Collaborative Project (large-scale integrating project)
Grant Agreement 226273
Theme 6: Environment (including Climate Change)
Duration: March 1st, 2009 – February 29th, 2012

SEVENTH FRAMEWORK PROGRAMME

WISER

Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery

DELIVERABLE

Deliverable D3.1-6: Manuscripts on composition & bloom metrics, uncertainty and integration

Lead contractor: **Natural Environment Research Council (NERC)**

Contributors: **Carvalho L., Lyche Solheim A., Borics G., Catalan J., De Hoyos C., Drakare S., Dudley B., Jarvinen M., Laplace-Treyture C., Maileht K., McDonald C., Mischke U., Moe J., Morabito G., Nõges P., Nõges T., Ott I., Pasztaleniec, A., Phillips G., Poikane S., Skjelbred B. & Thackeray S.J**.

Due date of deliverable: **Month 36 (February 2012)**
Actual submission date: **Month 37 (March 2012)**

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013) Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

Non-technical summary

Phytoplankton constitute a diverse array of algae that live suspended in the water column of lakes and reservoirs. They are short-lived organisms (generation times of days to weeks) and they derive their nutrients exclusively from the water column. These features make this biological quality element the most direct and earliest indicator of the impacts of changing nutrient conditions on lake ecosystems. It also makes them particularly suitable for measuring the success of restoration measures following reductions in nutrient loads. This report summarises the manuscripts that have been planned to disseminate the results from the WISER Lake Phytoplankton Work Package (WP3.1). Much of the work for these manuscripts was developed in close collaboration with national phytoplankton experts from across Europe, who have been involved in a cross-comparison (Intercalibration) of classification schemes for the Water Framework Directive. These regional groups of experts are known as Geographical Intercalibration Groups (GIGs). A joint WISER/GIG Lake Phytoplankton Workshop was held in Italy in October 2011 to elaborate the main publications from this joint work. Full titles, proposed authorships and brief paper aims are described in this report.

Introduction

A number of manuscripts have been planned as part of the dissemination of WISER WP3.1 results. Much of the work for these manuscripts was developed in close collaboration with phytoplankton experts in the GIGs. A joint WISER/GIG Lake Phytoplankton Workshop was held at CNR-ISE, Pallanza, Italy, in October 2011 to elaborate the main publications from this joint work. Table 1 below gives an abbreviated list of the planned papers outlined at this workshop. Full titles, proposed authorships and brief paper aims are given in the following two sections. The first section (papers 1-15) is a list of papers using the external datasets provided to WISER via the GIGs and the EC REBECCA Project. The second section (papers 16-18) lists papers based on the use of the internal WISER field dataset.

Data owners (see acknowledgements) have all been contacted with this list of papers and have agreed to publication of the results. Several data owners are represented as co-authors in many of the publications; all are acknowledged in submitted publications using their data. Five of these papers are being submitted to the WISER Special Issue and two others are completed for submission elsewhere. Many others are in an advanced draft stage for submission over the coming months.

Table 1: Abbreviated list of the planned WISER/GIG papers from WP3.1

No.	1st Author	Short Paper Title
1	Geoff Phillips	PTI composition metric[1]
2	Giuseppe Morabito	Functional traits of lake phytoplankton
3	Ute Mischke	Evenness & richness response to P
4	Geoff Phillips	Cyanobacterial bloom metric
5	Laurence Carvalho	Integrated assessment and uncertainty[1,2]
6	Anne LycheSolheim	Module paper on BQE responses to pressure[1,2]
7	Gabor Borics	Compositional responses to limitation
8	Tatiana Caraballo	The occurrence of colonial algae
9	Laurence Carvalho	Cyanobacteria responses to phosphorus
10	Laurence Carvalho	Nutrient limitation and cyanobacteria response
11	Marko Järvinen	Phytoplankton in reference sites[1]
12	Kairi Maileht	Richness responses to lat/long, altitude[1]
13	Geoff Phillips	How do phytoplankton metrics inform P targets?
14	Stephen Thackeray	Temporal variability in phytoplankton metrics
15	Jordi Catalan	The algal group index (IGA)
16	Jordi Catalan	HPLC pigment analysis for monitoring[2]
17	Stephen Thackeray	Spatial variability in phytoplankton metrics[2]
18	Stephen Thackeray	Outflow vs Open water sampling (UK/NO)[2]

[1] Planned for submission to the WISER Special Issue

[2] Uses the WISER field exercise data

Manuscripts being planned using external data (e.g. GIG datasets)

1. A phytoplankton trophic index to assess the status of lakes for the Water Framework Directive.

Geoff Phillips, Anne Lyche Solheim, Birger Skjelbred (+ co-authors contributing to writing)

We present here a new pan-European phytoplankton taxonomic index (PTI) that was developed for use as part of an independent common metric for the WFD Intercalibration process. The metric was developed from a dataset containing data from 21 European countries and over 1500 lakes. We selected a training set of data from the summer period (July – September) and used Canonical Correspondence Analysis with a single constraining environmental variable, total phosphorus, to produce a set of taxa optima from the 1st ordination axis. These optima were then used to generate sample scores using a weighted average of the proportion of the biovolume of each taxa present in the sample. The resulting index was shown to have a good relationship with pressure measured as total phosphorus (GAM model R2 = 0.667 $p<0.001$), but was different for lakes of low, moderate and high alkalinity. To allow for this, the index was converted to an Environmental Quality Ratio (EQR) by dividing by type specific reference values, derived from a model which used a population of reference lakes to predict reference PTI values. The metric was subsequently successfully used as the taxonomic component of a common metric to compare the status of European phytoplankton assessment systems through the Intercalibration process.

2. Functional traits of lake phytoplankton across Europe and their response to environmental gradients.

Giuseppe Morabito, Catalan, Mischke (+co-authors)

Traits-based community ecology in phytoplankton studies recently received new inputs and an increasing interest. These approaches hold the potential of increasing our ability to explain the organization of ecological communities and predict their reorganizations under different environmental constraints. The paper will focus on the use of some functional traits (size, motility, mixotrophy, colony formation, silica requirements) for describing phytoplankton responses to environmental gradients. The distribution and dominance of the different traits will be investigated across Europe, in a set of lakes with different geographic, morphological and chemical features. Some of the key questions to be addressed can be: how do the functional traits respond? Which traits are selected under certain environmental conditions? Is the response the result of the environmental heterogeneity? Is there any relationship with geographic distribution of the lakes? Do lakes/lake types with higher environmental patchiness offer more ecological niches and host an higher number of functional traits? Which traits give the better response to eutrophication pressure (considering the needs of the WFD)?

3. Lake Phytoplankton: Can Diversity serve as a Phytoplankton Index?

Ute Mischke, Laurence Carvalho, Birger Skjelbred, Caridad de Hoyos, Jose Pahissa, Christoph Laplace-Treyture, Anne Lyche-Solheim, Geoff Phillips, Kairi Maileht, Ingmar Ott, Jordi Catalan, Gabor Borics

The paper will introduce the distribution of diversity indices (species richness, evenness) in the eco-regions Northern Europe, Central Europe, and Mediterranean region along the main pressure eutrophication, e.g. total phosphorous (TP). The paper will summarize results from WISER on metric evenness and how it interacts with a critical bloom density. Lakes are split in types by mean water depth, alkalinity and turbidity. In summer plankton of Scandinavian lakes the diversity of phytoplankton is slightly higher and it is stronger influenced by TP than in CB-GIG. Problems associated with different skills in taxa determination and species number per sample will be discussed. Also selected cases will be described which lower the correlation to TP: Which taxa are dominant when evenness is low in some of the reference lakes? Which communities cause high evenness even under bloom conditions in some other lakes?

4. Assessing lake quality based on cyanobacteria abundance.

Geoff Phillips, Laurence Carvalho, Caridad de Hoyos, Anne Lyche Solheim, Ute Mischke, Gabor Borics, Stina Drakare, Marko Jarvinen, Sandra Poikane & Claire McDonald.

Cyanobacterial blooms are probably the most widely recognized ecological responses to eutrophication and one of few WFD elements to have explicit consequence for ecosystem services (access to safe, clean water for drinking and recreation). Application of a bloom metric in ecological assessment, therefore, is of great relevance to the general public and policy makers. We review cyanobacteria metrics proposed for the WFD and describe approaches adopted in IC for establishing reference conditions and setting status class boundaries. We recommend bloom metrics based on actual abundance, rather than % abundance, as these are more relevant to health risks and are shown to have a more robust relationship with nutrient pressures. We outline an approach to boundary-setting in relation to World Health Organisation (WHO) thresholds for recreational waters.

5. Lake Phytoplankton: integrated assessment and uncertainty

Carvalho L., Borics G., Catalan J., De Hoyos C., Drakare S., Dudley B., Jarvinen M., Laplace-Treyture C., Lyche Solheim A., Maileht K., Mischke U., Moe J., Morabito G., Nõges P., Nõges T., Ott I., Pasztaleniec, A., Phillips G., Poikane S., Skjelbred B. & Thackeray S.

The paper will introduce the philosophy of phytoplankton assessment i.e. that it should represent the wider impacts of eutrophication and not just TP. It will highlight the need for new metrics (e.g. bloom metrics) and the need for a common metric (combination of chlorophyll, PTI and ideally a bloom metric). The paper will summarise results from WISER on metric strength and uncertainty and recommend the strongest combination of metrics for IC purposes. The reasons why a range of metrics are needed will be discussed and why a perfect correlation with TP or with each other is not necessary. To illustrate this added value, case-studies of where particular metrics, and their combination, affect status assessment will be highlighted. Problems associated with different combination rules will also be discussed

6. Comparing responses of different Biological Quality Elements to different pressures in lakes, Synthesis of WISER results on Lakes

Anne Lyche Solheim, Laurence Carvalho, Agnieszka Kolada, Martin Pusch, Torben Lauridsen, Richard Johnson?, Sandra Poikane?, (+ co-authors contributing to writing)

Different biological quality elements in lakes have different sensitivities to different pressures. While phytoplankton responds primarily to eutrophication, the other biological quality elements may respond also to hydromorphological pressures. They also respond differently to the same pressure in terms of strong or weak relationships, as well as whether or not clear thresholds occur in the responses curves along the gradients. By compiling and comparing response curves for various metrics for each BQE to different pressures we will identify the most sensitive BQEs/metrics and provide some recommendations to managers as to which BQEs/metrics are preferable when assessing impacts of different pressures on WFD ecological status of lake ecosystems.

7. Phytoplankton compositional responses to nitrogen and phosphorus limitation

Gabor Borics, Laurence Carvalho, Ute Mischke, Bernard Dudley (+co-authors)

Analysis of WISER dataset to examine whether particular species or genera are associated with N- or P-limitation, co-limitation or nutrient excess. Particular focus on cyanobacteria species?

8. Trophic gradients in lakes and the occurrence of colonial forms across phytoplankton groups

Tatiana Caraballo, Jordi Catalan, Caridad de Hoyos, Birger Skjelbred, (+coauthor contributors)

In this study we aim to evaluate the hypothesis that colonial forms require higher nutrient concentrations than evolutionary-closed single cell forms. Based on the large WISER phytoplankton data set, we will look at the TP optimal of the species (and probably also other nutrient state data also) and whether there are significant difference according to evolutionary lineages between single cell species and colonial species of different types.

9. Cyanobacterial responses to phosphorus concentrations and their application to recreational health thresholds

Laurence Carvalho, Claire McDonald, Caridad de Hoyos, Ute Mischke, Geoff Phillips, Gábor Borics, Sandra Poikane, Birger Skjelbred, Anne Lyche Solheim & Ana Cristina Cardoso

A safe, clean water supply is critical for sustaining many important ecosystem services provided by freshwaters. The development of cyanobacterial blooms in lakes and reservoirs has a major impact on the provision of these services, particularly limiting their use for recreation and water supply for drinking and spray irrigation. Nutrient enrichment and climate change are thought to be the most important pressures responsible for the widespread increase in cyanobacterial blooms in recent decades. Quantifying how nutrients limit cyanobacterial abundance is, therefore, a key need for setting robust targets for the management of freshwaters. Using a dataset from over 1500 European lakes, we highlight the use of quantile regression modelling for understanding the maximum potential capacity of cyanobacteria in relation to phosphorus and the use of a range of quantile responses, alongside World Health Organisation (WHO) health alert thresholds for recreational waters, for setting robust phosphorus targets for lake management in relation to water use. The analysis shows that cyanobacteria exhibit a non-linear response to phosphorus with the sharpest increase in cyanobacterial abundance occurring in the TP range from about 20 µg L^{-1} up to about 100 µg L^{-1}. The likelihood of exceeding the WHO 'low health alert' threshold increases from about 5% exceedance at 16 µg L^{-1} to 40% exceedance at 54 µg L^{-1}. About 50% of lakes remain below the WHO low threshold, irrespective of increasing TP concentrations, highlighting the importance of other limiting factors affecting population growth and loss processes, such as high flushing rate. Developing a more quantitative understanding of the limiting effect of nutrients on cyanobacterial abundance in freshwaters provides important knowledge for restoring and sustaining a safe, clean water supply and can also support mitigation strategies in relation to the less manageable pressure of climate change. Our results can be used to set nutrient targets to sustain recreational services and provide different levels of precaution that can be chosen dependent on the importance of the service provision.

10. Assessing the extent of nitrogen-limitation in European lakes and the consequences for the development of cyanobacteria

Carvalho, L., Cardoso, A.C., Borics, G., De Hoyos, C., Drakare, S., Dudley, B., Järvinen, M., Laplace, C., Maberly, S.C., Mischke, U., Morabito, G. Phillips, G. & Solheim, A.-L.

This study of over 700 European lakes reveals that nitrogen-limitation of phytoplankton production is much more widespread than previously acknowledged, particularly from June to October. In many European lakes, concentrations of both phosphorus (P) and nitrogen (N) decline to low levels during the summer months, suggesting co-limitation is widespread. This is especially true in nutrient-poor lakes in Sweden & Finland. N-limitation alone is more prominent in lakes with high P concentrations, such as in Belgium and the Netherlands. Very few lakes have an excess of both N and P during the summer months. The widespread extent of N-limitation, particularly when phosphorus is in excess, may have important consequences for the development of toxic cyanobacterial blooms and recreational impacts. This study examines the widely accepted hypothesis that cyanobacteria are more abundant in N-limited lakes.

11. Phytoplankton composition and biomass in Northern European clear water and humic reference lakes

Marko Järvinen, Stina Drakare, Anne Lyche-Solheim, Ute Mischke, Geoff Phillips, Birger Skjelbred, Gary Free

Describe the phytoplankton species composition, indicator species, and total biomass in Northern European reference lakes, with special emphasis on N-GIG clear water and humic lake types.

12. Water colour, phosphorus and alkalinity are the major determinants of the dominant phytoplankton species in European lakes

Kairi Maileht, Tiina Nõges, Peeter Nõges, Ingmar Ott, Ute Mischke, Laurence Carvalho, & Bernard Dudley

Analysis of phytoplankton data from >1500 lakes in 20 European countries has revealed that two-thirds of the 151 most common summer dominant species are dominant in both Northern Europe and Central & Southern Europe. As these species were common, we were able to study how both habitat conditions in lakes and environmental factors over broad geographical scales explained their distribution. We hypothesized that temperature, trophic state, water colour, and alkalinity would be the major determinants of the dominance of different species in European lakes. Using Canonical Correspondence Analyses, as expected, which species dominated was best explained by water colour and latitude along the primary axis. Alkalinity and total phosphorus also appeared to be important explanatory factors determining which species dominated, independent from colour and latitude. Contrary to our hypothesis, water temperature from July to September had only a negligible impact on the distribution of dominants, showing the prevalence of rather homogeneous thermal conditions throughout Europe for this period of year. A comparison of the data from the northern ecoregion with the rest of Europe showed that chrysophytes and cryptophytes occurred more frequently among dominants in the north whereas cyanobacteria and dinophytes in the southern parts of Europe. Our analysis suggest that besides trophic conditions, other hydrochemical variables, such as alkalinity and the content of humic substances, have at least as important role in determining the distribution of the dominant phytoplankton species in European lakes.

13. How do phytoplankton metrics inform P targets?

Geoff Phillips, Anne Lyche Solheim (+ co-authors)

Early stages of planning – like paper 9 but using all metrics?

14. Temporal variability in lake phytoplankton community metrics

Stephen J. Thackeray, Peeter Nõges, Michael Dunbar, Bernard J. Dudley, Ana Negro, Jannicke Moe ((+ co-authors contributing to writing)

The values of phytoplankton metrics will vary temporally (within and among years), as a result of similar variations in the communities that they represent. Water body assessments based upon the calculation of these metrics from field samples will therefore have an inherent level of uncertainty, since different results would be obtained on different sampling dates. Here we will analyse data from a large number of European lakes in order to quantify the level of temporal variation in a range of candidate phytoplankton metrics, and compare this to the level of variation in metrics among lakes. We aim to establish whether temporal metric variation is of a similar magnitude to variation among lakes spanning a wide eutrophication pressure gradient. The results will help to reveal how uncertainties in water body assessment may be moderated by adopting sampling protocols with different temporal resolutions.

Manuscripts being planned using Internal WISER field data

15. The algal group index (IGA) for assessing ecological status of lakes and reservoirs within the Water Framework Directive.

Jordi Catalan, Caridad de Hoyos, José Pahissa, Tatiana Caraballo, (+ co-authors contributing to field dataset and manuscript)

The algal group index (IGA) is a phytoplankton index currently in use in several Mediterranean countries that has not been formally described in a scientific paper. Our goal is to provide the definition, the scientific rationale behind the index, the sources of uncertainty in its application, the potential shortcomings and possible improvements including the evaluation of lower cost alternative.

16. HPLC pigment analysis as a phytoplankton monitoring system: benefits and uncertainties

Jordi Catalan, Ute Mischke, Suzanne McGowan, Rene Freiberg, Giuseppe Morabito, Andrea Lami, Nicolas Mazella (+ other coauthors contributing to field dataset?)

Phytoplankton is a key element in the assessment of the ecological quality of freshwaters (e.g., EU Water Framework Directive). Both abundance and composition of the phytoplankton populations are relevant in determining the ecological status. Whereas chlorophyll a concentrations is a quick method to characterize phytoplankton amounts, composition features are commonly considered through indexes or metrics that involves time consuming counting methods. HPLC pigment analysis could provide an alternative approach in which a decline in taxonomic resolution may be largely compensated by the speed of the process and the standardization of the methods across labs. In this paper, we would examine the reliability of the method, different alternative approaches and uncertainties linked to them.

The study is based on a field campaign across European lakes and reservoirs covering a wide range of trophic states and environmental conditions. HPLC pigments analysis and phytoplankton counting were simultaneously performed in each lake.

17. Quantifying uncertainties in biologically-based water quality assessment: a pan-European analysis of phytoplankton community metrics

Stephen J. Thackeray, Peeter Nõges, Michael Dunbar, Bernard J. Dudley, Birger Skjelbred, Giuseppe Morabito, Laurence Carvalho, Geoff Phillips, Ute Mischke, Jordi Catalan, Caridad de Hoyos, Christophe Laplace, Martina Austoni, Tomasa Virdis, Kairi Maileht, Agnieszka Pasztaleniec, Marko Jarvinen, Stina Drakare & Anne Lyche Solheim.

The EU Water Framework Directive (WFD) states that attributes of biological communities should be used to assess the ecological status of fresh- and coastal/transitional waters. For lakes, the phytoplankton is a key biological community to be used for this purpose. It is therefore necessary to develop metrics that describe high-level properties of phytoplankton communities and that are sensitive to environmental pressures, such as nutrient enrichment. Assessment of the utility of such metrics demands a knowledge of the extent to which they can be affected by sampling and sample processing procedures e.g. where samples are collected from and who processes the samples. If metrics vary more with differences in sampling and sample processing within a lake than they do among lakes then they are unlikely to provide a sensitive means of describing differences in the biological impacts of an environmental stressor among lakes. Here we analyse the results of a multi-scale field campaign of 32 European lakes, to resolve the extent to which seven proposed phytoplankton metrics vary among lakes and with sampling/sample processing. We also relate these metrics to different environmental variables, including total phosphorus concentration as an indicator of eutrophication.

For all seven metrics, between 65% and 96% of the variance in metric scores was due to variability among lakes, much higher than variability occurring due to sampling/sample processing. Using multi-model inference, there was strong support for relationships between among-lake variation in three of the metrics and differences in total phosphorus concentrations. Three of the metrics were similarly related to mean lake depth. Unexplained among-lake metric variance indicated that metrics were additionally sensitive to unmeasured environmental factors. Differences among sub-samples and analysts accounted for much of the within-lake metric variance, suggesting that sub-sample replication and standardisation of analyst procedures may result in increased precision of ecological assessments based upon these metrics. The residual variance in most metrics, and therefore the uncertainty associated with them, changed as a function of among lake variations in the physical (mean depth) and chemical (total phosphorus concentration) environment, and lake location (altitude).

For three of the candidate phytoplankton metrics being considered for the WFD Intercalibration of lake phytoplankton metrics: chlorophyll a concentration, the Phytoplankton Trophic Index (PTI), and cyanobacterial biovolume, > 88% of the variance in metric scores was among-lakes, and, total phosphorus concentration was well supported as a predictor of this among-lake variation. Based upon this study, these proposed metrics may be considered robust for ecological status assessment and suitable for adoption in the WFD Intercalibration process.

18. A comparison of variability in metric scores between outflow and open-water sampling locations

Stephen J. Thackeray, Geoff Phillips, Sian Davies, Laurence Carvalho, Anne Lych Solheim

Analysis of UK (WISER & EA) and Norwegian (WISER) field data

Acknowledgements

The papers listed are all a result of the project WISER (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (contract No. 226273). We would like to thank Birger Skjelbred, Jannicke Moe and Bernard Dudley who supported data management and extraction and especially thank all the data providers. These include (figures in brackets give the number of WBs):

Mediterranean GIG (Data manager: Caridad de Hoyos, CEDEX-CEH)

Portugal: Instituto da Água, I.P. (INAG) (18)

Spain: Ministerio de Agricultura, Alimentación y Medio Ambiente (122), Centro de Estudios Hidrográficos (CEDEX-CEH) (46).

France: Institut national de recherche en sciences et rechnologies pour l´environnement et l´agriculture (IRSTEA) (6).

Italy: Università degli Studi di Sassari. Dipartimento di Scienze Botaniche, Ecologiche e Geologiche (DiSBEG) (18).

Cyprus: Ministry of Agriculture, Natural Resources and Environment, Water Development Department (MANRE-WDD) (7).

Greece: Maria Moustaka, Aristotle Unniversity of Thessaloniki (1)

Central-Baltic GIG (Data manager: Ute Mischke, IGB)

Estonia: Estonian University of Life Sciences (EMU) gathered for the state monitoring programme supported by the Estonian Ministry of Environment.

Latvia: Sandra Poikane, Latvian Environment, Geology and Meteorology Centre

Lithuania - EPA Lithuania

Denmark: National Environmental Research Institute, University of Aarhus.

Netherlands: Rijkswaterstaat (RWS)

Belgium: Jeroen Van Wichelen, Ghent University (UGENT)

France: Institut national de recherche en sciences et rechnologies pour l´environnement et l´agriculture (IRSTEA).

Germany: Data from German water bodies were kindly provided by the following institutions of the German Federal States: Landesamt für Umwelt, Gesundheit und Verbraucherschutz Brandenburg (LUGV; 127), Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg-Vorpommern (MLUV, Seenprogramm, 65), Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt (LHW, 5), Landesamt für

Landwirtschaft, Umwelt und ländliche Räume Schleswig-Holstein (LLUR, 13), Senatsverwaltung für Gesundheit, Soziales und Verbraucherschutz Berlin (SenGUV, 12), Niedersächsische Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN, Sulingen, 1).

Poland: Data were provided partly by the Institute of Environmental Protection - National Research Institute. The Inspection for Environmental Protection provided data obtained within the framework of state environmental monitoring

United Kingdom - Scottish Environment Protection Agency (SEPA) and the Environment Agency for England & Wales (EA)

Eastern-Continental GIG (Data manager: Gabor Borics, EPI)

Hungary: Environmental Protection Inspectorate for Trans-Tiszanian Region

Romania: Ministeriul Meduli şi Pădurilor (MMP) (10).

