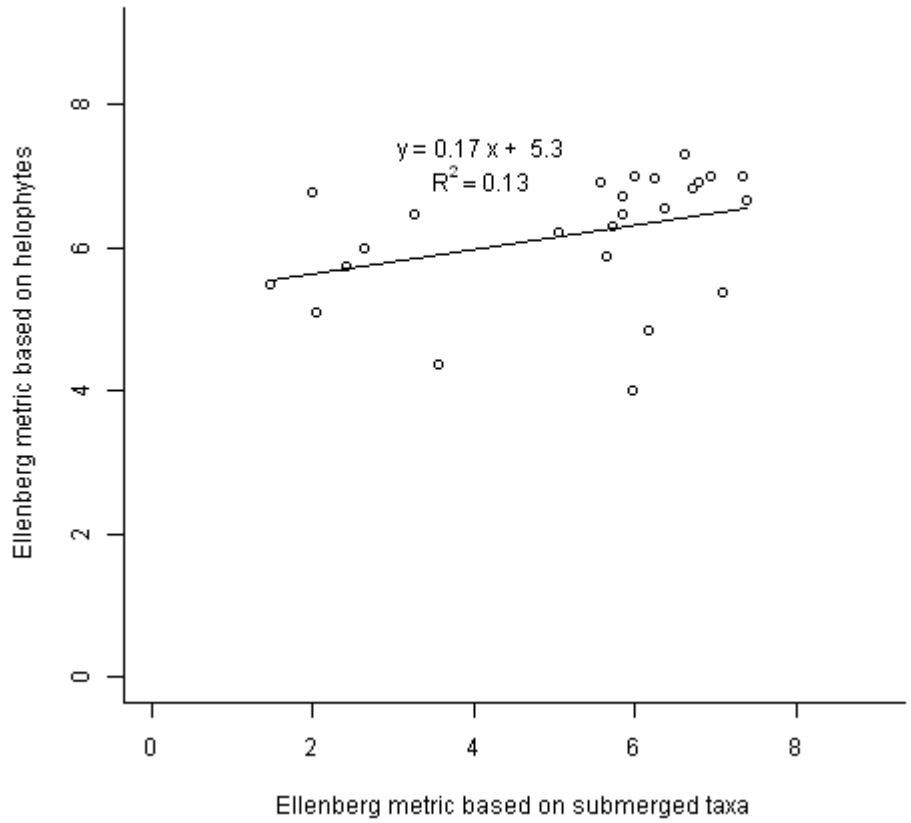

Fig. 4 Comparison of Ellenberg metrics calculated from only submerged taxa with the same metrics calculated from only helophytes, for each lake in this study. All metrics produced as a weighted average of taxa scores.