Northern GIG (Data manager: Geoff Phillips, EA)

Finland: Finnish Environment Institute (SYKE)

Sweden: Swedish University of Agricultural Sciences (SLU)

Norway: Norwegian Institute for Water Research (NIVA)

United Kingdom - Scottish Environment Protection Agency (SEPA) and the Environment Agency for England & Wales (EA)

Ireland: Environment Protection Agency (EPA)

Annex 1: Completed papers

Carvalho et al. (2012). Cyanobacterial responses to phosphorus concentrations and their application to recreational health thresholds

Laurence Carvalho[1,2*], Claire McDonald[2], Caridad de Hoyos[3], Sandra Poikane[1], Ute Mischke[4], Geoff Phillips[5], Gábor Borics[6], Birger Skjelbred[7], Anne Lyche Solheim[7] & Ana Cristina Cardoso[1]

[1]Institute for Environment & Sustainability, European Commission Joint Research Centre (JRC), Ispra, ITALY.

[2]Centre for Ecology & Hydrology (CEH), Bush Estate, Penicuik, EH26 0QB, UK

[3]CEDEX, C/ Paseo Bajo de la Virgen del Puerto nº 3; 28005 Madrid, Spain

[4]IGB, D-12489 Berlin, Justus-von-Liebig-Str. 7, Germany

[5]Environment Agency, Kings Meadow Road, Reading, Berkshire, RG1 8DQ, UK

[6] Balaton Limnological Research Institute, Department of Tisza River Research, Klebelsberg Kuno U. 3, Tihany, H-8237, Hungary

[7] NIVA, Gaustadalléen 21, N-0349 Oslo, Norway

*On secondment from CEH[2] to JRC[1]. Email address for correspondence: laca@ceh.ac.uk

Running title: Cyanobacterial responses to phosphorus

Word count: entire paper (5558), summary (330), main text (3794), acknowledgements (413), references (845), tables and figure legends ().

Number of tables and figures: 11

Number of references: 39

Summary

A safe, clean water supply is critical for sustaining many important ecosystem services provided by freshwaters. The development of cyanobacterial blooms in lakes and reservoirs has a major impact on the provision of these services, particularly limiting their use for recreation and water supply for drinking and spray irrigation. Nutrient enrichment and climate change are thought to be the most important pressures responsible for the widespread increase in cyanobacterial blooms in recent decades. Quantifying how nutrients limit cyanobacterial abundance is, therefore, a key need for setting robust targets for the management of freshwaters.

Using a dataset from over 1500 European lakes, we highlight the use of quantile regression modelling for understanding the maximum potential capacity of cyanobacteria in relation to phosphorus and the use of a range of quantile responses, alongside World Health Organisation (WHO) health alert thresholds for recreational waters, for setting robust phosphorus targets for lake management in relation to water use.

The analysis shows that cyanobacteria exhibit a non-linear response to phosphorus with the sharpest increase in cyanobacterial abundance occurring in the TP range from about 20 µg L^{-1} up to about 100 µg L^{-1}.

The likelihood of exceeding the WHO ‘low health alert’ threshold increases from about 5% exceedance at 16 µg L^{-1} to 40% exceedance at 54 µg L^{-1}. About 50% of lakes remain below the WHO low threshold, irrespective of increasing TP concentrations, highlighting the importance of other limiting factors affecting population growth and loss processes, such as high flushing rate.

Synthesis and applications Developing a more quantitative understanding of the limiting effect of nutrients on cyanobacterial abundance in freshwaters provides important knowledge for restoring and sustaining a safe, clean water supply and can also support mitigation strategies in relation to the less manageable pressure of climate change. Our results can be used to set nutrient targets to sustain recreational services and provide different levels of precaution that can be chosen dependent on the importance of the service provision.

Keywords: algal bloom, blue-green algae, ecosystem services, freshwater, lake, nutrient, quantile regression, WHO

Introduction

Currently there is much political drive to quantify ecosystem services provided by freshwaters (Millennium Ecosystem Assessment, 2005; Carpenter et al., 2009) although there is still a great deal of debate as to what primary data can best be used to map the provision and quality of services (Eigenbrod et al., 2010). In this respect, for freshwaters, a safe, clean water supply is a critical need. Cyanobacteria, specifically the toxins they produce, represent one of the most hazardous waterborne biological substances that produce a range of adverse health effects from mild skin irritations to severe stomach upsets and even fatal consequences (Codd et al., 1999; 2005). The widespread development of large cyanobacterial populations, or blooms, in lakes and reservoirs, therefore, has a major impact on the provision of many ecosystem services, particularly limiting their use for recreational activities in and around freshwaters (WHO, 2003) and water supply for drinking and spray irrigation (WHO, 2004). Their abundance can, therefore, be used as one indicator of the 'functional quality' of freshwater services

There is strong evidence that the development of cyanobacterial blooms has been increasing in recent decades (Smith 2003) and this is widely believed to be primarily due to nutrient enrichment, especially phosphorus (Downing et al., 2001; Schindler, 2008), but also in response to warmer and drier summer conditions (Paerl & Huisman 2009, Weyhenmeyer et al. 2002) and more stable stratification (Wagner & Adrian, 2009). Nutrient concentrations in the water set the capacity for cyanobacteria standing crops and are probably the most manageable pressure affecting their abundance. Developing a more quantitative understanding of the limiting effect of nutrients on cyanobacterial capacity within freshwaters would, therefore, provide important knowledge for restoring and sustaining a safe, clean water supply and could also support mitigation strategies in relation to the less manageable pressure of climate change.

There is a vast amount of quantitative empirical evidence demonstrating increasing phytoplankton abundance under increasing nutrient concentrations (Dillon & Rigler, 1974; OECD, 1982; Phillips et al., 2008). There are also a few studies examining more specifically, the relative (%) abundance of cyanobacteria in relation to nutrients (Downing et al., 2001; Ptacnik et al., 2008; Wagner & Adrian, 2009). There are, however, far fewer extensive empirical studies quantifying the actual abundance of cyanobacteria in relation to nutrients, despite this being more directly relevant to water use. Most studies of actual abundance are of individual or small groups of lakes (Mischke, 2003; Nõges et al., 2008), although Carvalho et al (2011) examined cyanobacteria abundance in about 150 lakes in the UK in relation to lake types, flushing rates and nutrients. In this paper, we analyse the actual abundance of cyanobacteria from a dataset of over 1500 European lakes, substantially larger than any other analysis reported in the literature. With this dataset we are able to provide robust quantification of the abundance of cyanobacteria in relation to phosphorus concentrations. In particular, we highlight the use of quantile regression modelling for understanding the maximum potential capacity, or upper quantile response, of cyanobacteria in relation to phosphorus and the use of a range of quantile responses, alongside the World Health Organisation (WHO) cyanobacteria thresholds for recreational waters (WHO, 2003), for setting robust phosphorus targets for lake management in relation to water use.

Materials and Methods

Data

The EC Water Framework Directive has enabled the collation of large biological datasets following standard sampling and counting methodologies (Moe et al., 2008). As part of the EC WISER Project (http://www.wiser.eu/), phytoplankton composition data were collated from 1506 lakes spanning 4 European biogeographical regions and 16 countries (Table 1). The bulk of the data were from low and medium alkalinity lakes in Northern Europe (855 lakes) and high alkalinity lakes in Central-European or Baltic countries (599 lakes) (Table 1). Both cyanobacterial abundance (biovolume) and nutrient data were summarised as a summer mean using data spanning the months July, August and September. For each lake, only the last year of available data was used in the analysis to avoid bias of lakes with many years of data. Phytoplankton samples were predominantly integrated tube samples from the middle of the lake, counted after preservation with Lugol's iodine solution. In general, 400 counting units were measured across magnifications usually using a combination of low magnification full-chamber counts, intermediate magnification transects and high magnification fields of view. Counts and biovolume estimates of cells, colonies and filaments broadly followed the approach outlined by CEN (2004).

Table 1. No. of lakes with cyanobacteria and TP data, by region, country and alkalinity type.

Statistical Analysis

The majority of biological response modelling approaches in current use [e.g. simple linear least squares regression, generalized linear models (GLM) or generalized additive modelling (GAM)] are based on the estimation of mean or median responses to environmental factors. One method which models the relationship of variables at different levels of a distribution is quantile regression (Koenker & Bassett 1978). Modelling high quantiles, such as the 95%, may better represent the maximum capacity of cyanobacteria for a given TP concentration. Other quantiles can also be estimated and can be used to identify relationships that least squares regression of mean responses may not effectively represent. Quantile regression was used to model responses of cyanobacterial abundance (actual biovolume) against TP concentrations. In this analysis, a number of percentile cyanobacterial responses were modelled, 10%, 25%, the median (50%), 75%, 90% and 95%. Linear and non-linear parametric, and non-parametric quantile regression were all applied to the data in R (R Development Core Team, 2010), using routines available in the *quantreg* pakcage (Koenker, 2009). Non-parametric quantile regression was applied using the function *rqss* in the *quantreg* package which fits a smoothing spline using a roughness penalty term. The parametric non-linear quantile regression models are described further below. Simple linear regression and GAMs of the mean response (Wood 2006, 2008) were also examined for comparison with the quantile models. GAM's are non-parametric modifications of generalized linear models with a linear predictor involving a sum of smooth functions of the explanatory variables (Hastie & Tibshirani, 1986). Cubic regression splines were used as the type of smooth function. The GAM was fitted using a normal distribution and an identity link function.

Non-parametric regression models are based on rank differences and, therefore, cannot, be used to describe or visualise the relationship and do not enable predictions from model equations. Therefore, parametric, non-linear quantile regression was applied to the datasets to enable this using the interior point algorithm for finding the best fitting model solution (Koenker & Park 1994) For the significant quantiles, the following 3-parameter asymptotic exponential equation was generally used:

$\text{Log}_{10}(\text{Cyanobacteria volume} + 1) = a/(1+b*\exp(-c*\text{Log}_{10}(\text{total phosphorus})))$

Where,a = cyanobacteria biovolume where the fitted curve begins to reach a maximum
b = a - position on the y-axis where the convex curve starts
c = position of x-axis where the initial change in slope occurs i.e where the concave curve starts.

For the 25% quantile model, a 2-parameter asymptotic exponential model was fitted to the data:

$\log(\text{cyano biovolume}+1) \sim (a * \exp(b * \log(TP)))$

Where a is the intercept of the line and b is the slope of the line.

Akaike's information criterion (AIC) values for the linear and the non-parametric quantile regression models were used to compare the different quantile model fits to the data to distinguish the best models for prediction purposes; the model having the lowest AIC being the best. For parametric non-linear quantile regression, AIC values cannot be calculated for each quantile, therefore, deviance is reported as a measure of goodness of fit (Crawley, 2009). For a continuous variable, such as cyanobacterial biovolume, deviance is calculated as:

$$D = \sum_{i=1}^{n} (y_i - \mu)^2$$

where n is the sample size, y_i is the observed data point and μ is the mean of the y variable. The lower the deviance value then the better the fit of the model to the dataset.

WHO Guidelines

In this study, the quantile modelling approach is combined with WHO thresholds for cyanobacterial abundance in recreational waters to identify the likelihood of exceeding health alert thresholds. WHO (1999; 2003) recommend “a series of guideline values associated with incremental severity and probability of health effects” and these values are then defined for three health alert categories: low, moderate and high. A high alert (or high probability of adverse health effects) is assigned when surface scums are present, where cell densities and toxin concentrations can be very high and severe health risks are possible. Cyanobacteria cell densities of 20,000 and 100,000 cells ml^{-1}, respectively, are associated with "relatively low" and "moderate" probabilities of adverse health effects, associated with less severe symptoms such as skin irritations and gastro-intestinal illness. These cell densities can be converted to a biovolume (mm^3 L^{-1}) by multiplying by a typical cyanobacterial cell volume. We have adopted here the equivalent biovolumes of 2 mm^3 L^{-1} and 10 mm^3 L^{-1}, outlined in WHO (1999), based on a spherical cell with a diameter of 5.7 µm.

Results

Exploratory analysis of the data highlighted that cyanobacteria are generally absent, or in very low abundance, in low alkalinity lakes (< 200 µequiv. L^{-1}) (Fig. 1). The mean and median abundance and upper percentiles all clearly increase with increasing alkalinity class (Fig. 1; Table 2). Because of their general absence or low abundance in low alkalinity lakes, the quantile regression analysis was carried out on a sub-set of 807 medium and high alkalinity lakes drawn from all regions.

Mean response

Considering the whole lake dataset together, there is a positive linear relationship between ($\log_{10}$) cyanobacterial biovolume and ($\log_{10}$) TP concentrations (r^2 = 0.295, p<0.001, deviance 138.7). Despite the significance of the relationship there is clearly still a large amount of scatter in the data. A GAM (Fig. 2; r^2_{adj} = 0.342, p = <0.001, deviance = 128.9) and a 3-parameter non-linear model (Fig. 3; deviance = 129.4) of the mean cyanobacterial response fit the data better. Both non-linear models indicate a take-off in the mean cyanobacterial response above a TP concentration of approximately 10 µg L^{-1}. For the GAM model a strong positive response is apparent up to about 300 µg L^{-1} (Fig. 2), whereas for the parametric non-linear model, there is a flattening of the mean response at a threshold of about 100 µg L^{-1} (Fig. 3). Below about 5 µg L^{-1} and above about 300 µg L^{-1}, there are few data points and, therefore, less confidence in the modelled relationships outside this TP range (Fig. 3).

Quantile responses – medium and high alkalinity lakes

A linear quantile model was initially fitted to the medium and high alkalinity lake data, however, comparison of AIC values for linear and non-linear non-parametric quantile regression models highlight the much poorer fit of linear models for most quantiles (Table 3). Models for the two lowest quantiles examined (0.05 and 0.10) were linear and had the lowest AIC values, due to the large proportion of low or zero values for cyanobacteria biovolume. These lower quantile relationships between cyanobacteria biovolume and TP were, however, more or less flat, and there was no significant relationship between the two variables e.g. 0.05 (p=0.98), 0.10 (p=1.00) quantiles. For all higher quantiles examined (0.25 and above), non-linear, non-parametric regression models were a better fit and all had a highly significant relationship between cyanobacteria biovolume and TP (p<0.001).

Three-parameter (asymptotic exponential models were the best fit for the 0.50 – 0.95 quantile models and the non-linear mean response, whereas only a 2-paramter model was selected for the 0.25 quantile (Table 4). The resulting non-linear parametric regression models for significant quantiles 0.25-0.95 are shown in Figure 3 and deviance values and parameter estimates are given in Table 4. The models shown are thosewith the deviance minimized.

Like the non-linear model for the mean response, all the quantile models indicate a take-off in the cyanobacterial response above a threshold TP of approximately 10 µg L^{-1} and a flattening of the response at a threshold of about 100 µg L^{-1} (Fig. 3). The biggest difference between the different quantiles is in the slope of the increase, with the 0.95 quantile showing the steepest increase, whilst the 0.50 quantile the shallowest increase. Additionally the quantiles differ greatly in terms of parameter a: the estimated cyanobacteria value where the fitted curve begins to reach a maximum (Table 4, Figure 3). For example, the 0.50 quantile plateaus at just below 2 mm^3 L^{-1}, the WHO (1999) low risk threshold, at TP concentrations of 100 µg L^{-1} or greater.

Application of quantile responses for predicting bloom capacity

The upper quantiles (e.g. 0.95) provide estimates of the potential maximum capacity of cyanobacteria in response to increasing TP concentrations (Table 5). The capacity for cyanobacteria increases with increasing TP, with the relationship levelling off at TP concentrations >150 µg L^{-1}. The 95% quantile model indicates that at 16 µg L^{-1}, 5% of lakes will exceed the low risk threshold and at 32 µg L^{-1} 5% of lakes will exceed the medium risk threshold (Table 5).

Nutrient targets in relation to health thresholds

The equations in Table 3.5 can be used to determine the proportion of lakes exceeding the low and medium risk WHO thresholds for cyanobacteria for a given TP concentration (Table 6; Figure 4). Only significant quantile curves which pass through these risk threshold levels can be used. The results indicate that at a TP concentration of about 22 µg L^{-1} 10% of lakes exceeded the WHO low risk threshold, at 31 µg L^{-1} this increased to 25% of lakes, and at 41 µg L^{-1} 33% of lakes were above the WHO low risk threshold. Similarly 10% of lakes exceeded the WHO medium risk threshold, at TP concentrations of 48 µg L^{-1}.

Discussion

Despite the wide variety of life strategies between different cyanobacterial species and the consequent variety of environmental factors shaping their abundance (Dokulil & Teubner 2000; Reynolds et al., 2002), it is still of great importance to understand more fully the response of this whole group of algae in relation to nutrient pressures. The reason for this is that many cyanobacterial species produce hazardous toxins and this has led to the WHO guidance for recreational and drinking waters that outline threshold densities of cyanobacteria as a whole, rather than for individual species, in relation to threats to water usage (WHO, 1999; 2003; 2004). There is widespread acceptance amongst freshwater ecologists that cyanobacteria increase in abundance with increasing nutrient concentrations. Almost all published literature quantifying the relationship has, however, focused on the relative % abundance of cyanobacteria (e.g. Downing et al., 2001; Ptacnik et al., 2008). It is, however, the actual biomass of cyanobacteria that affects the provision of safe, clean water for recreation and water supply (WHO 2003; 2004). Our study specifically addresses, providing robust quantitative relationships between TP and actual cyanobacterial biovolume in European lakes and reservoirs. The exploratory analysis highlighted that cyanobacteria are generally absent, or in very low abundance, in low alkalinity lakes. This result was as expected, as the preference of cyanobacteria for neutral to alkaline waters is generally recognised and has been nicely demonstrated in previous in-lake experimental studies (Reynolds & Allen, 1968; Shapiro 1984). For this reason, we carried out further analysis on data from medium and high alkalinity lakes only. A previous study (Carvalho et al., 2011) of lakes in the UK, 97 of which were medium and high alkalinity, indicated that the mean response of cyanobacteria to TP was linear. Our current analysis of more than 800 medium and high alkalinity lakes is much more extensive, covering a broader nutrient gradient than that of Carvalho et al. (2011). In this study, the mean response indicated a non-linear relationship with TP. One reason for the better fit of the non-linear model to the mean response in this study appears to be because there were more lakes included with very low (<10 µg L^{-1}) and very high (>100 µg L^{-1}) TP concentrations, and the response appears to flatten out at these extremes. Even though the relationship between TP and the mean response of cyanobacteria biovolume was highly significant, it was evident that there was a large amount of scatter in the data. For this reason, modelling the mean response is not the ideal approach to adopt; quantile regression is more appropriate when several factors may limit a population at many sites (Cade & Noon, 2003).

Quantile models

There are many possible factors limiting cyanobacteria abundance in freshwaters and many of these, such as flushing rate or water colour are not routinely recorded. This reality, means that there will be unequal variation across a dataset when describing the relationship between a population response and only one of these factors. Examining a number of quantile responses allows us to compare how a range of cyanobacteria responses, from the minimum to maximum response, are affected by TP. This range in responses was demonstrated by the fact that linear models fitted the lower quantiles, a 2-parameter non-linear model was the best fit for the 25% and a 3-parameter non-linear models were the best fit for mean, median and higher quantiles. The fact that the lower quantile relationships between cyanobacteria biovolume and TP were more or less flat, and not significant, indicates that a small percentage of lakes always have no, or little, cyanobacteria, irrespective of TP concentrations. Clearly other factors limit cyanobacteria populations in the summer months in these lakes. This could include factors limiting population growth (e.g. limitation by nitrogen or light), or factors affecting population loss processes (e.g. flushing, grazing, parasitism) (Reynolds, 2006; Carvalho et al., 2011). For example, long-term monitoring of individual lakes has demonstrated that cyanobacteria are never abundant in lakes or reservoirs with a retention time <30 days (Reynolds & Lund, 1988).

Variability in the dataset also reflects the fact that a number of cyanobacterial genera contribute significantly to total biovolume in European lakes and these may be affected by TP in the water column, or other limiting factors, differently from each other and also differently in different lake types. For example, colonial gas-vacuolate genera, such as *Microcystis*, are known to migrate vertically in response to nutrient limitation, potentially allowing them to exploit deep, hypolimnetic sources of P, irrespective of epilimnion concentrations (Brookes & Ganf, 2001). The slopes and plateaus of the different quantile models all vary, although all models show that the biggest increase in cyanobacterial abundance occurs in the TP range from about 20 µg L^{-1} up to about 100 µg L^{-1}. This is an important finding for achieving successful restoration, as it indicates that nutrient concentrations need to be within this range before any significant declining response is likely to be observed in cyanobacterial abundance. The use of these different quantile responses to two specific applications for lake management in relation to recreational services are described further below.

Application of quantile responses for predicting bloom capacity

In the context of harmful cyanobacterial blooms, it is important to know the maximum cyanobacterial abundance that a lake environment could potentially support, rather than the mean or relative % abundance. Modelling the upper bounds of species–environment relationships relates much more to the most limiting resource (Cade & Noon, 2003; Vaz et al., 2008), for phytoplankton in many temperate lakes this is often phosphorus (Phillips et al., 2008; Schindler, 2008). The wide scatter of points around the mean or median responses clearly indicate that TP is not the single dominating factor limiting cyanobacterial abundance in lakes, but the higher quantile models may still better represent the capacity for cyanobacterial abundance in relation to phosphorus, given a lack of other limiting factors, such as loss rates to grazing or flushing. Using the 95% to represent the potential maximum capacity, our results clearly demonstrate that there are small probabilities for quite substantial cyanobacterial populations that exceed WHO (1999) health thresholds at relatively low TP concentrations. The fact that 5% of medium and high alkalinity lakes exceeded the low and medium risk thresholds at TP concentrations less than 35 µg L^{-1} supports anecdotal accounts of blooms in relatively nutrient poor waters that often cause surprise to local lake managers. At the other extreme, the 95% shows that cyanobacterial populations reach a maximum capacity of about 30 mm^3 L^{-1} at TP concentrations of about 150 µg L^{-1}. Further increases in TP have little effect on capacity, indicating that some other factor is limiting their abundance, most likely light-limitation (self-shading) is important at these very high phytoplankton densities.

Nutrient targets in relation to health thresholds

Given analysis of such a large population of lakes, the quantiles can also be used to represent the likelihood of cyanobacteria exceeding the WHO health alert thresholds for a given TP concentration. Although only a small proportion of lakes exceed the low risk threshold at low TP concentrations, the steepest rise in % exceedance occurs between TP concentrations of about 20 and 30 µg L^{-1} TP (approximately 10% to 25% exceedance). It was also clear that in about 50% of lakes the low risk threshold is not exceeded, irrespective of increasing TP concentrations above 100 µg L^{-1}. What level of precaution is chosen to focus nutrient management on is a local, social or political decision and will be affected by the use of the water body. A TP target of 20 µg L-1 should result in a low probability of risk (<10% exceedance) and may be appropriate for lakes or reservoirs of high importance for recreation or water supply. It is, however, important to point out that these nutrient targets are for guiding land management at a regional or European landscape scale, for a population of lakes and reservoirs. Given the approach used to derive them, these models are not applicable for individual lake basins, as discussed by. Reynolds (1980) in relation to the limitations of Vollenweider models for predicting chlorophyll concentrations in lakes.

For many years environmental managers have had robust nutrient targets in relation to phytoplankton chlorophyll for reducing phytoplankton biomass in general. Here we now present robust nutrient targets in relation to cyanobacterial abundance and their relationship to exceedance of WHO health thresholds for recreational waters. The nutrient targets could, therefore, be applied to current attempts at mapping ecosystem services provided by freshwaters, giving a good indication of the functional quality of waters in terms of recreational use. More broadly, the targets could also help with mitigation measures for sustaining services in relation to climate change. Although the quantile models provide more information about the variability in cyanobacterial response in relation to TP, they do not identify any reasons for the possible causes of variability, many site-specific factors may be important. Wagner & Adrian (2011) in a detailed single lake study highlighted that climatic factors only had significant positive effects on cyanobacterial dominance when TP concentrations rose above 70 µg L^{-1}. This further emphasises the importance of maintaining TP concentrations below this value to help mitigate against cyanobacterial blooms if future warmer climates prevail.

Acknowledgements

This paper was completed whilst Laurence Carvalho was on a secondment from the Centre for Ecology & Hydrology to the European Commission Joint Research Centre's Institute for Environment & Sustainability. It is a contribution to the PEER PRESS Project on mapping ecosystem services (http://www.peer.eu/projects/press/). The paper is a result of the project WISER (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (contract No. 226273), and is based on a database collated as part of the WISER Project. We would like to thank Bernard Dudley and Jannicke Moe who supported data management and extraction and especially thank all the data providers. These include (figures in brackets give the number of WBs):

Mediterranean GIG (Data manager: Caridad de Hoyos, CEDEX-CEH)
Spain: Ministerio de Agricultura, Alimentación y Medio Ambiente (122), Centro de Estudios Hidrográficos (CEDEX-CEH) (46). Italy: Università degli Studi di Sassari. Dipartimento di Scienze Botaniche, Ecologiche e Geologiche (DiSBEG) (18). Portugal: Instituto da Água, I.P. (INAG) (18). Romania: Ministeriul Meduli şi Pădurilor (MMP) (10). Cyprus: Ministry of Agriculture, Natural Resources and Environment, Water Development Department (MANRE-WDD) (7). France: Institut national de recherche en sciences et technologies pour l´environnement et l´agriculture (IRSTEA) (6). Greece: Maria Moustaka, Aristotle University of Thessaloniki.

Central-Baltic GIG (Data manager: Ute Mischke, IGB)
Estonia: Estonian University of Life Sciences (EMU) gathered for the state monitoring programme supported by the Estonian Ministry of Environment. Latvia: Sandra Poikane, Latvian Environment, Geology and Meteorology Centre. Lithuania - EPA Lithuania. Denmark: National Environmental Research Institute, University of Aarhus. Belgium: Jeroen Van Wichelen, Ghent University (UGENT). Netherlands: Rijkswaterstaat (RWS). Germany: Data from German water bodies were kindly provided by the following institutions of the German Federal States: Landesamt für Umwelt, Gesundheit und Verbraucherschutz Brandenburg (LUGV; 127), Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg-Vorpommern (MLUV, Seenprogramm, 65), Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt (LHW, 5), Landesamt für Landwirtschaft, Umwelt und ländliche Räume Schleswig-Holstein (LLUR, 13), Senatsverwaltung für Gesundheit, Soziales und Verbraucherschutz Berlin (SenGUV, 12), Niedersächsische Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN, Sulingen, 1). Poland: Data were provided partly by the Institute of Environmental Protection - National Research Institute. The Inspection for Environmental Protection provided data obtained within the framework of state environmental monitoring. Hungary: Environmental Protection Inspectorate for Trans-Tiszanian Region.

Northern GIG (Data manager: Geoff Phillips, EA)
Finland: Finnish Environment Institute (SYKE). Sweden: Swedish University of Agricultural Sciences (SLU). Norway: Norwegian Institute for Water Research (NIVA). United Kingdom - Scottish Environment Protection Agency (SEPA) and the Environment Agency for England & Wales (EA). Ireland: Environment Protection Agency (EPA).

References

Brookes JD & Ganf GG, 2001. Variations in the buoyancy response of Microcystis aeruginosa to nitrogen, phosphorus and light. Journal of Plankton Research, 23: 1399-1411.

Cade, B.S. and Noon, B.R. 2003. A gentle introduction to quantile regression for ecologists, Frontiers in Ecology and the Environment 1: 412–420.

Carpenter SR et al., 2009. Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment Proceedings National Academy of Sciences, 106, 1305-1312.

Carvalho L, Miller CA, Scott EM, Codd GA, Davies PS and Tyler A, 2011. Cyanobacterial blooms: statistical models describing risk factors for national-scale lake assessment and lake management. Science of the Total Environment, Accepted Sep 2011.

CEN, 2004. Water quality – Guidance standard for the routine analysis of phytoplankton abundance and composition using inverted microscopy (Utermöhl technique) CEN TC 230/WG 2/TG 3/N83, May 2004.

Codd GA, Morrison LF, Metcalf JS. Cyanobacterial toxins: risk management for health protection. Toxicology and Applied Pharmacology 2005; 203:264-272.

Dillon P.J.and Rigler F.H., 1974. The phosphorus-chlorophyll relationship in lakes. Limnology & Oceanography, 28: 792-795.

Dokulil MT, Teubner K. 2000. Cyanobacterial dominance in lakes. Hydrobiologia, 438: 1-12.

Downing JA, Watson SB, McCauley E.. Predicting cyanobacteria dominance in lakes. Canadian Journal of Fisheries and Aquatic Sciences 2001; 58:1905-1908.

Eigenbrod, F., Armsworth, P.R., Anderson, B.J., Heinemeyer, A., Gillings, S., Roy, D.B., Thomas, C.D. & Gaston, K.J., 2010. The impact of proxy-based methods on mapping the distribution of ecosystem services. Journal of Applied Ecology, 47, 377–385.

Hastie, T. and Tibshirani, R. 1990. Generalized Additive Models. Chapman & Hall.

Koenker, R. and Bassett, G. 1978. Regression Quantiles. Econometrica, 46: 33-50.

Koenker, R. and Park, B.J. 1994. An Interior Point Algorithm for Nonlinear Quantile Regression, Journal of Econometrics, 71: 265-283.

Koenker, R. 2009. quantreg: Quantile Regression. R package version 4.44. http://CRAN.R-project.org/package=quantreg.

Millennium Ecosystem Assessment, 2005. Ecosystems and Human Well-Being: Wetlands and Water Synthesis. World Resources Institute, Washington, DC. 68 pp.

Mischke, U. (2003). Cyanobacteria associations in shallow polytrophic lakes: influence of environmental factors. Acta Oecologica 24, Supplement 1(0): S11-S23.

Moe, S.J., Dudley, B., Ptacnik, R., 2008. REBECCA databases: Experiences from compilation and analyses of monitoring data from 5,000 lakes in 20 European countries. Aquatic Ecology, 42, 183-201.

Nõges, T. Laugaste, R., Nõges P. & Tõnno I., 2008. Critical N:P ratio for cyanobacteria and N2-fixing species in the large shallow temperate lakes Peipsi and Võrtsjärv, North-East Europe. Hydrobiologia 599: 77–86.

OECD. Eutrophication Of Waters. Monitoring, Assessment And Control, OECD, Paris; 1982. 156 pp.

Paerl, H.W. and Huisman J. 2008: Blooms Like It Hot. Science 320 (5872), 57 – 58

Phillips, G.L., Pietilainen, O.,Carvalho, L., Solimini, A., Solheim, A. and Cardoso, A., 2008. Chlorophyll - nutrient relationships: a statistical evaluation of typed lakes using a large European data set. Aquatic Ecology, 42, 213-226.

Ptacnik R., Lepistö L., Willén E., Brettum P., Andersen T., Rekolainen S., Lyche Solheim A. and Carvalho L., 2008a. Quantitative responses of lake phytoplankton to eutrophication in Northern Europe. Aquatic Ecology, 42, 227-236.

Reynolds C, 1980. Eutrophication and the management of planktonic algae: what Vollenweider couldn’t tell us. In: Sutcliffe DW, Jones G (eds) Eutrophication: research and application to water supply. Freshwater Biological Association, Ambleside, pp 4–29

Reynolds CS, 2006. Ecology of Phytoplankton. Cambridge University Press, Cambridge.

Reynolds CS Allen SE, 1968. Changes in the phytoplankton of Oak Mere, following the introduction of base rich water. British Phycological Journal, 3: 451-462.

Reynolds CS & Lund JWG, 1988. The phytoplankton of an enriched, soft-water lake subject to intermittent hydraulic flushing (Grasmere, English Lake District). Freshwater Biology,19:379-404.

Schindler DW, Hecky RE, Findlay DL, Stainton MP, Parker BR, Paterson MJ, Beaty KG, Lyng M & Kasian SEM, 2008. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences,105: 11254-11258.

Shapiro J., 1984. Blue-green dominance in lakes: The role and management significance of pH and CO_2. Internationale Revue der Gesamten Hydrobiologie, 69: 765-780.

Smith VH, 2003. Eutrophication of freshwater and Coastal Marine Ecosystems. A global problem. Environmental Science & Pollution Research, 10:126-139.

Vaz, S., Martin, C. S., Eastwood, P. D., Ernande, B., Carpentier, A., Meaden, G. J., and Coppin, F. 2008. Modelling species distributions using regression quantiles. Journal of Applied Ecology, 45: 204– 217.

Wagner C. & Adrian R., 2009. Cyanobacteria dominance: quantifying the effects of climate change. Limnology & Oceanography, 54: 2460-2468.

Weyhenmeyer, G.A., Adrian, R., Gaedke, U., Livingstone, D.M. and Maberly, S.C. 2002. Response of phytoplankton in European lakes to a change in the North Atlantic Oscillation Verh. Internat. Verein. Limnol. 28 1436–1439.

Willen E., 2000 (see Smith 2003)

Willen et al., 2000. (see Smith 2003)

WHO, 1999. Toxic Cyanobacteria in water: A guide to their public health consequences, monitoring and management. Edited by Chorus & Bartram, E & FN Spon, London.

WHO, 2003. Guidelines for Safe Recreational Water Environments. World Health Organisation, Geneva.

WHO, 2004. Guidelines for Drinking-water Quality, 3rd edn. World Health Organisation, Geneva.

Wood, S.N. 2006. Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC, Boca Raton.

Wood, S.N. 2008. Fast stable direct fitting and smoothness selection for generalized additive models. Journal of the Royal StatisticalSociety (B) 70: 495-518

Table 1. No. of lakes with cyanobacteria and TP data, by region, country and alkalinity type.

Region	Country	L	M	H	U	Total
Central	BE			9		9
	DE			223		223
	DK		1			1
	EE	3	5	46		54
	FR		3			3
	IE		1	33	10	44
	LT				39	39
	LV				63	63
	NL			47		47
	PL			49		49
	UK		3	64		67
Central Total		3	13	471	112	599
EC	HU			18		18
Med	ES	9	8	16	1	34
Northern	FI	104	47		5	156
	IE	6	2			8
	NO	308	147	44	3	502
	SE	97	21	7		125
	UK	51	12	1		64
Northern Total		566	229	52	8	855
Grand Total		**578**	**250**	**557**	**121**	**1506**

Table 2. Summary statistics of cyanobacterial abundance in different alkalinity classes of lakes (L = Low, M = Medium, H = High and U = Unknown) all lakes

Statistic	L	M	H	U	All
N	578	250	557	121	1506
Lower Quartile (25%)	0	0.001	0.061	0.060	0.002
Mean	0.18	0.93	3.95	3.41	1.96
Median (50%)	0.004	0.019	0.652	0.206	0.037
Upper Quartile (75%)	0.03	0.22	3.50	1.67	0.64
90th Percentile	0.07	2.15	10.13	9.05	4.95

Table 3 AIC values for both linear and non-parametric quantile regression models relating cyanobacterial biovolume to TP concentrations in medium and high alkalinity lakes.

Quantile

Model type	**0.05**	**0.10**	**0.25**	**0.50**	**0.67**	**0.75**	**0.83**	**0.90**	**0.95**
Linear quantile	-185.5	-99.1	160.2	549.6	812.2	952.1	1110.1	1315.8	1560.9
Non-parametric quantile	6.5	20.4	205.1	496.4	685.4	798.6	962.2	1184.2	1427.4

Table 4 Parameter estimates derived using non-linear quantile regression for medium and high alkalinity lakes. Estimates for non-linear mean response also shown.

Model	Deviance	Parameter a ±s.e.	Parameter b ±s.e.	Parameter c ±s.e.
0.25	**61.95**	**-5.41 ± 0.42**	**1.04 ± 0.38**	
0.50	**97.23**	**0.47 ± 0.05**	**1500579 ± 0**	**8.97 ± 0.23**
mean (non-linear)	**102.92**	**0.56 ± 0.03**	9493 ± 15020^{NS}	**6.23 ± 1.12**
0.60	**100.91**	**0.64 ± 0.06**	**86850 ± 0**	**7.18 ± 0.18**
0.67	**98.90**	**0.80 ± 0.05**	**99913 ± 0**	**7.38 ± 0.21**
0.75	**90.22**	**0.92 ± 0.04**	**98649 ± 0**	**7.78 ± 0.16**
0.83	**75.51**	**1.03 ± 0.05**	**17577 ± 0**	**6.79 ± 0.15**
0.90	**55.48**	**1.28 ± 0.08**	**3695 ± 0.4**	**5.77 ± 0.17**
0.95	**34.15**	**1.51 ± 0.07**	1219 ± 529*	**5.23 ± 0.36**

*Coefficients in bold all highly significant ($p<0.01$), *=$p<0.05$, NS=not significant*

Table 5. 95% quantile fitted values showing the changing Cyanobacterial biovolume (mm^3 L^{-1}) with change in total phosphorous (µg L^{-1}). The fitted quantile reaches an asymptote at 31.5 mm^3 L^{-1} of cyanobacteria biovolume.

Total Phosphorus (µg L^{-1})	**Cyanobacteria capacity (95%) (mm^3 L^{-1})**
0	0
10	0.5
12	1
16	2
24	5
32	10
50	20
150	30
350	31

Table 6. TP concentrations for a given likelihood (quantile) of being below low and medium risk WHO threshold levels for cyanobacteria volume. TP concentrations are obtained from the fitted quantile regression models to the medium and high alkalinity lakes.

WHO Threshold	Quantile	% exceeded	TP
Low	0.57	43	57.8
	0.60	40	54.4
	0.63	37	45.8
	0.67	33	41.0
	0.75	25	30.7
	0.78	22	29.4
	0.83	17	26.2
	0.87	13	22.8
	0.90	10	21.6
	0.95	5	16.3
	0.98	2	13.2
Medium	0.87	13	58.3
	0.90	10	47.7
	0.95	5	32.4
	0.98	2	22.7

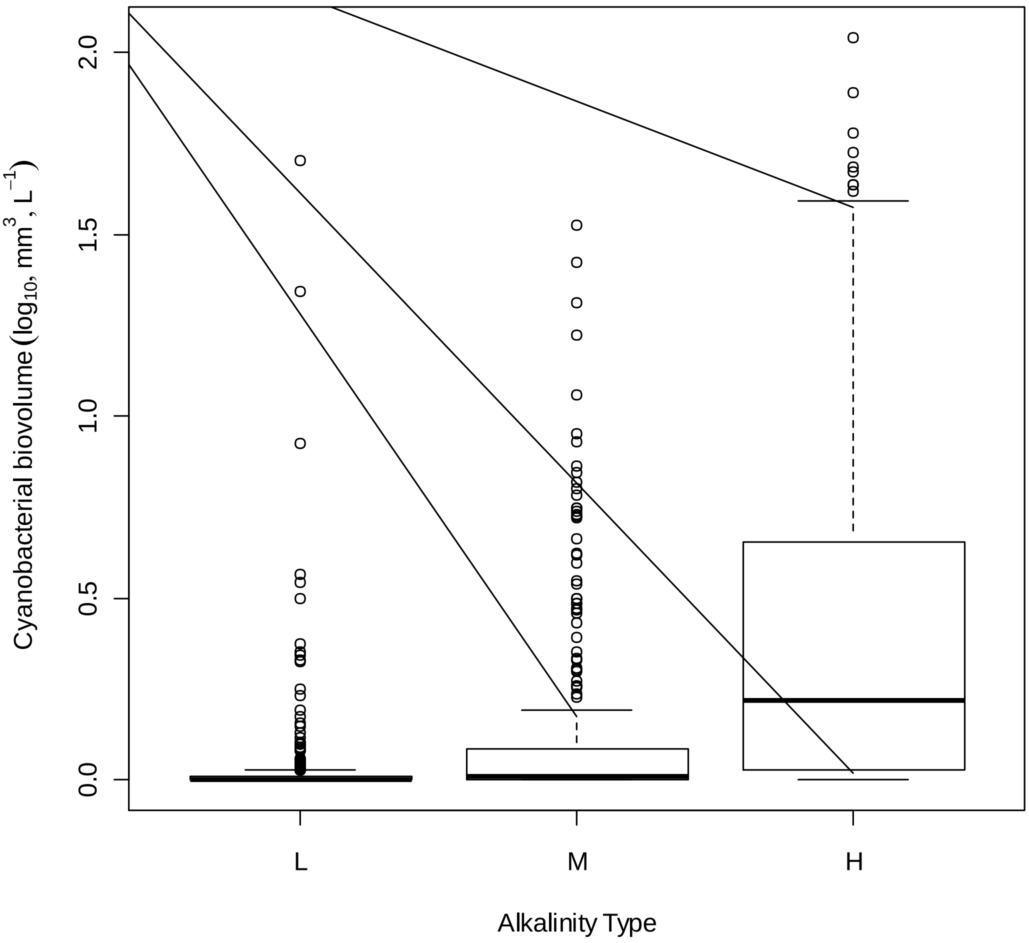


Fig. 1. Boxplot of cyanobacterial biovolume ($\log_{10}$ mm^3 L^{-1}) in lakes of low, medium and high alkalinity (<0.2, 0.2-1.0, >1.0 mequiv. L^{-1} respectively).

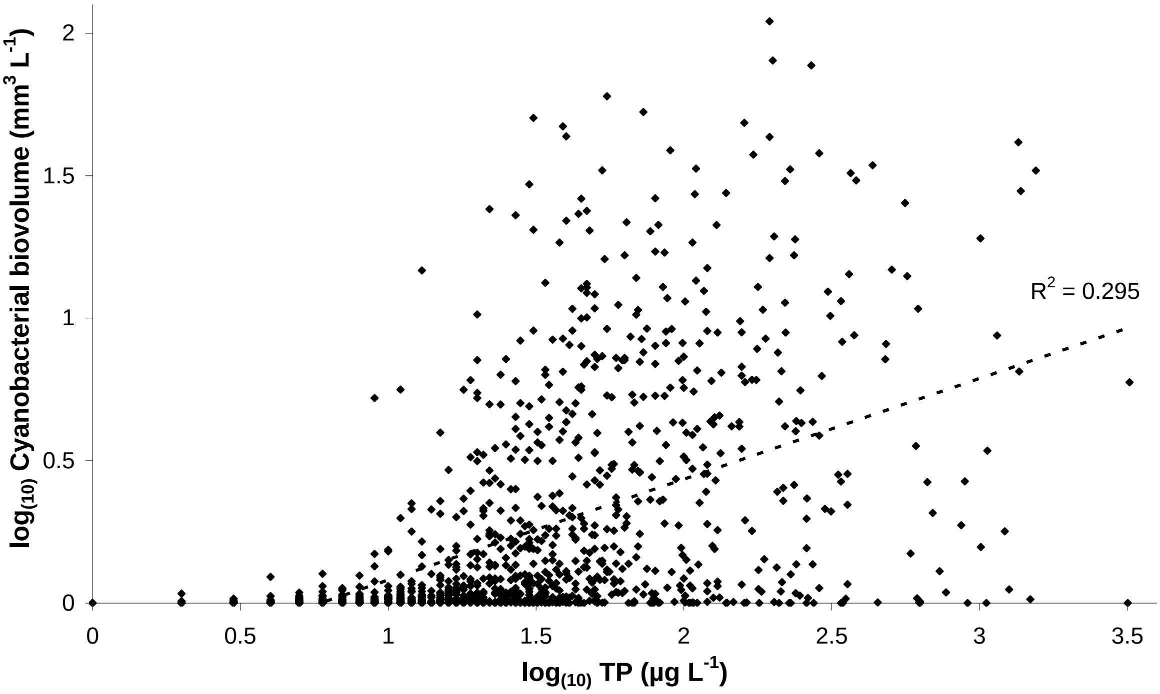


Fig. 2 Scatter plot for $\log_{10}$ *cyanobacteria and* $\log_{10}$ *total phosphorus (µg* L^{-1}*) with fitted linear regression*

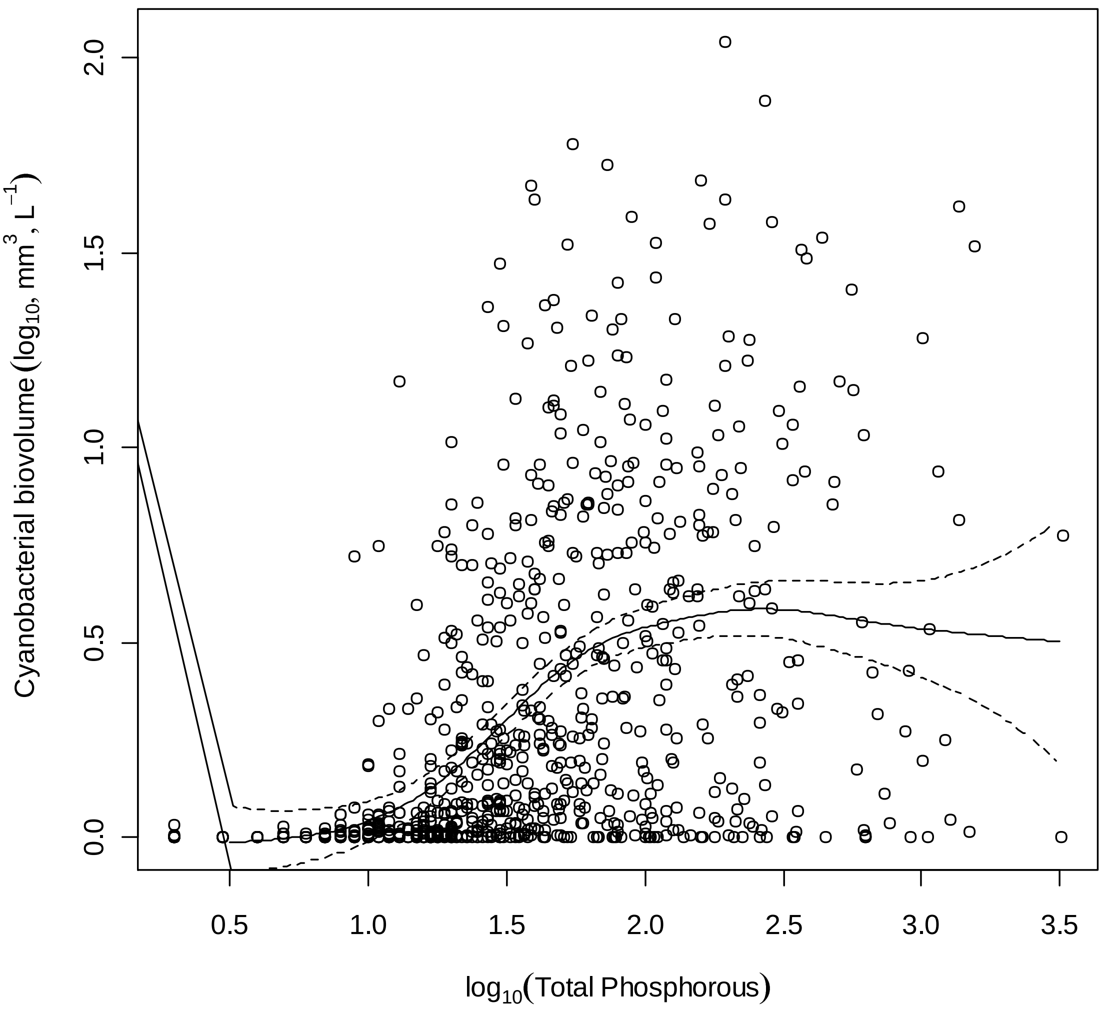


Figure 3. GAM for cyanobacteria biovolume in response to total phosphorus for medium and high alkalinity lakes.

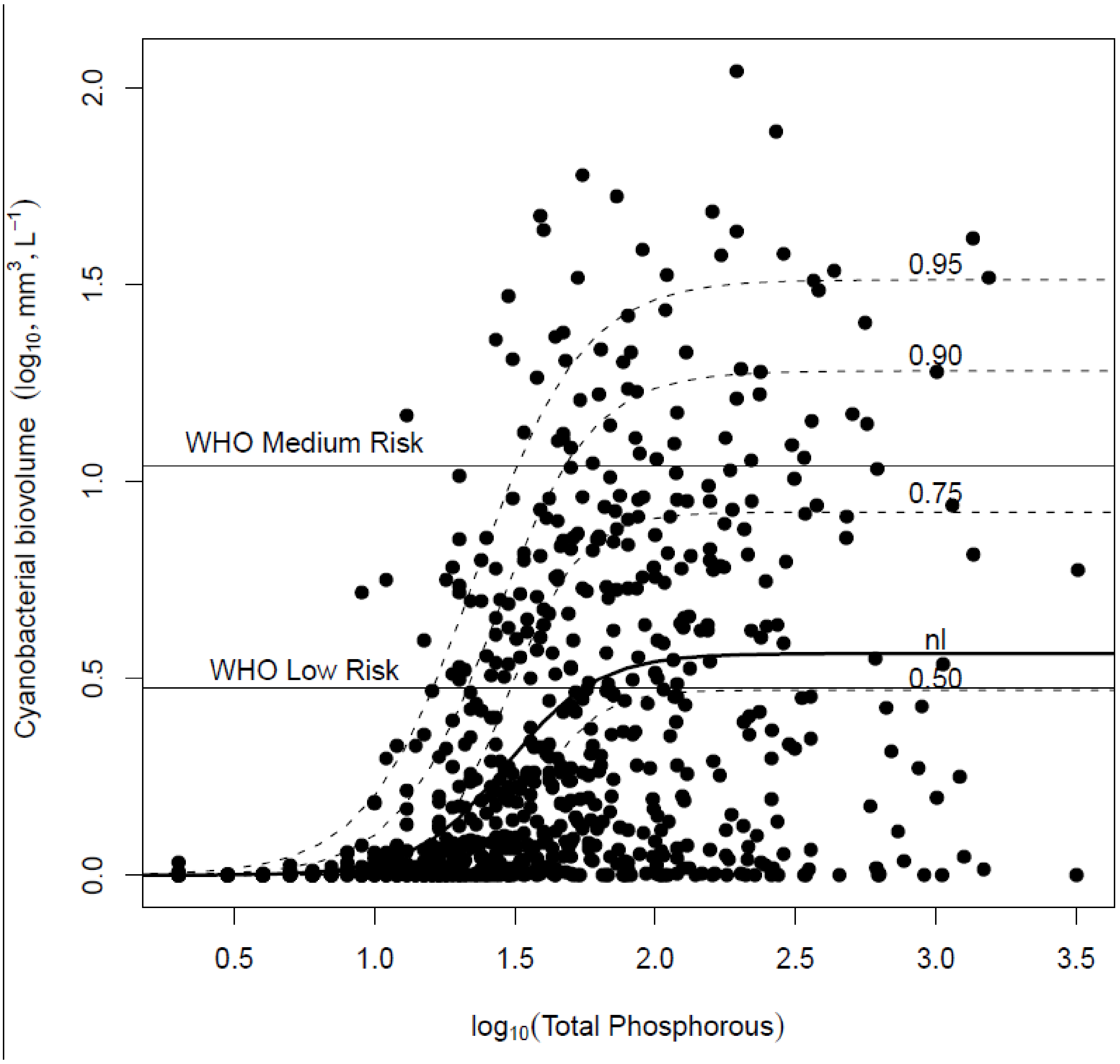


Figure 4. Scatter plot for $\log_{10}$ cyanobacteria and $\log_{10}$ total phosphorus for medium and high alkalinity lakes. Quantile regression curves (0.50 – 0.95) using a fitted 3-parameter sigmoid non-linear model are displayed. Nl = Non-linear regression fit to mean of data. Thresholds relating to approximate WHO (2003) low and medium risk thresholds are also indicated.

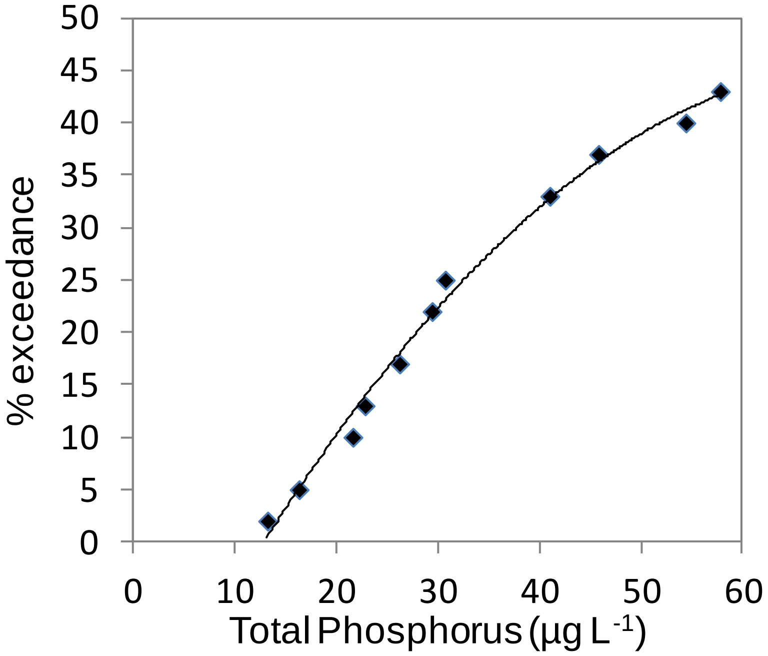


Figure 5. Relationship between % lakes exceeding WHO low/medium risk threshold for cyanobacterial biovolume (2 mm^3 L^{-1}) in relation to total phosphorus (TP)

Maileht et al. (2012). Water colour, phosphorus and alkalinity are the major determinants of the dominant phytoplankton species in European lakes.

Water colour, phosphorus and alkalinity are the major determinants of the dominant phytoplankton species in European lakes

Kairi Maileht[1], Tiina Nõges[1], Peeter Nõges[1], Ingmar Ott[1], Ute Mischke[2], Laurence Carvalho[3], Bernard Dudley[3]

[1] Estonian University of Life Sciences, Institute of Agricultural and Environmental Sciences, Centre for Limnology, 61117 Rannu Parish, Tartu County, Estonia.

[2] Leibniz-Institute of Freshwater Ecology and Inland Fisheries Department of Shallow Lakes and Lowland Rivers. Justus-von-Liebig-Str. 7. D-12489 Berlin, Germany.

[3] Centre for Ecology & Hydrology (CEH), Bush Estate, Penicuik, Midlothian EH26 0QB, UK.

Corresponding author: kairi.maileht@emu.ee

Keywords: phytoplankton, geographical gradients, nutrients, CCA, alkalinity, water colour

Abstract

Rather common idea is that the summer dominant species in lake phytoplankton are mostly correctly identified with well-known ecological requirements, studying their occurrence over broad geographical scales. This gives a robust picture about the distribution of habitat conditions in lakes. Based on climatic gradients in Europe and known geographic differences in water chemistry, we hypothesized that temperature, trophic state, water colour, and alkalinity could be the major determinants of the occurrence of different dominant species in European lakes. We included the 151 most common summer phytoplankton dominants found from 1558 lakes in 20 European countries in a Canonical Correspondence Analyses together with data on nutrients, water colour, alkalinity, and lake morphometry. As expected, the cloud of the dominant species had a strongly elongated shape in the direction determined by water colour on one hand and alkalinity and total phosphorus on the other hand. Contrary to our hypothesis, water temperature from July to September had only a negligible impact on the distribution of dominants, showing the prevalence of rather homogeneous thermal conditions throughout Europe for this period of year. A comparison of the data from the northern ecoregion with the rest of Europe showed that chrysophytes and cryptophytes occurred more frequently among dominants in the north whereas cyanobacteria and dinophytes in the southern parts of Europe. Our analysis suggest that besides trophic conditions, other hydrochemical variables, such as alkalinity and the content of humic substances, have at least as important role in determining the distribution of the dominant phytoplankton species in European lakes.

Introduction

Despite continuous efforts of generations of algologists, the biogeographical distribution of freshwater phytoplankton diversity and its driving factors are still largely unknown. Padisák et al. (2003) pointed out the often different research targets in small and large lakes as one of the controversies in lake phytoplankton research. While taxonomic and floristic work has a stronger focus in small lakes, the bulk of our knowledge on the ecology of phytoplankton originates from middle-sized or large lakes. The major problem has been the absence of phytoplankton databases of a comparable resolution and harmonized taxonomy covering continental scales. Large variability is driven by local environmental factors that vary along latitudinal, longitudinal and altitudinal gradients and differs in lakes of different morphometry. Phytoplankton abundance and taxa richness might display some regularities in their spatial distribution, but these patterns are often blurred because of regionally different taxonomic resolution, counting routines and traditions in taxonomic work.

Nevertheless, a number of important studies have been carried out based on analyses of data compiled from published sources. Using this approach, Watson et al. (1997) studied the patterns in phytoplankton taxonomic composition across 91 temperate lakes of differing nutrient status. Dodson et al. (2000) investigated the relationship between species richness and productivity in a survey of 33 well-studied lakes on different continents for which data on six major taxonomic groups were available.

Several studies of phytoplankton community structure have been carried out on regional or country-scale. For example, based on 165 lakes located throughout Florida, Duarte et al. (1992) demonstrated the existence of smooth gradients of change in phytoplankton community structure with increasing lake trophic status, from dominance by green algae in oligotrophic lakes to dominance by cyanobacteria in eutrophic and hypertrophic lakes, with a peak in diatom abundance in mesotrophic lakes. Willén et al. (1990) studied summer phytoplankton in 73 nutrient-poor Swedish lakes. Dominant species and functional assemblages in late summer phytoplankton were studied in 80 Hungarian small shallow lakes (Padisák et al., 2003). Lepistö et al. (2004) studied type-specific and indicator taxa of phytoplankton as a quality criterion for assessing the ecological status of Finnish boreal lakes. Ptacnik et al. (2008) studied the responses of three major phytoplankton classes to eutrophication in lakes from Scandinavia and the United Kingdom and found significant differences between humic and clearwater lakes, and between low- and moderately alkaline lakes.

The only successful attempt to expand a species-level analysis to a continental scale was by Stomp et al. (2011), who analyzed comparable data collected in 1973-1975 from 540 lakes and reservoirs on the continental United States and found strong latitudinal, longitudinal and altitudinal gradients in phytoplankton biodiversity.

In Europe the implementation of the Water Framework Directive (Directive, 2000) has given a new impetus to freshwater ecological studies at species and community levels and the need for comparisons over broad geographical ranges shifted to the forefront of research. Chemical and biological data from more than 5,000 lakes in 20 European countries were compiled into databases within the EU REBECCA Project (Moe et al., 2008) and complemented by new data during the EU WISER Project (www.wiser.eu). This database is now the largest combined dataset on phytoplankton composition in Europe. It is always challenging to study phytoplankton community responses to changes in the environment due to the high variability of phytoplankton species structure. Common ways to cope with this complexity is to “boil it down” to major taxonomic groups (e.g. Duarte et al., 1992; Ptacnik et al., 2008), functional groups (Reynolds et al., 2002, Padisák et al., 2003) or strategist groups (Grime, 1979; Reynolds, 1988) or to calculate various indices to characterize different aspects of the community structure, such as diversity (Shannon, 1948; Simpson, 1949; Margalef, 1958) or evenness (Pielou, 1975). For many purposes even the simplest parameter, the number of species (Hill, 1973), may be the most useful measure of local or regional diversity.

In our study, we decided to follow a different approach and focus on the dominant species. Studying the dominants is interesting for several reasons. The stability of ecological communities depends upon the population dynamics of the dominant species (Grime, 1998; Flöder et al., 2010). As the winners of competition for resources, the dominants can give a robust picture of resource availability. In this respect, studying the summer phytoplankton is most promising, as community equilibria occur most prominently during summers when less flushing allows competitively stabilized associations to develop (Padisák et al., 2003). Many of the dominants tend to be nuisance species, so the distribution and understanding of their controlling factors remains a high priority research topic. Finally, selecting the dominant species should guarantee reduced taxonomic uncertainty as researchers tend to pay more attention to abundant species and their high abundance in the sample should minimize misidentification errors as sufficient material is observed to cover phenotypic variability in the species.

A recent analysis of the 1,337 lakes included in the European Environment Agency (EEA) database (Nõges, 2009) showed that lakes at higher latitudes are larger but shallower and have smaller catchment areas. Northern lakes have lower alkalinity, pH and conductivity, and also lower concentrations of nitrogen and phosphorus while the concentration of organic matter is higher compared to southern lakes. Several gradients in lake environments were found also along longitudinal and altitudinal scales. As Europe extends from arctic to sub-tropical areas, and from maritime to continental climates, the temperature and ice regimes of inland waters vary within a wide range.

The present study aims to assess the impact of hydrochemical, climatic, and morphometric factors affecting lake environments over broad geographical scales, on the dominant taxa of lake phytoplankton and their functional attributes. As differences in humic matter content and alkalinity have been shown to be the major factors modifying phytoplankton response to eutrophication within the Nordic countries (Ptacnik et al., 2008), we hypothesise that expanding the geographical range to the south and west and focusing on dominant species, we should see an even stronger impact of these factors. We also hypothesise that the effect of water temperature will be clearly manifested in the occurrence of different phytoplankton dominants.

Materials and methods

Twenty countries have provided data to the EU 7th Framework Programme project WISER database. Data has been gathered during a long period (1972-2009). July, August and September were selected for the species analysis, comprising a total of 6120 samples from 1558 water bodies from Belgium (BE) 11, Cyprus (CY) 7, Germany (DE) 217, Denmark (DK) 64, Estonia (EE) 46, Spain (ES) 135, Finland (FI) 156, France (FR) 5, Greece (GR) 1, Hungary (HU) 13, Ireland (IE) 40, Italy (IT) 14, Lithuania (LT) 36, Latvia (LV) 58, The Netherlands (NL) 43, Norway (NO) 401, Poland (PL) 39, Romania (RO) 10, Sweden (SE)

113 and United Kingdom (UK) 149. More than half of these data (62%) originated from the last ten years.

Phytoplankton samples were analyzed according Utermöhl technique (CEN EN 15204, 2006). Very seldom additional slides preparation for identification of diatom species was carried out in parallel, so diatom taxa list is based on different approaches. Therefore for example *Cyclotella* taxa are distinguished only for a small proportion of all samples on species level. The European WISER phytoplankton list was created as an operational list to merge European data (http://www.freshwaterecology.info/). So the phytoplankton list is not kept up to date with new valid names, but was harmonized for a status available in common determination keys in Europe in 2010.

We focused our study on the most dominant taxa, which we defined as the single species with the largest biovolume from each sample. Only species which were recorded in at least five samples in the dataset were included. We also examined the dominant species in terms of their belonging to 11 algal classes: Bacillariophyceae (Bac), Chlorophyceae (Chlor), Chrysophyceae (Chrys), Conjugatophyceae (Conj), Cryptophyceae (Crypt), Cyanophyceae (Cyan), Dictyochophyceae (Dict), Dinophyceae (Dino), Euglenophyceae (Eug), Prymnesiophyceae (Prym) and Raphidophyceae (Raph) and assigned them to functional groups according to Reynolds et al. (2002) and Padisák et al. (2009).

To study the occurrence of dominant species in Europe, we split the data into two parts – countries belonging to the Nordic Geographical Intercalibration Group (N-GIG: FI, SE, NO, IE, part of UK) and the rest. The GIG boundaries were delineated within the WFD implementation process and reflect the eco-regions which share common types of surface water bodies. This split divided the data into more or less comparable parts with 4071 samples collected from 859 N-GIG lakes and 2049 samples collected from 699 lakes located in the rest of Europe. The fact that many lakes were represented by a number of samples in which the dominant species could be either the same or different, complicated the calculation of occurrence frequencies of different dominant species: calculation by lakes became impossible whereas calculating by samples would have caused a bias towards lakes for which there were more samples in the database. To overcome this, we considered the occurrences of different dominant species in the same lake as different counting units or occasions, but if the same species dominated in all samples from a lake, it was considered as one counting unit. We got 1897 such counting units for N-GIG and 1341 counting units for the rest of Europe that were analyzed for the frequency of dominant species belonging to different algal classes.

The database included the following environmental parameters: latitude, longitude, altitude, alkalinity, maximum depth, mean depth, surface area, colour, total nitrogen (TN), total phosphorus (TP) and water temperature. We used data on phytoplankton species biovolumes in order to select the dominant taxa by the largest biovolumes for each sample.

For evaluation of the relationships between the distribution of dominant species and the environmental variables, we ran a Canonical Correspondence Analysis (CCA), using the multivariate statistical package (MVSP; KCS, 2007).

Results

Among the most abundant phytoplankton species by biovolume we found altogether 151 taxa, 130 of which were identified to species level and 21 to genus level. We handled all of them as unique taxa. Occurrence of dominant species by country and information of functional groups (Padisák et al., 2009) is presented in table 1.

The distribution frequency of the taxa among the 20 countries had a positive skew (Fig. 1) with only 2-4 countries sharing many of the dominant species. The most widespread taxa, occurring as dominant in 15 or more countries, were *Ceratium hirundinella, Cyclotella* sp., *Aulacoseira granulata* and *Cryptomonas* sp. Among dominants, 132 taxa occurred in N-GIG lakes, with 29 of these being restricted only to the N-GIG and 126 dominant taxa in the rest of Europe with 16 taxa restricted to this area. Hence about 2/3 of taxa occurred as dominant in both parts of Europe.

The division of the dominant taxa between algal classes was rather similar in the two parts of Europe (Fig. 2). There were slightly more diatom, chrysophyte and chlorophyte taxa and slightly less cyanobacterial taxa among dominants in the north than in the south. The same differences appear much stronger in terms of the frequency of occurrence of dominants between the two regions. Chrysophytes occurred 3 times and cryptophytes nearly twice more frequently among dominants in the N-GIG than in the rest of Europe and cyanobacteria and dinophytes occurred about twice less frequently. The biggest difference, however, was revealed for *Gonyostomum semen*, the single representative of raphidophyta, which dominated in N-GIG lakes 5 times more frequently than elsewhere.

The first two CCA axes accounted for 62.1% of the total variance, with the Axis 1 alone explaining 33.4%, indicating a relevant gradient in the data set. The CCA biplot (Fig. 3a,b) revealed a strongly intercorrelated group of factors describing lake morphometry (mean depth, maximum depth, surface area), which was positively related to altitude and negatively to TP and TN. Water colour increased strongly with increasing latitude and longitude and was negatively related to alkalinity. Water temperature, which was the weakest among variables, increased with decreasing lake size and depth and being directed almost perpendicularly to the arrows of latitude and longitude, showed no relationship with the geographic location.

The cloud of the dominant species had a strongly elongated shape in the direction determined by water colour on one hand and alkalinity and TP on the other hand. Taxa associated with high water colour were in rank order *Crucigenia tetrapedia*, *Peridinium umbonatum var. goslaviense* and *Urosolenia longiseta* and those associated with high alkalinity *Aphanizomenon aphanizomenoides, Cryptomonas curvata* and *Staurastrum pingue.* The three species associated most strongly to large lake size and greatest depth, were the diatoms *Asterionella formosa, Tabellaria fenestrata* and *Cyclotella comensis.*

If plotted by algal classes, the stronger dependence of diatoms on lake morphometry compared to other algal classes was expressed in the much broader vertical spread of the cloud (not shown). Chrysophytes (Fig. 4) instead had a strongly skewed distribution towards increasing latitude and water colour.

Among the functional groups, large motile species in eutrophic (Lm) and in mesotrophic (Lo) conditions were distributed along colour, latitude, longitude and alkalinity gradients, but with the Lm group clearly associated with more alkaline and eutrophic lakes, whilst Lo group was observed to dominate in high colour/latitude and low alkalinity/TP lakes (Fig. 5a). The cloud of functional groups of small nannoplankton (X1, X2, X3) was clearly elongated in the direction of colour/latitude, with X3 dominating more frequently in clear water alkaline/eutrophic lakes at lower latitudes and X1 and X2 in lower alkalinity/TP and more coloured lakes at higher latitudes (Fig. 5b). Distribution of other functional groups (not shown) was much broader.

Discussion

Dominant taxa

The fact that, in general, about 2/3 of the dominant species in Northern Europe were the same taxa that dominated in the rest of Europe is surprising. This suggests that broad geographical-scale gradients, such as the effects of climate and daylength or length of growing season are less important in determining the dominant species than more local lake- and catchment-specific factors, such as depth and alkalinity. Less surprising is that dominant species spanned many algal classes reflecting a diverse range of lake types and broad alkalinity and nutrient gradients across Europe. The fact that chrysophytes, cryptophytes, diatoms and raphidophytes were more frequently dominant in Northern Europe, whilst cyanobacteria and dinoflaggelates more frequently dominated Central and southern Europe reflected the broad distinction between Northern Europe and the rest of Europe. Northern lakes are generally larger and shallower with smaller catchment areas, lower alkalinity, pH and conductivity and with less nutrients and more dissolved organic compounds than southern lakes (Nõges, 2009). Chrysophytes are common in softwater lakes with low or moderate productivity and lakes with low pH (Nicholls & Wujek, 2003), which is in good correspondence with our analyses (Fig. 3). Cryptophytes are common species with a widespread distribution in many lake types, but our analysis supports individual lake studies that show they often dominate in the summer and autumn in humic lakes (Arvola et al., 1999). The higher frequency of cyanobacteria and dinophytes (Fig. 2) as dominants in southern Europe is clearly explained by the distribution of lakes of higher trophic state and higher alkalinity in this region. The impact of these specific gradients in geography, morphology and water quality on species and functional groups are discussed in more detail below.

Colour, latitude and longitude

Latitude, longitude and colour gradients in European lakes are correlated, but the strongest factor is colour, since the others describe location. The majority parts of Scandinavian lakes have acid and coloured waters in correspondence with catchment areas covered mostly by forests, swamps and mires. Thin soils lie directly on bedrock and buffer capacities are relatively low. Hereafter colour and latitude, and to a lesser degree longitude, should be considered together, not as separate parameters.

Many taxa, common only in highly productive lakes, are also more frequently recorded in lakes of higher humic content (Arvola et al., 1999). Arvola et al. (1999) presented a list of species, more frequently occurring in brown coloured lakes. These species also occurred as the dominant species in our dataset: *Acanthoceras zachariasii*, *Anabaena planktonica*, *Aphanizomenon flos-aquae*, *Eunotia zasuminensis*, *Mallomonas caudata*, *Melosira varians*. *Gonyostomum semen*, *Botryococcus braunii*, *Crucigenia tetrapedia*, *Tabellaria flocculosa*, *Monoraphidium griffithii*, *Dinobryon pediforme*, *Synura* sp., *Aulacoseira alpigena*, *Spondylosium planum*, *Peridinium umbonatum*, *Urosolenia longiseta* and *Aulacoseira italica*. *Anabaena lemmermannii* is known as a characteristic species of soft water lakes (Ott & Kõiv, 1999). *Trichormus catenula* is widely distributed (Zabelina et al., 1951). Most of these species dominate in northern parts of Europe, in countries like NO, SE, FI, UK, EE and DK, with some exceptions like *Trichormus catenula* and *Synura* sp.

Rosen (1981) identified *Oocystis submarina* (Arvola et al., 1999) and small naked chryso- and dinoflagellates as typical of humic conditions. *Anabaena macrospora* and *Woronichinia* *compacta* are also common in northern temperate zone (Komárek & Anagnostidis, 1999; Komárek & Zapomelova, 2008).

Gonyostomum semen is a well-known nuisance alga with widespread distribution (Figueroa & Rengefors, 2006) and has increased in Scandinavian (Willén, 2003; Figueroa & Rengefors, 2006; Trigal et al., 2011) and Baltic soft water lakes (Rakko et al., 2008). Our database revealed that *Gonyostomum semen* was a dominant species of soft water lakes (alkalinity between -0,067 to 1,055 meq/l) in DK, EE, ES, FI, LV, NO, SE and UK. The highest biovolumes of *G. semen* were recorded in SE, NO, EE and DK brown water lakes (average water colour of these lakes was 112 mg/l Pt). Our analysis highlighted other species as being capable of dominating dark acid waters, like *Chrysosphaerella longispina* (Dillard, 2008; Trigal et al., 2011), *Botryococcus terribilis* (Trigal et al., 2011), *Peridinium inconspicuum* (Willén, 2003) and *Dinobryon sociale var. americanum* (Canter-Lund & Lund, 1995). Species which showed a good relationship with longitude, like *Dinobryon pediforme,* are also common in acid lakes (Willén, 2003).

Polyhumic lakes usually have a very specific phytoplankton composition, where dominant species are adapted to low light and large fluctuations and gradients of temperature and oxygen. Generally a moderate increase of humic content results in higher phytoplankton biovolume (Arvola et al., 1999; Carvalho et al., 2008; 2009), while in polyhumic lakes (>100 g Pt m^{-3}) this trend stops. Moderate content of humic matter seems to affect positively phytoplankton abundance. One explanation of this is that environmental resources are enriched in coloured waters. If besides moderate humic matter, mineral nutrients are enriched, and there are enough C resources, phytoplankton have been shown to be more rich in comparison with low coloured lakes (Ott & Kõiv, 1999).

At low latitudes and longitude, i.e. southern and western Europe there is another cluster of species like *Planctonema lauterbornii*, *Dictyosphaerium subsolitarium*, *Cyclotella ocellata*, *Mougeotia* sp., *Coenochloris fotti* and *Cryptomonas erosa*. All these species had their peak biovolume in southern countries, most of these in ES. One of these species, *Planctonema lauterbornii,* has shown in other studies a strong relationship with temperature (Gomes et al., 2004).

Water temperature, TP and alkalinity

Our analysis of the distributions of the dominant phytoplankton functional groups provides strong empirical support for some of the expert-judgement based associations outlined in Padisak et al. (2009) and Reynolds et al. (2002). In particular, it supports the opinion that the Lm group is associated generally with more enriched lakes than the Lo group (Fig. 5a). Interestingly our analysis also shows that the bulk of the Lm taxa have a tendency to dominate in smaller and shallower lakes, whereas Lo taxa were more frequently dominant in larger, deeper lakes. The small nanoflaggelate functional groups X1, X2 and X3 were most frequently dominant in a range of lakes along a North-East to South-West gradient across Europe (Fig. 5b) but particularly related to lakes of lower alkalinity and TP where nutrients are less available and where taxa that can feed heterotrophically may be particularly favoured. *Aphanizomenon aphanizomenoides* and *Cylindrospermopsis raciborskii* showed very good relationships with these vectors. The wide distribution of *C. raciborskii* and *A. aphanizomenoides* in the temperate zone is widely cited as a response to global warming (Briand et al., 2004; Stüken et al., 2006). *C. raciborskii* is a common species in tropical and pantropical regions (Cronberg & Annadotter, 2006). Our analyses showed that *C. raciborskii* was a dominant species of samples from ES, HU and NL. This species has rapidly increased all over the world from tropical to temperate zones (Fabbro & Duivenvoorden., 1996; Chapman & Schelske., 1997; Lagos et al., 1999; Shafik et al., 2001; Briand et al., 2004; Valerio et al., 2005; Bouvy et al., 2006; Fastner et al., 2007; Moustaka-Gouni et al., 2009; Alster et al., 2010; Kokociński et al., 2010; Moisander et al., 2012) except Antarctica (Padisák et al., 2003). *C. raciborskii* prefers highly eutrophic waters, when water temperature is high and light conditions are poor (Moustaka-Gouni et al., 2006; 2009), but it can also survive in water bodies with lower trophic status, because of its effective storage capacity for phosphorus. This species can dominate under different abiotic conditions, like high concentration of dissolved minerals or salinity, but temperature appears to be the most important factor. *A. aphanizomenoides* is also recorded from tropical and subtropical regions, but has expanded its distribution to the temperate zone (Stüken et al., 2006). Our database showed that *A. aphanizomenoides* was a dominant in DE and ES.

Assemblage of water temperature, TP and alkalinity also showed a very good relationship with some cyanobacteria like *Microcystis flos-aquae*, *Anabaena viguieri*, *Aphanizomenon gracile*, *Planktothrix agardhii*, *Pseudanabaena limnetica*, *M. viridis*, *Limnothrix redekei*, *Chroococcus limneticus* and *Anabaena danica*. All these species, except *C. limneticus*, are known in meso- and eutrophic water bodies and may form water blooms (Mischke & Nixdorf 2003; Nixdorf et al., 2003; Reynolds et al., 2002; Cronberg & Annadotter, 2006; Willén, 2007). Phillips et al. (2010) classified phytoplankton genera (some genera were divided to more detailed groups) into very tolerant, tolerant, sensitive and very sensitive taxa. Many of the cyanobacterial genera were classified as very tolerant or tolerant. Only *Chroococcus* sp. was classified as a slightly sensitive genus. Dominance by *Euglena* sp. also showed a strong relationship with water temperature and TP. This taxon does not generally reach a big biovolume in large lakes, but in small lakes their biovolume can be great (Padisák et al., 2003), as was supported in this study. *Pandorina morum* and *Cryptomonas curvata* are common in nutrient rich water bodies (Reynolds et al., 2002; Padisák et al., 2003) and *C. curvata* is tolerant of low light (Reynolds et al., 2002). *Ceratium furcoides* and *Aphanizomenon gracile* dominance showed strong relationships with alkalinity in our study but both species are described by Reynolds et al. (2002) as tolerant of low carbon concentrations, although this may be the case in waters of very high alkalinity.

Lake morphometry (surface area, mean and maximum depth) and altitude

Lake morphometry and altitude appeared important in favouring the dominance of the following species in rank order: *Asterionella formosa*, *Tabellaria fenestrata*, *Cyclotella comensis* and *Dinobryon bavaricum*. *Asterionella formosa* and *Tabellaria fenestrata* showed very strong relationship with mean depth, surface area and maximum depth. Both species are dominant in deep lakes with large surface area. Despite heavy frustules they are best adapted to float in the water column due to long, thin cells, or the belt- and star-like structure of their colonies. *Cyclotella comensis* showed good relationship with surface area. This species is common in alpine lakes in Switzerland and France (Zabelina et al., 1951). Our analyses showed that this alga was dominant in higher altitude locations of ES, IT, NO and SE (with average altitude 397.6 m). Also the average maximum depth of these lakes was relatively high (164.7 m).

Conclusions

We found 151 phytoplankton taxa mostly identified to species level, which occurred as the most dominant taxa by biovolume at least in five of the 6120 samples collected between July and September from 1558 lakes in 20 countries of Europe.

2/3 of the dominant species in Northern Europe (including Finland, Sweden, Norway, Ireland and part of UK) were the same taxa that dominated in the rest of Europe. The dominant species spanned all algal classes and a large variety of functional groups in both parts of Europe reflecting a diverse range of lake types across Europe.

There were slightly more diatom, chrysophyte and chlorophyte taxa and slightly less cyanobacteria taxa among dominants in the north than in the south.

Chrysophytes occurred 3 times and cryptophytes nearly twice more frequently among dominants in the North European lakes than in the rest of Europe whereas cyanobacteria and dinophytes occurred about twice less frequently.

The CCA ranked water colour, alkalinity, and TP as the most influential factors determining the large-scale distribution patterns of lake phytoplankton dominants in Europe suggesting that besides trophic conditions, other hydrochemical variables, have at least as important role in determining phytoplankton community composition in lakes. Water temperature from July to September had only a negligible impact on the distribution of dominants, showing the prevalence of rather homogeneous thermal conditions throughout Europe for this period of year.

Cryptophytes and especially chrysophytes revealed a clear affinity to more coloured and less alkaline waters of Northern Europe. The higher frequency of cyanobacteria and dinophytes as dominants in southern Europe can be explained by the higher trophic state and higher alkalinity of lakes in this region.

Our analysis of the distribution of the dominant species provided strong empirical support for the habitat requirements of some phytoplankton functional groups.

Acknowledgements

This paper is a result of the project WISER (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (contract No. 226273), www.wiser.eu. The results are based on a database collated as part of the WISER Project. We would like to thank Birger Skjelbred, Jannicke Moe and Bernard Dudley who supported data management and extraction and especially thank all the data providers. These include (figures in brackets give the number of WBs):

Mediterranean GIG (Data manager: Caridad de Hoyos, CEDEX-CEH)

Spain: Ministerio de Agricultura, Alimentación y Medio Ambiente (122), Centro de Estudios Hidrográficos (CEDEX-CEH) (46).

Italy: Università degli Studi di Sassari. Dipartimento di Scienze Botaniche, Ecologiche e Geologiche (DiSBEG) (18).

Portugal: Instituto da Água, I.P. (INAG) (18).

Romania: Ministeriul Meduli şi Pădurilor (MMP) (10).

Cyprus: Ministry of Agriculture, Natural Resources and Environment, Water Development Department (MANRE-WDD) (7).

France: Institut national de recherche en sciences et rechnologies pour l´environnement et l´agriculture (IRSTEA) (6).

Greece: Maria Moustaka, Aristotle University of Thessaloniki.

Central-Baltic GIG (Data manager: Ute Mischke, IGB)

Estonia: Estonian University of Life Sciences (EMU) gathered for the state monitoring programme supported by the Estonian Ministry of Environment.

Latvia: Sandra Poikane, Latvian Environment, Geology and Meteorology Centre.

Lithuania - EPA Lithuania.

Denmark: National Environmental Research Institute, University of Aarhus.

Belgium: Jeroen Van Wichelen, Ghent University (UGENT).

Netherlands: Rijkswaterstaat (RWS).

Germany: Data from German water bodies were kindly provided by the following institutions of the German Federal States: Landesamt für Umwelt, Gesundheit und Verbraucherschutz Brandenburg (LUGV; 127), Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz Mecklenburg-Vorpommern (MLUV, Seenprogramm, 65), Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt (LHW, 5), Landesamt für Landwirtschaft, Umwelt und ländliche Räume Schleswig-Holstein (LLUR, 13), Senatsverwaltung für Gesundheit, Soziales und Verbraucherschutz Berlin (SenGUV, 12), Niedersächsische Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN, Sulingen, 1).

Poland: Data were provided partly by the Institute of Environmental Protection - National Research Institute. The Inspection for Environmental Protection provided data obtained within the framework of state environmental monitoring.

Hungary: Environmental Protection Inspectorate for Trans-Tiszanian Region.

Northern GIG (Data manager: Geoff Phillips, EA)

Finland: Finnish Environment Institute (SYKE).

Sweden: Swedish University of Agricultural Sciences (SLU).

Norway: Norwegian Institute for Water Research (NIVA).

United Kingdom - Scottish Environment Protection Agency (SEPA) and the Environment Agency for England & Wales (EA).

Ireland: Environment Protection Agency (EPA).

References

Alster, A., R. N. Kaplan-Levy, A. Sukenik & T. Zohary, 2010. Morphology and phylogeny of a non-toxic invasive *Cylindrospermopsis raciborskii* from a Mediterranean Lake. Hydrobiologia 639: 115–128.

Arvola., L., P. Eloranta, M. Järvinen, J. Keskitalo & A.-L. Holopainen, 1999. Food webs of humic waters. Phytoplankton. In: Limnology of humic waters. J. Keskitalo & P. Eloranta (Eds.). Blackhyus Publishers, Leiden. 284 pp.

Briand, J. F., C. Leboulanger, J. F. Humbert, C. Bernard, P. Dufour, 2004. *Cylindrospermopsis raciborskii* (cyanobacteria) invasion at mid-latitudes: selection, wide physiological tolerance, or global warming? Journal of Phycology 40: 231–238.

Bouvy, M., N. Ba, S. Ka, S. Sane, M. Pagano, R. Arfi, 2006. Phytoplankton community structure and species assemblage succession in a shallow tropical lake (Lake Guiers, Senegal). Aquatic Microbial Ecology 45: 147–161.

Canter-Lund, H. & J. W. G. Lund, 1995. Freshwater Algae their microscopic world explored. Biopress Ltd., Bristol. 360 pp.

CEN EN 15204, 2006. Water quality – Guidance standard for the routine analysis of phytoplankton abundance and composition using inverted microscopy (Utermöhl technique).

Carvalho L, A. Solimini, G. Phillips, M. van den Berg, O-P. Pietiläinen, A. Lyche, S. Poikane, U. Mischke, 2008. Chlorophyll reference conditions for European Intercalibration lake types. Aquatic Ecology 42: 203–211.

Carvalho L, A. G. Solimini, G. Phillips, O-P. Pietiläinen, J. Moe, A. C. Cardoso, A. Lyche Solheim, I. Ott, M. Sondergaard, G. Tartari & S. Rekolainen, 2009. Site-specific chlorophyll reference conditions for lakes in Northern and Western Europe. Hydrobiologia 633: 59–66.

Chapman, A. D. & C. L. Schelske, 1997. Recent appearance of *Cylindrospermopsis* (Cyanobacteria) in five hypereutrophic Florida lakes. Journal of Phycology 33: 191–195.

Cronberg, G. & H. Annadotter, 2006. Manual on aquatic cyanobacteria. A photo guide and synopsis of their toxicology. ISSHA & UNESCO, Copenhagen. 106 pp.

Dillard, G. E., 2008. Common Freshwater Algae of the United States. An Illustrated Key to the Genera (Excluding the Diatoms). Gebrüder Borntraeger, Berlin, Stuttgart. 188 pp.

Duarte, C. M., S. Agusti & D. E. Canfield Jr., 1992. Patterns in phytoplankton community structure in Florida lakes. Limnology and Oceanography 37: 155–161.

Directive, 2000. Directive of the European Parliament and of the Council 2000/60/EC establishing a framework for Community action in the field of water policy. Official Journal 2000 L 327/1, European Commission, Brussels.

Dodson, S. I., S. E. Arnott & K. L. Cottingham, 2000. The relationship in lake communities between primary productivity and species richness. Ecology 81: 2662–2679.

Fabbro, L. D. & L. J. Duivenvoorden, 1996. Profile of a bloom of the cyanobacterium *Cylindrospermopsis raciborskii* (Woloszynska) Seenaya and Subba Raju in the Fitzroy River in tropical central Queensland. Marine and Freshwater Research 47: 685–694.

Fastner, J., J. Rucker, A. Stuken, K. Preußel, B. Nixdorf, I. Chorus, A. Köhler, C. Wiedner, 2007. Occurrence of the cyanobacterial toxin cylindrospermopsin in northeast Germany. Environmental Toxicology 22: 26–32.

Figueroa, R. I. & K. Rengefors, 2006. Life cycle and sexuality of the freshwater raphidophyte *Gonyostomum semen* (Raphidophyceae). Journal of Phycology 42: 859–871.

Flöder, S., S. Jaschinski, G. Wells & C.W. Burns, 2010. Dominance and compensatory growth in phytoplankton communities under salinity stress. Journal of Experimental Marine Biology and Ecology 395: 223–231.

Gomes, N., P. R. Hualde, M. Licursi, D. E. Bauer, 2004. Spring phytoplankton of Rio de la Plata: a temperate estuary of South America. Estuarine Coastal & Shelf Science 61: 301–309.

Grime, J. P., 1979. Plant strategies and vegetation processes. John Wiley & Sons, New York, 222 pp.

Grime, J. P., 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. Journal of Ecology 86: 901–910.

Hill, M. O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54: 427–432.

KCS, 2007. MVSP 3.1. Kovach Computing Services, Pentraeth, Wales, U.K.

Kokociński, M., K. Stefaniak, J. Mankiewicz-Boczek, K. Izydorczyk, J. Soininen, 2010. The ecology of the invasive cyanobacterium *Cylindrospermopsis raciborskii* (Nostocales, Cyanophyta) in two hypereutrophic lakes dominated by *Planktothrix agardhii* (Oscillatoriales, Cyanophyta). European Journal of Phycology 45(4): 365–374.

Komárek, J. & K. Anagnostidis, 1999. Süßwasserflora von Mitteleuropa. Cyanoprokaryota Teil 1/Part 1: Chroococcales. Gustav Fischer Verlag Jena, Lübeck, Ulm. 548 S.

Komárek, J. & E. Zapomelova, 2008. Planktic morphospecies of the cyanobacterial genus Anabaena = subg. Dolichospermum – 2. part: straight types. Fottea, Olomouc, 8 (1): 1–14.

Lagos, N., H. Onodera, P. A. Zagatto, D. Andrinolo, S. Azevedo, Y. Oshima, 1999. The first evidence of paralytic shellfish toxins in the freshwater cyanobacterium *Cylindrospermopsis raciborskii,* isolated from Brazil. Toxicon 37: 1359–1373.

Lepistö, L., A.-L. Holopainen & H. Vuoristo, 2004. Type-specific and indicator taxa of phytoplankton as a quality criterion for assessing the ecological status of Finnish boreallakes. Limnologica 34: 236–248.

Margalef, D. R., 1958. Temporal succession and spatial heterogeneity in phytoplankton. In Perspectives in marine biology. Buzzati-Traverso, A. (Eds), University of California Press, Berkeley.

Mischke, U. & B. Nixdorf, 2003. Equilibrium phase conditions in shallow German lakes: How Cyanoprokaryota species establish a steady state phase in late summer. Hydrobiologia 502: 123–132.

Moe, S. J., B. Dudley & R. Ptacnik, 2008. REBECCA databases: experiences from compilation and analyses of monitoring data from 5,000 lakes in 20 European countries. Aquatic Ecology 42: 183–201.

Moisander, P. H., L. A. Cheshire, J. Braddy, E. Calandrino, H. Hoffman, M. F. Piehler, H. W. Paerl, 2012. Facultative diazotrophy increases *Cylindrospermopsis raciborskii* competitiveness under fluctuating nitrogen availability. Federation of European Microbiological Societies. 79: 800–811.

Moustaka-Gouni, M., E. Vardaka, E. Tryfon, 2006. Phytoplankton species succession in a shallow Mediterranean lake (L. Kastoria, Greece): steady-state dominance of *Limnothrix* *redekei, Microcystis aeruginosa* and *Cylindrospermopsis raciborskii*. Hydrobiologia 575: 129–140

Moustaka-Gouni, M., K. A. Kormas, E. Vardaka, M. Katsiapi, S. Gkelis, 2009. *Raphidiopsis mediterranea* SKUJA represents non-heterocystous life-cycle stages of *Cylindrospermopsis raciborskii* (WOLOSZYNSKA) SEENAYYA et SUBBA RAJU in Lake Kastoria (Greece), its type locality: evidence by morphological and phylogenetic analysis. Harmful Algae 8: 864–872.

Nicholls, K. H & D. E. Wujek, 2003. Chrysophycean algae. Freshwater Algae of North America. Ecology and Classification. J. D. Wehr & R. G. Sheath (Eds). Academic Press. California, London, Massachusetts. 918 pp.

Nixdorf, B., U. Mischke & J. Rücker, 2003. Phytoplankton assemblages and steady state in deep and shallow eutrophic lakes - an approach to differentiate the habitat properties of Oscillatoriales. Hydrobiologia 502: 111–121.

Nõges, T., 2009. Relationships between morphometry, geographic location and water quality parameters of European lakes. Hydrobiologia 633: 33–43.

Ott, I. & T. Kõiv. 1999. Estonian small lakes: Special features and changes. Estonian Environment Information Centre. Estonian Academy of Sciences. Institute of Zoology and Botany of the Estonian Agricultural University. Tallinn. 128 pp.

Padisák, J., G. Borics, G. Fehér, I. Grigorszky, I. Oldal, A. Schmidt, Z. Zámbóné-Doma, 2003. Dominant species, functional assemblages and frequency of equilibrium phases in late summer phytoplankton assemblages in Hungarian small shallow lakes. Hydrobiologia 502: 157–168

Padisák, J., L.O. Crossetti & L. Naselli-Flores, 2009. Use and misuse in the application of the phytoplankton functional classification: a critical review with updates. Hydrobiologia 621: 1–19.

Phillips, G., G. Morabito, L. Carvalho, A. Lyche Solheim, B. Skjelbred, J. Moe, T. Andersen, U. Mischke, C. de Hoyos, G. Borics, 2010. Deliverable D3.1-1: Report of lake phytoplankton composition metrics, including a common metric approach for use in intercalibration by all GIGs. Deliverable D3.3-1: Report on lake phytoplankton composition metrics. 56 pp.

Pielou, E.C., 1975. Ecological Diversity. Wiley InterScience, New York.

Ptacnik, R., L. Lepistö, E. Willén, P. Brettum, T. Andersen, S. Rekolainen, A. Lyche Solheim & L. Carvalho, 2008. Quantitative responses of lake phytoplankton to eutrophication in Northern Europe. Aquatic Ecology 42: 227–236.

Reynolds, C. S., 1988. Functional morphology and the adaptive strategies of freshwater phytoplankton. In Sandgren C.D. (Eds.), Growth and reproductive strategies of freshwater phytoplankton. Cambridge University Press, Cambridge: 388–433.

Reynolds, C. S., V. Huszar, C. Kruk, L. Naselli-Flores, S. Melo, 2002. Towards a functional classification of the freshwater phytoplankton. Journal of Plankton Research. 24: 417–428.

Rakko, A., R. Laugaste & I. Ott, 2008. Algal blooms in Estonian small lakes. Evangelista, V., L. Barsanti, A. M. Frassanito,P. Vincenzo & P. Gualtieri, (Eds.). Algal Toxins: Nature, Occurrence, Effect and Detection. Springer. Italy. 211–220

Rosen, G., 1981. Phytoplankton indicators and their relations to certain chemical and physical factors. Limnologica 13: 263–290.

Shannon, C. E., 1948. The mathematical theory of communication. In Shannon C. E. & W. Weaver (eds), The mathematical theory of communication. University. Illinois Press, Urbana.

Simpson, E.H., 1949. Measurement of diversity. Nature 163: 688.

Stomp, M., J. Huisman, G.G. Mittelbach, E. Litchman & C. A. Klausmeier, 2011. Large-scale biodiversity patterns in freshwater phytoplankton. Ecology 92: 2096–2107.

Stüken, A., J. Rücker, T. Endrulat, K. Preussel, M. Hemm, B. Nixdorf, U. Karsten, C. Wiedner, 2006. Distribution of three alien cyanobacterial species (Nostocales) in northeast Germany: *Cylindrospermopsis raciborskii, Anabaena bergii* and *Aphanizomenon aphanizomenoides*. Phycologia 45 (6): 696–703.

Shafik, H. M., S. Herodek, M. Presing & L. Voros, 2001. Factors effecting growth and cell composition of cyanoprokaryote *Cylindrospermopsis raciborskii* (Woloszynska) Sheenayya et Subba Raju. Algological Studies 102: 75–93.

Trigal, C., W. Goedkoop & R. K. Johnson, 2011. Changes in phytoplankton, benthic invertebrate and fish assemblages of boreal lakes following invasion by *Gonyostomum semen*. Freshwater Biology 56: 1937–1948.

Valerio, E., P. Pereira, M. L. Saker, S. Franca, R. Tenreiro, 2005. Molecular characterization of *Cylindrospermopsis raciborskii* strains isolated from Portuguese freshwaters. Harmful Algae 4: 1044–1052.

Watson, S. B., E. McCauley & J. Downing, 1997. Patterns in phytoplankton taxonomic composition across temperate lakes of differing nutrient status. Limnology and Oceanography 42: 486–495.

Willén E., S. Hajdu & Y. Pejler, 1990. Summer phytoplankton in 73 nutrient-poor Swedish Lakes. Classification, ordination and choice of long-term monitoring objects. Limnologica 20: 217–227.

Willén, E., 2003. Dominance patterns of planktonic algae in Swedish forest lakes. Hydrobiologia 502: 315–324.

Willén, E., 2007. Växtplankton i sjöar. Bedömningsgrunder. SLU, Institutionen för Miljöanalys, Rapport 5. 33 pp.

Zabelina, M. M., I. A. Kiselev, A. I. Proshkina-Lavrenko, V. C. Sheshukova, 1951.Opredelitel presnovodnych vodoroslei SSSR. Diatomovye vodorosli. (Identification key for freshwater algae of USSR. Diatoms). Moscow. 619 pp.

Fig. 1 Sharing of the dominant taxa in lake phytoplankton in European countries

Fig. 2 Distribution of the dominating lake phytoplankton taxa among algal classes (two left columns) and their relative frequency of occurrence (two right columns) compared between countries belonging to the Nordic Geographical Intercalibration Group (N-GIG) and the rest of Europe

Fig. 3 Biplot of the Canonical Correspondence Analysis (CCA) results on factors determining the distribution of dominant phytoplankton taxa in lakes of Europe. a – the large picture, b – the central part magnified. The arrows in the biplot representing the environmental variables indicate the direction of maximum change of that variable across the diagram and the length of the arrow is proportional to the rate of change. Each point representing a dominant species lies at the centroid of the samples in which it was found

Fig. 4 CCA biplot showing the factors determining the distribution of Chrysophytes in lakes of Europe

Fig. 5 CCA biplot showing the factors determining the distribution of dominant phytoplankton taxa in lakes of Europe plotted by different functional groups

Table 1.Occurrence of the most dominant species and information of functional groups (Reynolds et al., 2002; Padisák et al., 2009) in European countries

Thackeray et al. (2012). Quantifying uncertainties in biologically-based water quality assessment: a pan-European analysis of phytoplankton community metrics

Stephen J. Thackeray[1], Peeter Nõges[2], Michael Dunbar[3], Bernard J. Dudley[4], Birger Skjelbred[5], Giuseppe Morabito[6], Laurence Carvalho[4], Geoff Phillips[7], Ute Mischke[8], Jordi Catalan[9], Caridad de Hoyos[10], Christophe Laplace[11], Martina Austoni[6], Tomasa Virdis[12], Kairi Maileht[2], Agnieszka Pasztaleniec[13], Marko Jarvinen[14] & Anne Lyche Solheim[5].

[1]Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK.
[2]Centre for Limnology, Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, 61117 Rannu, Tartumaa, Estonia.
[3]Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK.
[4]Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK.
[5]Norsk Institutt for Vannforskning, Gaustadalléen 21, NO-0349 Oslo, Norway.
[6]CNR Istituto per lo Studio degli Ecosistemi, Largo V. Tonolli 50, 28922 Verbania Pallanza, Italy.
[7]Environment Agency, Cobham Road, Ipswich, Suffolk, IP3 9JD, UK.
[8]Leibniz Institute of Freshwater Ecology and Inland Fisheries, Justus-von-Liebig-Straße 2, 12489, Berlin, Germany.
[9] Center for Advanced Studies of Blanes (CEAB-CSIC), Accés Cala St. Francesc 14, Blanes 17300, Spain.
[10]Centro de Estudios Hidrográficos del CEDEX, PO Bajo de la Virgen del Puerto, 3 28005, Madrid, Spain.
[11]Research Institute for Agricultural and Environmental Engineering, CEMAGREF, av de Verdun 50, 33612 Cestas-Gazinet, France.
[12]University of Sassari, Piazza d'Armi, 17, 07100 Sassari, Italy.
[13]Institute of Environmental Protection, 01-692 Warszawa, Kolektorska 4, Poland.
[14]Finnish Environment Institute (SYKE), Jyväskyla Office, P.O. Box 35, 40014 Jyväskyla, Finland.

Summary

1) The EU Water Framework Directive (WFD) states that attributes of biological communities should be used to assess the ecological status of fresh- and coastal/transitional waters. For lakes, the phytoplankton is a key biological community to be used for this purpose. It is therefore necessary to develop metrics that describe high-level properties of phytoplankton communities and that are sensitive to environmental pressures, such as nutrient enrichment.

2) Assessment of the utility of such metrics demands a knowledge of the extent to which they can be affected by sampling and sample processing procedures e.g. where samples are collected from and who processes the samples. If metrics vary more with differences in sampling and sample processing within a lake than they do among lakes then they are unlikely to provide a sensitive means of describing differences in the biological impacts of an environmental stressor among lakes. Here we analyse the results of a multi-scale field campaign of 32 European lakes, to resolve the extent to which seven proposed phytoplankton metrics vary among lakes and with sampling/sample processing. We also relate these metrics to different environmental variables, including total phosphorus concentration as an indicator of eutrophication.

3) For all seven metrics, between 65% and 96% of the variance in metric scores was due to variability among lakes, much higher than variability occurring due to sampling/sample processing. Using multi-model inference, there was strong support for relationships between among-lake variation in three of the metrics and differences in total phosphorus concentrations. Three of the metrics were similarly related to mean lake depth. Unexplained among-lake metric variance indicated that metrics were additionally sensitive to unmeasured environmental factors. Differences among sub-samples and analysts accounted for much of the within-lake metric variance, suggesting that sub-sample replication and standardisation of analyst procedures may result in increased precision of ecological assessments based upon these metrics.

4) The residual variance in most metrics, and therefore the uncertainty associated with them, changed as a function of among lake variations in the physical (mean depth) and chemical (total phosphorus concentration) environment, and lake location (altitude).

5) For three of the candidate phytoplankton metrics being considered for the WFD Intercalibration of lake phytoplankton metrics: chlorophyll *a* concentration, the Phytoplankton Trophic Index (PTI), and cyanobacterial biovolume, > 88% of the variance in metric scores was among-lakes, and, total phosphorus concentration was well supported as a predictor of this among-lake variation. Based upon this study, these proposed metrics may be considered robust for ecological status assessment and suitable for adoption in the WFD Intercalibration process.

Keywords: ecological quality assessment, eutrophication, linear mixed effects models, multi-model inference, Water Framework Directive

Introduction

The Water Framework Directive [WFD; (EC 2000)] has revolutionised the assessment of anthropogenic impacts upon fresh- and coastal-transitional waters of the member states of the European Union. The central tenet of the Directive is that the assessment of human impacts on the surface water environment, rather than being based solely upon chemical parameters, should be based upon the attributes of key communities (Biological Quality Elements, BQEs) that are sensitive to environmental pressures such as eutrophication and physical habitat modification.

For lakes, the phytoplankton has been identified as a key BQE to be used in ecological status assessment and is already widely used as an important water quality indicator because of rapid replication rates (ensuring rapid responses to environmental stressors), direct sensitivity to physical and chemical environmental factors, and high diversity with species and/or functional types showing markedly variable responses to changes in the surrounding environment (Murphy et al. 2002, Reynolds 2006). Furthermore, sampling of these communities is simple and inexpensive, with minimal impacts on co-existing biota. As a result of these features, phytoplankton was included in the WFD monitoring scheme as a relevant quality element for all surface water categories. As parameters to be studied, the WFD prescribes phytoplankton abundance, composition, and the frequency and intensity of blooms. While phytoplankton community composition and diversity are regulated by a complex interplay of intrinsic and extrinsic drivers such as climate, resource availability, patterns of competition and predation, and dispersal (Reynolds 2006) they may also act as sensitive indicators of environmental pressures such as eutrophication as a result of increased nutrient loading (Kuemmerlin 1998, Padisák and Reynolds 1998). Phytoplankton abundance, composition and the frequency/intensity of blooms are all considered to undergo changes along this pressure gradient (Carvalho et al. 2006). The extent of these changes can be “translated” into WFD normative definitions. To this end, the WFD requires quantitative high-level indicators, or metrics, of these complex systems which can be used to monitor the status of freshwater communities in the face of anthropogenic pressures, and identify improvements to ecological status as a result of management interventions. As part of the EU project WISER (http://www.wiser.eu/) a number of existing, or newly developed, metrics have been considered for this purpose (Mischke et al. 2010, Phillips et al. 2010).

However, there is an urgent need to assess the likely uncertainty in ecological status assessments when using such metrics (Hering et al. 2010). Phytoplankton communities show marked spatial heterogeneity within lakes, over a range of spatial scales, as a result of patterns in lake circulation and mixing, and spatial gradients in flushing, grazing and nutrient availability (Pinel-Alloul and Ghadouani 2007). In addition, variation in phytoplankton metrics may occur due to differences in the analysts processing samples and sub-sampling procedures (Vuorio et al. 2007). Therefore, it is highly likely that the choice of sampling location within a lake and sample processing will affect the values of metrics based upon phytoplankton community data. Where metric values fall close to ecological status class boundaries, then these variations may fundamentally influence the overall assessment of a waterbody (Clarke et al. 2006b). This has led to suggestions that results of ecological status classification should be given in terms of probabilities (Hering et al. 2010). Analyses of riverine macroinvertebrate community metrics have shown that the level of metric variability due to sampling may itself change with the ecological quality of a site (Clarke et al. 2006a), along the environmental pressure gradient that separates minimally and extensively impacted sites. If the candidate phytoplankton metrics are to be used to distinguish between lakes of differing ecological quality, then among-lake variations in metric scores must be maximised and variation due to sampling/sample-processing minimised in order to give the best chance for the former to be related to differences the intensity of key ecological pressures acting upon those lakes. It is also important to know whether these metrics become inherently more or less variable (uncertain) along this pressure gradient.

Until now, there has not been a formal assessment of the multiple sources of uncertainty that are inherent in phytoplankton metrics. The statistical tools to make this assessment exist (Carvalho et al. 2006, Clarke and Hering 2006) but there has been a need for new data, collected according to a sampling design that allows distinction of different and independent sources of variability in metric scores. Knowledge of the relative importance of different sources of metric variability will guide the design of sampling campaigns aimed at ecological quality assessment. For example if a large component to the total variance in a metric is associated with sub-sampling of field samples, then the precision of assessments based upon this metric could be improved by analysing a larger number of sub-samples to derive a more representative average metric score for the lake. Herein, we present the results of a novel analysis of seven established phytoplankton community metrics based on a pan-European field sampling campaign of 32 lakes. Rigorous standardisation of sampling and sample processing procedures, along with a hierarchical sampling design targeted at uncertainty estimation, allow an entirely consistent analysis of sources of phytoplankton metric variation within and among European lakes. Specific objectives are to address the following questions; do candidate phytoplankton community metrics:

Q1: show greater variability among lakes than within lakes or as a result of differences in sample processing?

Q2: differ significantly along a gradient in lake nutrient status, once accounting for within-lake and sample-processing variation?

Q3: show systematic changes in their level of variability along gradients in physical, chemical and geographic attributes of lakes?

Materials and methods

Field survey

The present analysis is based upon water samples collected from 32 lakes in eleven European countries during the spring and summer of 2009 (Table 1). These collectively represented lake types found within Member States and Norway comprising the Alpine, Northern, Central/Baltic and Mediterranean Geographical Intercalibration Groups [GIGs(WISE 2008)]. All lakes were less than 10 km^2 in surface area, but varied widely in mean depth (3.5 - 34 m) and altitude (15 – 970 m a.s.l.). The lakes also differed markedly in productivity/trophic status, with wide variation in alkalinity (0.06 – 4.40 meq L^{-1}) and total phosphorus concentration (4 - 151 mg m^{-3}) at the time of sampling.

Lakes were all sampled according to the same standardised protocol. The sampling design allowed the total variability in phytoplankton community structure, as indicated by a range of metrics, to be decomposed into a series of independent variance components, each indicating a potential source of uncertainty. The sampling design was as follows (Fig. 1):

(i) Within each lake, water samples were collected at three stations: above the deepest point of the open water zone, and at points representing the mean depth of the lake and a depth intermediate to the mean and maximum depths. This allowed quantification of within-lake spatial heterogeneity in phytoplankton community composition and metric scores, at the basin scale.

(ii) Two water samples were collected at each of the three stations. This allowed quantification of errors associated with repeated sampling at a specific location, as a result of smaller-scale heterogeneity in the phytoplankton community.

(iii) Each sample was sub-sampled in order to quantify variations in phytoplankton metric scores due to sub-sampling errors and differences in the analyst identifying and enumerating phytoplankton in the sub-samples. For analyses of phytoplankton composition, three sub-samples were collected from the first sample. Two of these were processed by the same analyst (revealing sub-sampling error), while the third was processed by a different analyst (to evaluate variability in metric scores due to differences in the approach used by different analysts). From the second sample, only one sub-sample was collected, to allow comparison with metric scores derived from the first sample. Prior to microscopic examination an aliquot (sub-sub-sample) of each sub-sample was collected and put into a sedimentation chamber. Any variation associated with this sub-sub sampling is of course confounded with sub-sample variation in what follows, as no replication is available at this level of the hierarchy. For chlorophyll *a* (Chl-*a*) analysis, which followed a rigorously standardised spectrophotometric protocol, the effect of the analyst was not addressed and only two sub-samples were taken from the first sample to evaluate the sub-sampling error.

This design allowed the identification of elements of field sampling campaigns that, through greater replication or standardisation, could be modified in order to improve the precision of ecological status assessments. For example, would the precision of such assessments be improved if we collected more samples, samples from more stations throughout the lake, processed more sub-samples or standardised taxonomic skills among analysts?

At each station, water samples were collected using an integrated tube sampler. If a lake was thermally stratified samples were taken from the euphotic layer (estimated as 2.5 x Secchi depth). When the water column was mixed samples were collected from throughout the whole water column, down to 0.5m above the sediment surface. Sub-samples were collected from each sample after thorough mixing, If immediate extraction of Chl-*a* samples was not possible, they were stored in a refrigerator or ice box for as short a time as possible. Samples for microscopic analysis were preserved using a solution of Lugol's iodine (final concentration approximately 0.5% by volume) and stored in the dark.

A further separate water sample was collected at the deepest point of each lake and analysed for alkalinity and concentrations of total phosphorus (TP). TP was measured following sulphuric acid-potassium persulphate digestion of unfiltered samples, according to Murphy & Reilly (1962). For some lakes multiple determinations of each variable were made and these were averaged prior to statistical analyses. Whilst data on total phosphorus concentrations were available for all lakes, alkalinity values were missing for some lakes and so representative values were necessarily derived from data collected under a parallel hierarchic macrophyte survey (Dudley et al. 2010). Secchi depth was also recorded at the deepest point of each lake.

In the following analyses TP concentrations were used to indicate where the sampled lakes fell on a gradient of nutrient enrichment. Latitude, longitude and altitude of each lake were also included, as proxies for broad climatic gradients that might impact upon phytoplankton communities via effects on lake physical processes. Alkalinity and mean lake depth were included in the study as they are the primary determinants of the fundamental lake “types” described the WFD. Different combinations of high-low alkalinity and mean depth have been used to categorise these lake “types”, capturing the fact that lakes show natural variability in their phytoplankton communities, due to their catchment setting and morphometry, irrespective of differences in nutrient enrichment (Pinel-Alloul et al. 1990).

Sample processing for Chl-a analysis

A fixed volume of water, dependent on the amount and type of seston present in each lake, was filtered through 47-mm GF/F filters and the filter was placed into 10 ml of 96% ethanol for pigment extraction at 4 °C for 24 hours. Following extraction, samples were shaken and centrifuged for 5 minutes to remove filter and cell debris.

The extract was decanted from the centrifuge tube to 1-cm cuvette. Readings of spectrophotometric absorption (A) against 90% ethanol were taken in the range of the Chl-*a* absorption maximum, 662-665 nm (A_{peak}), and at 750 nm as a correction for light scattering (A_{750}). The sample was then acidified with 2 drops of 1.2 M HCl and readings at 662-665 nm and 750 nm were repeated resulting, correspondingly, in acA_{peak} and acA_{750}. To correct the absorbance for scattering, the A_{750} was subtracted from readings in non-acidified samples and acA_{750} from the acidified samples, i.e.

$E_{peak} = A_{peak} – A_{750}$ and
$acE_{peak} = acA_{peak} – acA_{750}$

Finally the concentration of Chl-*a* was calculated according to the equation given in ISO10260 (1992):

Chl-*a* [mg/m3] = $29.6 * (E_{peak} - acE_{peak}) * a/(L * V)$

Where a = final extraction volume (ml)
V = volume of water filtered (L)
L = length of the light path through the cuvette (cm)

Sample processing for microscopic examination of phytoplankton

Microscopic examination of phytoplankton followed the same standardised protocol across Member States, and was based upon procedures outlined in CEN 15204 (2006), National Rivers Authority (1995) and Brierly *et al.* (2007). Briefly, samples were examined in sedimentation chambers with an inverted microscope, according to the Utermöhl technique (Utermöhl 1958). For each sample, a low magnification (40x or 100x) whole chamber count, two intermediate magnification (200x or 250x) transect counts and 50-100 field of view counts at high magnification (400x or greater) were completed. Phytoplankton taxa were identified to the highest possible level. Counts of each taxon were converted to biovolumes by measuring cell/colony dimensions and approximating each taxon to a simple geometric shape (Brierly et al. 2007). Phytoplankton cells were measured using eye-piece graticules, after calibration with a stage micrometer. All subsequent phytoplankton metric calculations were based upon the biovolume data.

Phytoplankton metrics

Seven candidate phytoplankton metrics are considered herein, a brief description of which is given below. Full details on each metric are provided in Phillips et al. (2010) and Mischke et al (2010). These metrics have been categorised according to whether they relate to phytoplankton abundance or composition, or to features of blooms.

1. Chl-*a* concentration (Abundance metric, in mg m^{-3}) is a measure of phytoplankton abundance, commonly used to represent the ecological status of a lake with respect to eutrophication pressures.
2. Phytoplankton Trophic Index (PTI, Composition metric). This has been developed, using an independent data set, from the "trophic scores" of phytoplankton taxa along a eutrophication gradient. After a Canonical Correspondence Analysis (CCA) constrained by total phosphorus, taxa optima on the first ordination axis were derived indicating the TP concentration for the mean occurrence of each taxon. For each sub-sample, PTI was calculated as the weighted average of these taxa optima, where the weighing factor is the proportional biovolume of each taxon. The PTI increases with increasing lake trophic state.
3. Size Phytoplankton Index (SPI, Composition metric). The phytoplankton taxa within a sub-sample are grouped into a series of size categories, each one encompassing a doubling of cell biovolume e.g. $\leq 0.5 \mu m^3$, 0.5-1.0 μm^3, 1.0-2.0 μm^3, 2.0-4.0 μm^3 etc (Kamenir and Morabito 2009). The SPI is then calculated as a function of the size categories and "trophic scores"/"indicator values" for those categories (Phillips et al. 2010). Trophic scores indicate the position of a size class along the trophic spectrum and indicator values estimate the "power" of each size class as a biotic indicator. The SPI tends to increase with increasing lake trophic state, due to a shift towards increased dominance of larger, rather than smaller, phytoplankton (Phillips et al. 2010).
4. Morpho-Functional Group Index (MFGI, Composition metric). The phytoplankton taxa within a sub-sample are grouped into a series of categories ("Morpho-Functional Groups") based upon their morphological attributes e.g. presence/absence of flagella, colonial or unicellular, large or small size (Salmaso and Padisak 2007). The MFGI is then calculated as a function of the Morpho-Functional Groups and the "trophic scores"/"indicator values" for those groups (Phillips et al. 2010). The MFGI tends to increase with increasing lake trophic state, due to an increase in the dominance of colonial cyanobacteria, large diatoms/chlorophytes/conjugatophytes, and unicellular/colonial chlorococcales (Phillips et al. 2010).
5. Functional Traits Index (FTI, Composition metric). This is the arithmetic mean of the SPI and MFGI, and thus combines information on both the size spectrum and morpho-functional traits of the phytoplankton community. Phillips et al (2010) recommend the use of the FTI for water quality assessment.
6. Evenness metric (Bloom metric). This is Pielou's evenness index, which expresses the ratio between the Shannon diversity of a sub-sample and the maximum possible value of the Shannon diversity index (Pielou 1969, 1975). Evenness has been shown to decline under bloom conditions in more productive lakes, due to an increase in the dominance of a small number of tolerant species with high growth rates (Mischke et al. 2010).
7. Cyanobacterial abundance (Bloom metric). This is the total cyanobacterial biovolume (mm^{-3} L^{-1}) within a sub-sample, and is expected to increase with increasing lake trophic status (Mischke et al. 2010).

Statistical modelling

Q1: Do metrics show greater variability among lakes than within lakes or as a result of differences in sample processing?

These analyses aimed to resolve whether metrics had the potential to be sensitive to variations in the intensity of environmental pressures acting at the lake level (among-lake vs. within-lake metric variance), and to identify aspects of sampling campaigns that might be modified to improve the precision of ecological status assessments (comparison of components of within-lake metric variance). Linear mixed effects models (LME) were used to analyse metric scores (*m*) based upon samples collected during the hierarchic sampling campaign. A nested random effects structure was used to emulate the hierarchical nature of the sampling campaign and model non-independence between metric scores based upon data from the same sub-sample, sampling station, lake, etc. In this structure, lake was nested within country, sampling station within lake and sample within station. Sub-sample was modelled as the error ("unexplained") variability and analyst was included (except for analyses of Chl-*a* concentration) as a crossed random effect, representing the fact that it does not naturally nest within the sample hierarchy. More formally, the model structure can be denoted:

$$m_{\mathrm{austwc}} = \beta_0 + v_{\mathrm{ustlc}} + v_{\mathrm{stlc}} + v_{\mathrm{tlc}} + v_{\mathrm{lc}} + v_{\mathrm{c}} + v_{\mathrm{a}} + e_{\mathrm{austlc}} \qquad \text{eqn. 1}$$

where m_{austlc} is the value of the metric m for analyst a, for sub-sample u, in sample s, in station t, in lake l, in country c. Thus, m_{austlc} is the sum of a series of components that each contribute to metric variation. β_0 is the intercept common to all data points. The components of metric variation are modelled as independent, Normally distributed, variance components for analyst, sub-sample, sample, station, lake and country variance:

$v_{\mathrm{ustlc}} \sim \mathrm{N}(0, \sigma^2_{\mathrm{u}})$

$v_{\mathrm{stlc}} \sim \mathrm{N}(0, \sigma^2_{\mathrm{s}})$

$v_{\mathrm{tlc}} \sim \mathrm{N}(0, \sigma^2_{\mathrm{t}})$

$v_{\mathrm{lc}} \sim \mathrm{N}(0, \sigma^2_{\mathrm{l}})$

$v_{\mathrm{c}} \sim \mathrm{N}(0, \sigma^2_{\mathrm{c}})$

$v_{\mathrm{a}} \sim \mathrm{N}(0, \sigma^2_{\mathrm{a}})$

$e_{\mathrm{austlc}} \sim \mathrm{N}(0, \sigma^2)$

In order to determine the levels of the sampling hierarchy at which metric values showed the greatest variability (e.g. among vs. within lakes), a null model (i.e. including only the random effects and no potential explanatory variables as fixed effects) was fitted for each metric. The estimated variance parameters for each level in the random effects hierarchy (σ^2_{u}, σ^2_{s}, σ^2_{t}, σ^2_{l}, σ^2_{c}, σ^2_{a}, σ^2) were extracted in order to compare metric variation among lakes with that among samples and sub-samples/analysts. The proportion of the total metric variation occurring at each level in the sampling hierarchy was calculated from these parameter estimates. During this stage of the analysis, restricted/residual maximum likelihood (REML) estimation was used during model fitting in order to give unbiased estimates of the random effects.

Q2: Do metrics differ significantly along a gradient in lake nutrient status, once accounting for within-lake and sample-processing variation?

In order to investigate the extent to which phytoplankton metrics were sensitive to variations in measured morphometric, chemical and geographical features, models were then re-run to include the measured environmental variables (TP, alkalinity, mean lake depth, latitude, longitude and altitude) as fixed effects. Secchi depth was omitted since the direction of causality between this variable and the phytoplankton community is equivocal. In order to explicitly take account of uncertainties in model selection, arising since both model formulation and parameters are estimated from the sample data, we used multi-model inference (Burnham and Anderson 2002). For each metric, a "global" model was constructed containing the same random effects structure given above (equation 1) and all the predictor variables, so that lake-level variance could be modelled as a function of these predictors:
$v_{lc} \sim N(V_l, \sigma^2_l)$

and $V_l = \text{alkalinity}_l + \text{latitude}_l + \text{longitude}_l + \text{altitude}_l + \text{meandepth}_l + TP_l$ eqn. 2

Models were then run including all possible subsets of these variables, and ranked by the Akaike Information Criterion (AIC). A subset of top models, receiving progressively lower levels of support from the data, was determined by finding the model with the most optimal combination of environmental predictor variables (i.e. lowest AIC value) and other candidate models with AIC values differing from this "top" model by ≤ 4 (Burnham and Anderson 2002, Zuur et al. 2009). Model-averaged parameters (with 95% confidence intervals) were calculated using the parameter estimates in models within this top model subset. Maximum likelihood (ML) estimation was used when fitting models with different combinations of predictor variables.

To estimate the amount of the total among-lake variation in metric scores that could be "explained" by the selected environmental variables we compared the residual among-lake metric variance estimated by the model with the most optimal combination of environmental predictors (i.e. lowest AIC value), with that estimated in the corresponding null model (i.e. with no environmental predictors) thus:

$Prop_e = 1-(\sigma^2_{l,\ fitted}/\ \sigma^2_{l,\ null})$ eqn. 3

Where, for each metric, $Prop_e$ is the proportion of the among-lake variance explained by the fitted environmental predictors and $\sigma^2_{l,fitted}$ and $\sigma^2_{l,null}$ are the residual among-lake variance components estimated, respectively, in the most optimal fitted model and the null model. $\sigma^2_{l,\ fitted}$ therefore represents the among lake variation in a metric that cannot be explained by the predictor variables in the top fitted model, while $\sigma^2_{l,null}$ represents the total among-lake variation in that metric. This approach is conceptually similar to that employed by Clarke *et al*. (2006b) to compare variance components of invertebrate metric scores gathered from hierarchic sampling designs. Since $\sigma^2_{l,\ fitted}$ and $\sigma^2_{l,null}$ are themselves estimated parameters, and therefore each have a level of uncertainty associated with them, $Prop_e$ must also be considered an estimate with a level of uncertainty. Herein, we do not calculate the uncertainty associated with the estimate of $Prop_e$ and merely use the values as broadly indicative of the explanatory power of the selected predictor variables.

During the model fitting exercise, it was necessary to simplify the random effects structure to retain only crossed effects of "Lake" and "Analyst". Preliminary analyses revealed that the inclusion of the full random effects hierarchy when comparing models with different fixed effect structures resulted in convergence errors, due to high levels of model complexity. Furthermore, fitting of null models (see results) demonstrated that the omitted random effects consistently accounted for little of the total metric variance.

Q3: do metrics show systematic changes in their level of variability along gradients in physical, chemical and geographic attributes of lakes?

As a final step in the analysis, we examined whether metric scores became more or less variable as a function of among-lake changes in predictor variables, such as TP concentration or mean depth. If metric variability is not constant across lakes with different environmental attributes, then this could mean that sampling campaign design (in terms of sample replication, level of standardisation) might also need to vary among lakes. This was done by adding additional variance structures to previously fitted models that allowed for changes in residual metric variability as a function of the measured environmental predictors. For each metric, we worked with the model with the most optimal combination of environmental predictor variables (lowest AIC) and added these extra variance structures based upon each of the predictors within this top model. These structures took the form (Zuur et al. 2009):

$$\mathrm{var}(\varepsilon) = \sigma^2 e^{2\delta x} \qquad \text{eqn. 4}$$

so that the residual variance [$\mathrm{var}(\varepsilon)$] was allowed to vary as an exponential function of explanatory variable x and the estimated parameter δ. For each metric, we compared the top fitted model with none of these additional variance structures, with models including structures that allowed for residual "spreading" with respect to each of the explanatory variables present in the top model. So, for example, if the top model for a particular metric included predictors x_1 and x_2, we compared models i) without structures to capture spreading of residual metric variation, ii) with residual spreading as a function of x_1, iii) with residual spreading as a function of x_2 and, iv) with residual spreading as a function of x_1 and x_2. The most optimal solution was found by comparing the AIC values of each of these models, after fitting using REML estimation.

All analyses were conducted using the *base*, *gplots*, *lme4*, *MuMIn* and *nlme* packages of R version 2.13.1 (Pinheiro et al. 2010, Warnes 2010, Barton 2011, Bates et al. 2011, R Development Core Team 2011) and the Variance Estimation and Precision (VEPAC) package of STATISTICA 8.0 (StatSoft. Inc. 1984-2007).

Results

Sources of metric variability

Exploratory analyses of the metrics data revealed that Chl-*a* and total cyanobacterial biovolume were positively skewed and so, prior to statistical modelling, we $\log_{10}(x+0.1)$ transformed these metrics in order to reduce the potential influence of the minority of relatively high values in the dataset. The overall level of variability in each metric, across all sub-samples, was compared using the metric standard deviation after re-scaling all metrics to zero mean (i.e. subtracting the mean metric value from each metric series). This suggested that total cyanobacterial biovolume, Chl-*a* and PTI were most variable overall (Table 2). Results from null models of all seven metrics (Table 2) suggest that the majority of metric variance occurred between lakes. The Country (σ^2_c) and Lake (σ^2_l) random effects accounted for between 65% and 96% of the total metric variance. It is noteworthy that the Analyst (σ^2_a) and Error (sub-sample level, σ^2) variance components were the major contributors to the within-lake component, such that analyst differences and sub-sampling errors accounted for more variation in the phytoplankton community metrics than differences between samples and stations within a lake.

Relationships between metrics and lake characteristics

The seven metrics varied widely in their apparent relationship to total phosphorus concentration; indicating the primary among-lake pressure gradient of nutrient enrichment (Fig. 2). Visual inspection of the data suggested that metric-phosphorus relationships were strongest for the abundance metric Chl-*a*, PTI composition metric and total cyanobacterial biovolume bloom metric. This was confirmed by the structure of the most optimal models for these metrics, which included effects of total phosphorus concentration and mean lake depth (Table 3). Delta AIC values for these models, all ≥13.5, indicated a significant improvement in model fit compared to (null) models with no predictors. All three of these metrics increased in value in lakes with higher phosphorus concentrations and in shallow lakes. Top models for the three remaining composition metrics (MFGI, SPI and FTI) suggested that all three metrics increased in shallow lakes and in lakes at higher altitudes. While ΔAIC values ≥9 indicated that top models were considerably better supported than null models for MFGI and FTI, this was not the case for SPI (ΔAIC =2). Similarly the top model for the evenness, suggestive of a reduction in this bloom metric with increasing phosphorus concentration and at low alkalinity, represented only a modest improvement on a model with no fitted predictor variables (ΔAIC = 2.3). The majority of the among-lake variance in Chl-*a* concentration was accounted for by the fitted predictors in the top model (Table 3, Fig. 3). For total cyanobacteria and the PTI metric, the amount of among-lake variance "explained" by the fitted predictors in the top model was less, at 43-47%, while for the remaining metrics <40% of the among lake metric variance was accounted for in the fitted models.

However, relatively low Akaike weights for the top models for all metrics (0.06-0.19, Table 3) suggested that the top models did not receive overwhelming support within each model set and that, for each metric, other candidate models collectively received support from the data. Using a multi-model inference approach to calculate model averaged parameters for the relationships between each metric and the selected environment predictors confirmed strong support for an increase in Chl-*a* concentration, PTI and total cyanobacterial biovolume at high phosphorus concentrations (Figs. 4-6). Across many of the metrics there was a support for an effect of mean lake depth on metric scores. With the exception of evenness, all metrics decreased with an increase in mean lake depth i.e. a negative slope parameter for their relationship (Figs. 4-6). For MFGI, FTI and total cyanobacterial biovolume there was strong support for this effect, while for the remaining metrics support for this effect was relatively weaker. With the exception of Chl-*a* concentration there was also consistent, though weak, support for an effect of altitude on metric scores. Tables summarising the model sets used to derive these averaged parameters for each metric can be found in the Supplementary Information.

Changes in metric variability as a function of among-lake variations in physical, chemical and geographical attributes

For all but one of the metrics (FTI) the fit of the most optimal statistical model (from Table 3) was improved by allowing residual metric values to vary as a function of certain explanatory variables (phosphorus concentration, lake depth, Table 4). In the case of SPI and MFGI the difference in AIC between models including and excluding these structures (5.7 and 2.7 respectively) was much lower than for Chl-*a* concentration, PTI, evenness and total cyanobacteria biovolume (20.9 - 44.8). While residual Chl-*a* concentrations and evenness appeared to become more variable at lower phosphorus concentration (negative δ estimates), cyanobacterial biovolume showed the reverse pattern; with residuals being more variable at higher phosphorus concentrations (positive δ estimate). Residual Chl-*a* concentrations also became more variable at greater mean lake depths (positive δ estimate), while residual PTI and MFGI became less variable in these deeper lakes (negative δ estimates). Both residual SPI and PTI became more variable in higher altitude lakes (positive δ estimates). The model selection process, using multi-model inference to find the most well supported predictors of among-lake variations in each of these metrics, was repeated after including these additional variance structures, although the final parameter estimates for the fixed effects were affected minimally (results not shown).

Discussion

Herein, we analysed the results of a unique hierarchic field sampling campaign to resolve sources of variation in seven phytoplankton metrics that have been proposed as measures of lake ecological status, to serve the requirements of the EU Water Framework Directive (WFD). A specific focus was the relative magnitude of among-lake variation in each metric, compared to within-lake/sample-processing variation. Further, we assessed the evidence for relationships between these metrics and the primary environmental pressure gradient recognized by the WFD; eutrophication (as indicated by the proxy of total phosphorus concentration).

Comparison of sources of variation in metric scores showed that among-lake variation was by far the dominant component of variability for all seven metrics. This suggested that, all other things being equal, the capability of the metrics to respond to pressures acting at the lake level should not be limited by sampling variation arising from within-lake spatial variation. Differences in locations around a lake, or sampling and analytical variability, only accounted for a relatively small proportion of the variance in metric scores. These results are especially true for the four candidate phytoplankton metrics being considered for Intercalibration: chlorophyll, PTI, MFGI and cyanobacterial blooms, for which 86% or more of the variance in metric scores occurred at the among-lake level of the sampling hierarchy. Between-analyst and between sub-sample variation accounted for most of the remaining, within-lake, variation. Little variation was attributable to differences between lake stations and repeated sampling from each station. This was despite the fact that lake stations were treated as “random” in the modelling approach even though they were selected: which should lead to an over-estimate of the station-to-station variability. Lake stations were selected to represent water columns of mean depth or greater in the present study, and it is plausible that a greater station level effect might have been observed if stations had been selected in shallower waters or from outflow or edge samples. Processes in inshore regions of lakes, such as flushing by influent waters (Mackay et al. 2011), enhanced zooplankton grazing facilitated by structurally complex macrophyte refugia (Schriver et al. 1995) or chemical interactions with macrophytes (Wium-Andersen et al. 1982, Jasser 1995) may generate differences in phytoplankton communities between these areas and the deeper, open-water, zone. Furthermore, if sampling stations were distributed among the multiple interconnected basins of some lakes, it is conceivable that more station-level metric variation would be observed.

Though within-lake metric variance was relatively low compared to among-lake variance, the relative magnitude of the components of the former indicates potential areas for the refinement of field sampling campaigns, which could improve the precision of ecological assessments of lakes. Increasing the number of open water sampling stations visited, or the number of samples collected at each station, would do little to improve the precision of ecological assessments based upon these phytoplankton metrics. The representativeness of ecological assessments based upon the metrics, with respect to the impact of lake level pressures, could instead be improved by processing greater numbers of replicate sub-samples from each sample and standardising either *i)* analyst identity for samples from different lakes, or *ii)* taxonomic skills and laboratory procedures among different analysts. In fact, all analysts had attended workshops that aimed to standardise sample processing techniques and algal identification/enumeration. Furthermore, counters followed standard procedures based upon CEN 15204 (2006), National Rivers Authority (1995) and Brierley *et al.* (2007). It may therefore be the case that increasing sub-sample replication is the most feasible means by which to diminish the uncertainty associated with ecological assessments based upon the phytoplankton metrics. Nevertheless, the results of this study indicate that rigorous standardisation of sample mixing and sedimentation protocols, as well as of taxonomic procedures, can help minimise sampling and analytical variability and help make more meaningful comparisons of ecological status among different lakes.

We should also note that, in the current sampling design, the effects of analyst and sub-sampling variation were crossed such that it was not possible to compare results derived from different analysts counting exactly the same fields of view from the same sub-sample, or the same analyst counting different fields of view from the same sub-sample. Furthermore, the sub-samples were actually sub- sub-sampled prior to microscopic examination; another source of potential metric variability that was unquantifiable in this study. It is, therefore, difficult to truly isolate the effect of analyst variation upon metric scores in this study. Future studies targeting sources of variation arising from sampling processing and analyst variation alone would allow more accurate assessment of the extent to which metrics are influenced by these factors.

Total phosphorus concentration, as a proxy for among-lake variation in trophic status, featured in the most optimal statistical models for four of the seven metrics (Chl-*a* concentration, PTI, evenness and total cyanobacterial biovolume). However support for these "top" models was not overwhelming, as indicated by low Akaike weights; suggestive of high levels of model selection uncertainty. Taking a multi-model inference approach, there was strong support for a response of metric scores to phosphorus concentrations for three of the seven metrics: Chl-*a* concentration, PTI and total cyanobacterial biovolume. This would suggest that these proposed metrics are indeed responsive to the eutrophication pressure gradient apparent across the lakes sampled. These relationships suggested a general increase in Chl-*a* concentration and cyanobacterial abundance with increased phosphorus availability. The finding that Chl-*a* concentration increases with lake phosphorus concentration is consistent with the idea that the availability of this nutrient determines the supportive capacity of a lake system for phytoplankton biomass (Reynolds 2006); a relationship embodied in the results of previous empirical (Dillon and Rigler 1974, Schindler 1978, Phillips et al. 2008, Sondergaard et al. 2011), and process-based modelling studies (Elliott et al. 2006). Indeed, among lake variations in total phosphorus concentration have been found to be more powerful predictors of phytoplankton biomass than similar variations in total nitrogen concentrations (Brown et al. 2000, Phillips et al. 2008, Sondergaard et al. 2011), though this difference may be dependent on the relative availability of these two nutrients (McCauley et al. 1989, Brown et al. 2000, Phillips et al. 2008). The observation of increased cyanobacterial biomass at higher phosphorus concentrations is similarly consistent with the findings of previous studies (Smith 1985, Watson et al. 1997, Elliott et al. 2006). PTI scores were also higher in lakes with higher phosphorus concentrations, as shown by Phillips *et al* (2010)....which taxa are increasing with trophic state to cause this?

Comparison of results across metrics also revealed consistent support for an effect of mean lake depth, particularly for FTI, MFGI and total cyanobacterial biovolume (though there was also weaker support for this effect for PTI, SPI and Chl-*a* concentration). Mean lake depth acts as a surrogate for a variety of physical and chemical attributes, such as maximum depth, the likelihood of thermal stratification, flushing rate and underwater light availability (Kalff 2002). Furthermore, inverse relationships between among-lake variations in lake depth and Chl-*a* concentrations/cyanobacterial abundance have been noted in a number of previous studies (Pridmore et al. 1985, Smith 1985, Smith et al. 1987, Phillips et al. 2008). The fact that lake depth covaries with so many other physical and chemical determinants of phytoplankton production, renders hypothesising the mechanism behind the observed relationships difficult. That depth and total phosphorus concentration co-occur as predictors in the top models for Chl-*a* concentration and total cyanobacterial biovolume would suggest that depth offers “unique” explanatory power for these phytoplankton metrics compared to phosphorus availability on the day of sampling. The significance of this depth effect may lie in the fact that shallow lakes may be subject to frequent episodic wave mixing events, that disturb sediments and increase fluxes of nutrients into overlying waters (Hamilton and Mitchell 1988). A single spot measurement of total phosphorus concentration in surface waters, as collected for this study, will not adequately capture the effects of such events on time-integrated nutrient supplies to the phytoplankton community. Therefore, the higher observed Chl-*a* concentrations and cyanobacterial biovolumes in shallower lakes could be related to the increased average nutrient supply in these systems, due to frequent wave-induced nutrient fluxes. However, we cannot completely discount the possibility that in deep lakes, deep mixing and subsequent light limitation of primary production, results in a lower phytoplankton/cyanobacterial biomass (Sakamoto 1966, Berger et al. 2006, Phillips et al. 2008).

Effects of mean depth were also strongly supported in analyses of composition metrics (MFGI, FTI), suggesting systematic changes in community structure and trait representation with changes in lake depth. The final value of the MFGI metric is dependent upon the biomass, trophic score and indicator value of each morpho-functional group. In general, functional groups with high indicator values give more precise trophic classifications. Focusing on groups with top indicator values, increases in the MFGI metric can be explained by an increasing biomass contribution of functional groups with high trophic score. Conversely, an increasing biomass contribution of groups with low trophic score can decrease the metric value. High values of MFGI (such as in shallow lakes) indicate an increasing biomass of large, colonial, vacuolated Chroococcales or Nostocales. Low MFGI values (deep lakes) indicate an increasing biomass of xanthophytes, small pennate diatoms, small centric diatoms or Oscillatoriales. The inverse relationship between MFGI and depth seems to be driven by the trophic preferences of these functional groups, with the most eutrophic colonial Chroococcales and Nostocales being more abundant in shallow lakes. The results for these trait metrics may therefore suggest that the effect of mean depth is via correlated changes in the frequency of episodic nutrient release, as hypothesized above for Chl-*a* and cyanobacterial biovolume.

However, for each metric, considerable among-lake variation remained unexplained by the available environmental data. This was particularly the case for the composition (PTI, MFGI, SPI, FTI) and bloom (total cyanobacterial biovolume, evenness) metrics. While some of this variation might arise due to measurement errors in some of the environmental variables, this would also suggest the existence of important unmeasured drivers of phytoplankton community structure. Geographic variables were included in the analysis as a proxy for the effects of broad climatic gradients upon community structure, via lake physical processes, but the effects of grazing, flushing, water colour, silica or even other parameters associated with eutrophication pressure, such as dissolved nitrogen and turbidity, are all likely to be influential. However, these variables were not recorded consistently enough to resolve their effects in the current analysis. Such unexplained among-lake variability is also likely to arise due to the temporal dimension inherent in these interactions. Current phytoplankton community structure is a biological response to previous environmental conditions (Madgwick et al. 2006), with the time lag of the relationship determined by the time-scale over which phytoplankton gather resources and replicate. For this reason, the relationship between metrics and environmental drivers might be better resolved when these variables can be integrated over the growing season. In lakes with suitable time-series data it would, in principle, be possible to model temporal variability in metric scores as a further source of uncertainty, and also include the temporal relationship between metrics and drivers. Explicit consideration of these temporal aspects could not be achieved here due to the sampling design, but this is highly recommended for future research.

For six of the seven metrics there was evidence that not only mean values, but also variability, changed systematically with among-lake variations in physical, chemical and geographical attributes. Residual variability in metrics was not constant with respect to total phosphorus concentration (Chl-*a*, evenness, total cyanobacterial biovolume), mean depth (Chl-*a*, PTI, MFGI) or altitude (PTI, SPI). Furthermore, the association of this variability with specific drivers differed among metrics e.g. increases in total phosphorus concentration led to increased variability in total cyanobacterial biovolume, but decreases in variability in evenness and Chl-*a*. These findings are similar to the observations of Clarke *et al*. (2006a), who found that the sampling variability of macroinvertebrate community metrics can vary as a function of the overall ecological quality of a site (i.e. the average metric score). Plots of residual metric variability against predictor variables for some of the metrics in the present analysis suggested that a greater spread of metric variation for only a small proportion of the 32 study lakes compared to rest was sufficient for the inclusion of these variance structures to result in an improvement in overall model fit, as judged by AIC. If a future study were to compile data from a larger number of lakes it would be possible to assess how robust these among-lake gradients in metric variability are. For now, the present results suggest that phytoplankton metric variability, and therefore uncertainty, may differ with attributes of the environment from which the phytoplankton samples were drawn and that this may be an important consideration when planning monitoring programmes.

By analysing the results of a unique pan-European hierarchical sampling programme we have shown that seven candidate phytoplankton community metrics, being considered for Intercalibration under the Water Framework Directive, show the potential to indicate among lake variations in the effects of environmental pressures. This is particularly true for Chl-*a* concentration, PTI and total cyanobacterial biovolume, which appear to respond to variations in total phosphorus concentration as a proxy of eutrophication. These metrics are clearly also responsive to variations in other among-lake attributes, some unidentified. In order to further assess the performance of such metrics, it is essential to examine the temporal dimension of their variability (Sondergaard et al. 2011) and also the extent to which uncertainty in water body assessment may vary systematically among lakes differing in their physico-chemical and ecological attributes. These should be considered priorities for future research into freshwater ecological quality assessment.

Acknowledgements

This paper is a result of the project WISER (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) funded by the European Union under the 7th Framework Programme, Theme 6 (Environment including Climate Change) (contract No. 226273), www.wiser.eu. We are indebted to all who processed the samples from the field campaign: Ana Negro (University of Salamanca), Tatiana Caraballo (CEAB), Roser Farres (CEAB), Pierisa Panzani (CNR), Maria Antonietta Mariani (University of Sassari), Aimar Rakko (EMU), Małgorzata Poniewozik (IEP), Barbara Meinck (IGB), Andreas Nicklisch (IGB), Annette Tworeck (LBH Freiburg), Mitzi De Ville (NERC-CEH), Rene Groben (NERC-CEH), Sarah Pritchard (Beacon Biological) & Bill Brierly (EA).

References

Barton, K. 2011. MuMIn: Multi-model inference. R package version 1.0.0. http://CRAN.R-project.org/package=MuMIn.

Bates, D., M. Maechler, and B. Bolker. 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-41. http://CRAN.R-project.org/package=lme4.

Berger, S. A., S. Diehl, T. J. Kunz, D. Albrecht, A. M. Oucible, and S. Ritzer. 2006. Light supply, plankton biomass, and seston stoichiometry in a gradient of lake mixing depths. Limnology and Oceanography **51**:1898-1905.

Brierly, B., L. Carvalho, S. Davies, and J. Krokowski. 2007. Guidance on the quantitative analysis of phytoplankton in freshwater samples. Edinburgh.

Brown, C. D., M. V. Hoyer, R. W. Bachmann, and D. E. Canfield Jr. 2000. Nutrient-chlorophyll relationships: an evaluation of empirical nutrient-chlorophyll models using Florida and north-temperate lake data. Canadian Journal of Fisheries and Aquatic Sciences **57**:1574-1583.

Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.

Carvalho, L., L. Lepisto, J. Rissanen, O.-P. Pietiläinen, S. Rekolainen, L. Torok, A. Lyche Solheim, T. Saloranta, R. Ptacnik, G. Tartari, A. C. Cardoso, G. Premazzi, I. Gunn, E. Penning, J. Hanganu, S. Hellsten, I. Orhan, and I. Navodaru. 2006. Nutrients and eutrophication in lakes. Pages 3-32 *in* S. A., C. A.C., and H. A.-S., editors. Indicators and methods for the ecological status assessment under the Water Framework Directive: Linkages between chemical and biological quality of surface waters. European Commission, Luxembourg.

CEN EN 15204. 2006. Water quality – Guidance standard for the routine analysis of phytoplankton abundance and composition using inverted microscopy (Utermöhl technique).

Clarke, R. T., J. Davy-Bowker, L. Sandin, N. Friberg, R. K. Johnson, and B. Bis. 2006a. Estimates and comparisons of the effects of sampling variation using "national" macroinvertebrate sampling protocols on the precision of metrics used to assess ecological status. Hydrobiologia **566**:477-503.

Clarke, R. T. and 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.

Clarke, R. T., A. Lorenz, L. Sandin, A. Schmidt-Kloiber, J. Strackbein, N. T. Kneebone, and P. Haase. 2006b. Effects of sampling and sub-sampling variation using the STAR-AQEM sampling protocol on the precision of macroinvertebrate metrics. Hydrobiologia **566**:441-459.

Dillon, P. J. and F. H. Rigler. 1974. The phosphorus-chlorophyll relationship in lakes. Limnology and Oceanography **19**:767-773.

Dudley, B., M. Dunbar, E. Penning, A. Kolada, S. Hellsten, and A. Kanninen. 2010. WISER Deliverable D3.2-2: Report on uncertainty in macrophyte metrics.

EC. 2000. Directive 2000/60/EC of the European Parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities **L327**:1-72.

Elliott, J. A., I. D. Jones, and S. J. Thackeray. 2006. Testing the sensitivity of phytoplankton communities to changes in water temperature and nutrient load, in a temperate lake. Hydrobiologia **559**:401-411.

Hamilton, D. P. and S. F. Mitchell. 1988. Effects of wind on nitrogen, phosphorus, and chlorophyll in a shallow New Zealand lake Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie **23**:624-628.

Hering, D., A. Borja, J. Carstensen, L. Carvalho, M. Elliott, C. K. Feld, A.-S. Heiskanen, R. K. Johnson, J. Moe, D. Pont, A. L. Solheim, and W. van de Bund. 2010. The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. Science of the Total Environment **408**:4007-4019.

ISO 10260. 1992. Water quality – Measurement of biochemical parameters – Spectrometric determination of the chlorophyll-*a* concentration. Page 6 Int. Org. Standard., Geneva, 1st edn. 1992–07–15.

Jasser, I. 1995. The influence of macrophytes on a phytoplankton community in experimental conditions. Hydrobiologia **306**:21-32.

Kalff, J. 2002. Limnology. Prentice-Hall Inc., Upper Saddle River, NJ

Kamenir, Y. and G. Morabito. 2009. Lago Maggiore oligotrophication as seen from the long-term evolution of its phytoplankton taxonomic size structure. Journal of Limnology **68**:146-161.

Kuemmerlin, R. E. 1998. Taxonomical response of the phytoplankton community of Upper Lake Constance (Bodensee-Obersee) to eutrophication and re-oligotrophication. Advances in Limnology **53**:109-117.

Mackay, E. B., I. D. Jones, A. M. Folkard, and S. J. Thackeray. 2011. Transition zones in small lakes: the importance of dilution and biological uptake on lake-wide heterogeneity. Hydrobiologia **678**:85-97.

Madgwick, G., I. D. Jones, S. J. Thackeray, J. A. Elliott, and H. J. Miller. 2006. Phytoplankton communities and antecedent conditions: high resolution sampling in Esthwaite Water. Freshwater Biology **51**:1798-1810.

McCauley, E., J. A. Downing, and S. Watson. 1989. Sigmoid relationships between nutrients and chlorophyll among lakes. Canadian Journal of Fisheries and Aquatic Sciences **46**:1171-1175.

Mischke, U., L. Carvalho, C. McDonald, B. Skjelbred, A. L. Solheim, G. Phillips, C. de Hoyos, G. Borics, and J. Moe. 2010. WISER Deliverable D3.1-2: Report on phytoplankton bloom metrics.45.

Murphy, J. and J. P. Reilly. 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chim, Acta **27**:31-36.

Murphy, K. J., M. P. Kennedy, V. McCarthy, M. T. Ó'Hare, K. Irvine, and C. Adams. 2002. A review of ecology based classification systems for standing freshwaters. SNIFFER Project Number: W(99)65, Environment Agency R&D Technical Report: E1-091/TR.

National Rivers Authority. 1995. Test Methods and Procedures: Freshwater Phytoplankton. National Rivers Authority, Peterborough.

Padisák, J. and C. S. Reynolds. 1998. Selection of phytoplankton associations in Lake Balaton, Hungary, in response to eutrophication and restoration measures, with special reference to the cyanoprokaryotes. Hydrobiologia **384**:41-53.

Phillips, G., O.-P. Pietiläinen, L. Carvalho, A. Solimini, A. Lyche Solheim, and A. C. Cardoso. 2008. Chlorophyll–nutrient relationships of different lake types using a large European dataset Aquatic Ecology **42**:213-226.

Phillips, G., B. Skjelbred, G. Morabito, L. Carvalho, A. L. Solheim, T. Andersen, U. Mischke, C. de Hoyos, and G. Borics. 2010. WISER Deliverable D3.1-1: Report on phytoplankton composition metrics, including a common metric approach for use in intercalibration by all GIGs.

Pielou, E. C. 1969. An introduction to mathematical ecology. Wiley-Interscience, New York.

Pielou, E. C. 1975. Ecological diversity. Wiley Interscience, New York.

Pinel-Alloul, B. and A. Ghadouani. 2007. Spatial heterogeneity of planktonic microorganisms in aquatic systems: mutiscale patterns and processes.*in* R. Franklin and A. Mills, editors. The Spatial Distribution of Microbes in the Environment. Springer.

Pinel-Alloul, B., G. Méthot, G. Verrault, and V. Y. 1990. Phytoplankton in Quebec lakes: Variation with lake morphometry, and with natural and anthropogenic Acidification. Canadian Journal of Fisheries and Aquatic Sciences **47**:1047-1057.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Development Core Team. 2010. nlme: Linear and nonlinear mixed effects models. R package version 3.1-97.

Pridmore, R. D., W. N. Vant, and J. C. Rutherford. 1985. Chlorophyll-nutrient relationships in North Island lakes (New Zealand). Hydrobiologia **121**:181-189.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, ISBN 3-900051-07-0, url = http://www.R-project.org/, Vienna, Austria.

Reynolds, C. S. 2006. Ecology of phytoplankton. Cambridge University Press, Cambridge.

Sakamoto, M. 1966. Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. Archive fur Hydrobiologie **62**:1-28.

Salmaso, N. and J. Padisak. 2007. Morpho-Functional Groups and phytoplankton development in two deep lakes (Lake Garda, Italy and Lake Stechlin, Germany). Hydrobiologia **578**:97-112.

Schindler, D. W. 1978. Factors regulating phytoplankton production and standing crop in the world's freshwaters. Limnology and Oceanography **23**:478-486.

Schriver, P., J. Bogestrand, E. Jeppesen, and M. Sondergaard. 1995. Impact of submerged macrophytes on fish-zooplankton-phytoplankton interactions: large-scale enclosure experiments in a shallow eutrophic lake. Freshwater Biology **33**:255-270.

Smith, V. H. 1985. Predictive models for the biomass of blue-green algae in lakes. Water Resources Bulletin **21**:433-439.

Smith, V. H., E. Willen, and B. Karlsson. 1987. Predicting the summaer peak biomass of four species of blue-green algae (Cyanophyta/Cyanobacteria) in Swedish lakes. Water Resources Bulletin **23**:397-402.

Sondergaard, M., S. E. Larsen, T. B. Jorgensen, and E. Jeppesen. 2011. Using chlorophyll a and cyanobacteria in the ecological classification of lakes. Ecological Indicators **11**:1403-1412.

Utermöhl, H. 1958. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. Mitteilungen der Internationale Vereinigung für Theoretische und Angewandte Limnologie **9**:1-38.

Vuorio, K., L. Lepistö, and A.-L. Holopainen. 2007. Intercalibrations of freshwater phytoplankton analyses. Boreal Environment Research **12**:561-569.

Warnes, G. R. 2010. gplots: Various R programming tools for plotting data. R package version 2.8.0, url = {http://CRAN.R-project.org/package=gplots.

Watson, S. B., E. McCauley, and J. A. Downing. 1997. Patterns in phytoplankton taxonomic composition across temperate lakes of different nutrient status. Limnology and Oceanography **42**:487-495.

WISE. 2008. Water Note 7. Intercalibration: A common scale for Europe's waters.

Wium-Andersen, S., U. Anthoni, C. Christophersen, and G. Houen. 1982. Allelopathic effects on phytoplankton bysubstances isolated from aquatic macrophytes (Charales). Oikos **39**:187-190.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York.

Table 1. Lakes sampled in the field campaign. GIG indicates the Geographical Intercalibration Group within which each lake falls: AL = Alpine, CB = Central/Baltic, M = Mediterranean, N = Northern. Only Chl-*a* data were available for lakes marked with an asterisk.

Lake	***Country***	***GIG***	***Latitude (°N)***	***Longitude (°W)***	***Mean depth (m)***	***Maximum depth (m)***	***Altitude (m a.s.l.)***	***Total phosphorus (mg m^{-3})***	***Alkalinity (meq L^{-1})***
Nordborgsø	*Denmark*	*CB*	*55.06*	*9.76*	*5.0*	*8.5*	*20*	*62.67*	*2.30*
Fussingsø	*Denmark*	*CB*	*56.47*	*9.88*	*12.6*	*31.0*	*15*	*45.67*	*1.50*
Saadjärv	*Estonia*	*CB*	*58.54*	*26.65*	*8.0*	*21.7*	*85*	*14.00*	*2.53*
Viljandi	*Estonia*	*CB*	*58.35*	*25.60*	*5.5*	*9.5*	*75*	*21.50*	*4.40*
Sääksjärvi	*Finland*	*N*	*62.17*	*25.73*	*9.3*	*15.2*	*121*	*12.00*	*0.23*
Vuojärvi	*Finland*	*N*	*62.41*	*25.94*	*4.4*	*10.2*	*91*	*35.5*	*0.54*
Iso-Jurvo	*Finland*	*N*	*62.60*	*25.93*	*8.6*	*29.6*	*139*	*8.00*	*0.06*
Salagou	*France*	*M*	*43.66*	*3.40*	*15.6*	*49.3*	*139*	*21.76*	*2.77*
Caramany	*France*	*M*	*42.74*	*2.59*	*14.5*	*36.0*	*170*	*26.80*	*2.96*
Glindower See	*Germany*	*CB*	*52.36*	*12.92*	*4.9*	*14.3*	*24*	*151.00*	*2.40*

Grienericksee	*Germany*	*CB*	*53.10*	*12.89*	*4.7*	*11.5*	*55*	*19.00*	*2.20*
Roofensee	*Germany*	*CB*	*53.11*	*13.02*	*9.0*	*19.1*	*59*	*18.00*	*2.00*
Alserio	*Italy*	*AL*	*45.78*	*9.21*	*5.0*	*8.0*	*243*	*24.00*	*2.34*
Bidighinzu	*Italy*	*M*	*40.56*	*8.66*	*7.5*	*21.8*	*330*	*65.00*	*2.24*
Candia	*Italy*	*AL*	*45.33*	*7.92*	*5.0*	*7.5*	*226*	*16.50*	*1.00*
Monate	*Italy*	*AL*	*45.80*	*8.66*	*18.0*	*34.0*	*266*	*8.50*	*0.88*
Segrino	*Italy*	*AL*	*45.83*	*9.27*	*3.5*	*8.0*	*374*	*12.50*	*2.23*
Nøklevann	*Norway*	*N*	*59.88*	*10.88*	*19.0*	*31.0*	*163*	*4.00*	*0.17*
Longumvatnet	*Norway*	*N*	*58.49*	*8.76*	*14.0*	*35.5*	*34*	*7.50*	*0.28*
Temse	*Norway*	*N*	*58.38*	*8.64*	*6.0*	*10.2*	*15*	*17.00*	*0.32*
Rumian	*Poland*	*CB*	*53.38*	*20.00*	*6.0*	*14.0*	*152*	*88.00*	*2.60*
Lidzbarskie	*Poland*	*CB*	*53.26*	*19.80*	*10.0*	*24.0*	*128*	*56.50*	*2.45*
Kielpinskie	*Poland*	*CB*	*53.35*	*19.79*	*5.8*	*10.0*	*120*	*63.50*	*2.90*

Vencías, Las	*Spain*	*M*	*41.43*	*-3.96*	*8.0*	*14.8*	*869*	*20.46*	*2.43*
Vega de Jabalón	*Spain*	*M*	*38.76*	*-3.79*	*6.6*	*10.8*	*635*	*54.65*	*2.26*
Arquillo de San Blás	*Spain*	*M*	*40.36*	*-1.21*	*34.0*	*38.0*	*970*	*6.90*	*2.80*
*Fiolen**	*Sweden*	*N*	*57.08*	*14.53*	*3.8*	*10.0*	*226*	*10.00*	*0.10*
*Skirösjön**	*Sweden*	*N*	*57.36*	*15.38*	*5.2*	*8.0*	*146*	*45.33*	*0.63*
*Västra Solsjön**	*Sweden*	*N*	*59.08*	*12.29*	*12.3*	*40.0*	*147*	*10.00*	*0.16*
Loweswater	*UK*	*N*	*54.58*	*-3.36*	*8.0*	*14.8*	*125*	*9.97*	*0.22*
Grasmere	*UK*	*N*	*54.45*	*-3.02*	*8.4*	*19.4*	*61*	*9.15*	*0.21*
Rostherne mere	*UK*	*CB*	*53.35*	*-2.39*	*11.5*	*29.7*	*27*	*121.00*	*2.44*

Table 2. Proportions of metric variance at different levels in the sampling hierarchy, for null models of the six different metrics. Total between = Country + Lake, Total within = Station + Sample + Analyst + Error (sub-sample). Models fitted using REML estimation. Also included are standard deviations (SD) of each metric, after re-scaling to zero mean.

Metric	**SD**	**Country**	**Lake**	**Station**	**Sample**	**Analyst**	**Error (sub-sample)**	**Total within**	**Total between**
Log_{10} Chl-*a*	0.55	0.00	0.96	0.01	0.01	-	0.02	0.04	0.96
PTI	0.53	0.00	0.88	<0.01	0.00	0.04	0.07	0.12	0.88
SPI	0.03	0.00	0.65	0.03	0.00	0.19	0.13	0.35	0.65
MFGI	0.05	0.00	0.86	0.02	<0.01	0.05	0.08	0.14	0.86
FTI	0.04	0.00	0.81	0.02	0.00	0.09	0.08	0.19	0.81
Evenness	0.17	0.00	0.69	0.04	0.00	0.17	0.10	0.31	0.69
Log_{10} total cyanobacteria	0.74	0.09	0.86	0.01	0.00	0.02	0.03	0.06	0.94

Table 3. Summary of the most optimal linear mixed-effects models for each of the seven phytoplankton metrics. Shown are the number of estimated model parameters (*k*), the predictors present in the model, the difference in AIC between the most optimal model and the corresponding null model (ΔAIC_{null}) and the Akaike weight; a measure of the relative level of support for the most optimal model, compared to other candidate models, given the data. For the Akaike weight, values close to 1 indicate overwhelming support for the corresponding model, while lower values indicate the presence of other models with similar levels of support. For each predictor, the sign of the corresponding relationship is given as positive (+) or negative (-). Models fitted using ML estimation.

Metric	***k***	**Predictors**	Δ**AIC**$_{\mathbf{null}}$	**Akaike weight**
Log_{10} Chl-*a*	6	Log_{10} Mean lake depth (-) Log_{10} total phosphorus (+) Latitude (+)	35.5	0.12
PTI	7	Log_{10} Mean lake depth (-) Log_{10} total phosphorus (+) Log_{10} Altitude (+)	13.5	0.11
SPI	6	Log_{10} Mean lake depth (-) Log_{10} Altitude (+)	2.0	0.12
MFGI	6	Log_{10} Mean lake depth (-) Log_{10} Altitude (+)	10.0	0.12
FTI	6	Log_{10} Mean lake depth (-) Log_{10} Altitude (+)	9.0	0.19
Evenness	6	Log_{10} total phosphorus (-) Alkalinity (+)	2.3	0.06
Log_{10} total cyanobacteria	6	Log_{10} Mean lake depth (-) Log_{10} total phosphorus (+)	16.2	0.13

Table 4. AIC comparison of the most optimal linear mixed-effects models for each of the seven phytoplankton metrics (see Table 3), when including/excluding variance structures to account for changes in residual metric variance as a function of the fitted predictors. Shown are the predictors that residual variability is modeled as a function of (Predictor), the estimated delta parameter for the exponential function describing the relationship between residual variance and the named predictor (δ) and the AIC for each model. For each metric, the most optimal model is indicated in bold. Models fitted using REML estimation.

Metric	Model No.	Predictor	δ	AIC
Log$_{10}$ Chl-*a*	1	None	-	-195.1
	2	Log$_{10}$ Mean lake depth	0.88	-205.7
	3	Log$_{10}$ total phosphorus	-0.70	-230.7
	4	Latitude	0.02	-198.3
	5	**Log$_{10}$ Mean lake depth**	**0.57**	**-233.8**
		Log$_{10}$ total phosphorus	**-0.65**	
	6	Log$_{10}$ total phosphorus	-0.70	-228.8
		Latitude	<0.01	
	7	Log$_{10}$ Mean lake depth	0.75	-205.3
		Latitude	0.01	
	8	Log$_{10}$ Mean lake depth	0.57	-231.8
		Log$_{10}$ total phosphorus	-0.66	
		Latitude	<-0.01	
PTI	1	None	-	-138.7
	2	Log$_{10}$ Mean lake depth	-0.44	-144.9
	3	Log$_{10}$ total phosphorus	-0.40	-147.9
	4	Log$_{10}$ Altitude	0.66	-180.4
	5	Log$_{10}$ Mean lake depth	-0.53	-156.3
		Log$_{10}$ total phosphorus	-0.43	
	6	Log$_{10}$ total phosphorus	-0.11	-179.0
		Log$_{10}$ Altitude	0.62	
	7	**Log$_{10}$ Mean lake depth**	**-0.39**	**-183.5**
		Log$_{10}$ Altitude	**0.65**	
	8	Log$_{10}$ Mean lake depth	-0.43	-183.1
		Log$_{10}$ total phosphorus	-0.17	
		Log$_{10}$ Altitude	0.59	
SPI	1	None	-	-1682.9
	2	Log$_{10}$ Mean lake depth	0.19	-1682.8
	3	**Log$_{10}$ Altitude**	**0.23**	**-1688.6**
	4	Log$_{10}$ Mean lake depth	-0.06	-1686.7
		Log$_{10}$ Altitude	0.25	
MFGI	1	None	-	-1760.6
	2	**Log$_{10}$ Mean lake depth**	**-0.43**	**-1763.3**
	3	Log$_{10}$ Altitude	-0.12	-1760.7
	4	Log$_{10}$ Mean lake depth	-0.42	-1763.3
		Log$_{10}$ Altitude	-0.12	
FTI	**1**	**None**	**-**	**-1854.2**
	2	Log$_{10}$ Mean lake depth	-0.15	-1853.1
	3	Log$_{10}$ Altitude	0.01	-1852.2
	4	Log$_{10}$ Mean lake depth	-0.19	-1851.3
		Log$_{10}$ Altitude	0.04	
Evenness	1	None	-	-621.7
	2	**Log$_{10}$ total phosphorus**	**-0.51**	**-642.6**

	3	Alkalinity	-0.13	-633.8
	4	Log_{10} total phosphorus	-0.42	-641.6
		Alkalinity	-0.04	
Log_{10} total cyanobacteria	1	None	-	-171.6
	2	Log_{10} Mean lake depth	-0.52	-177.1
	3	**Log_{10} total phosphorus**	**0.71**	**-214.4**
	4	Log_{10} Mean lake depth	-0.23	-214.0
		Log_{10} total phosphorus	0.67	

Figure legends

Fig. 1. The hierarchic sampling design employed in each lake. Samples were collected from three stations, above the deepest point (z_{max}), the mean depth (z_{mean}) and a depth intermediate between the maximum and mean depths (z_{int}). Two samples (S1, S2) were collected at each station. At each station, three sub-samples (Sub1, Sub2, Sub3) were collected from sample 1 and one sub-sample from sample 2. In each case, two sub-samples from the first sample and the only sub-sample from the second sample were processed by one analyst (An1 or An2), while the third sub-sample from sample one was processed by a different analyst (An1 or An2).

Fig. 2. Scatterplots of lake-averaged values of the seven phytoplankton metrics against $\log_{10}$ total phosphorus concentration.

Fig. 3. The proportion of the among-lake variance in metric scores “explained” in top models, with the most optimal combination of environmental predictor variables. REML estimation used in model fitting.

Fig. 4. Model-averaged parameter estimates for the relationships between the modelled environmental predictors and the phytoplankton abundance metric ($\log_{10}$ Chl-*a* concentration). Filled circles indicate the model-averaged slope parameter estimate for each metric-predictor relationship, and whiskers indicate the 95% confidence interval for the estimate. Dashed horizontal line indicates zero. ML estimation used in model fitting.

Fig. 5. Model-averaged parameter estimates for the relationships between the modelled environmental predictors and the four phytoplankton composition metrics. Filled circles indicate the model-averaged slope parameter estimate for each metric-predictor relationship, and whiskers indicate the 95% confidence interval for the estimate. Dashed horizontal line indicates zero. ML estimation used in model fitting.

Fig. 6. Model-averaged parameter estimates for the relationships between the modelled environmental predictors and the two phytoplankton bloom metrics. Filled circles indicate the model-averaged slope parameter estimate for each metric-predictor relationship, and whiskers indicate the 95% confidence interval for the estimate. Dashed horizontal line indicates zero. ML estimation used in model fitting.

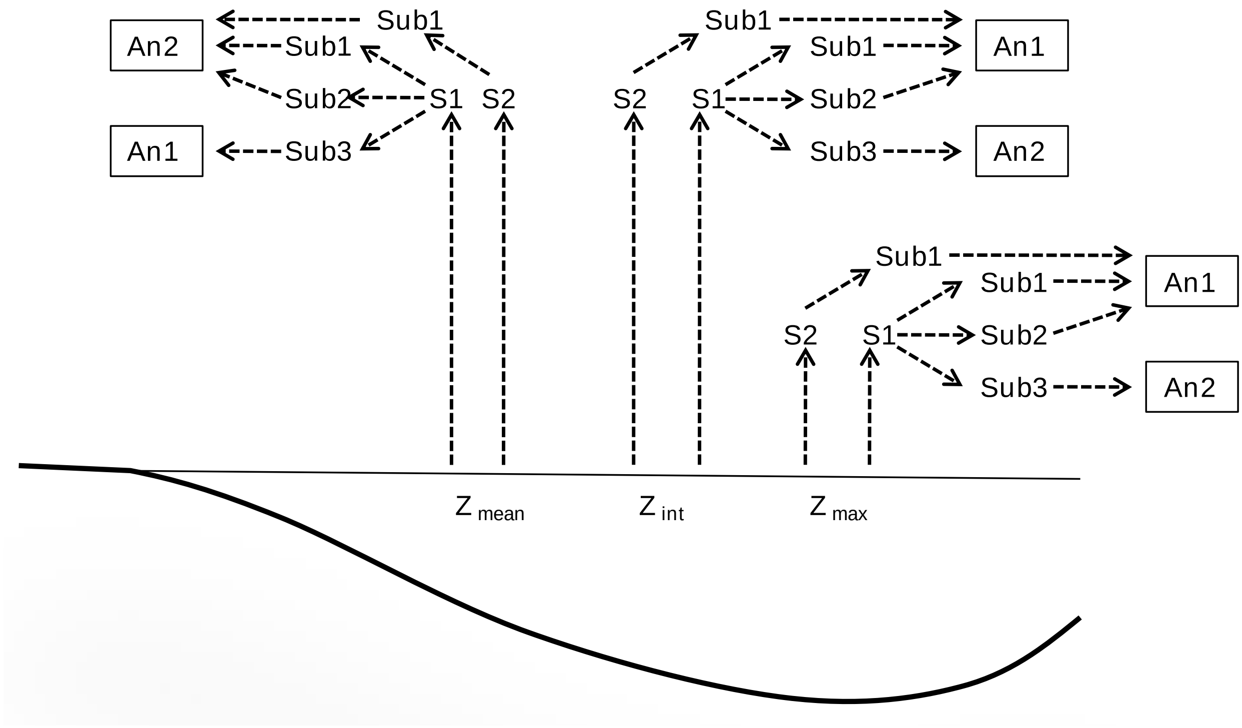
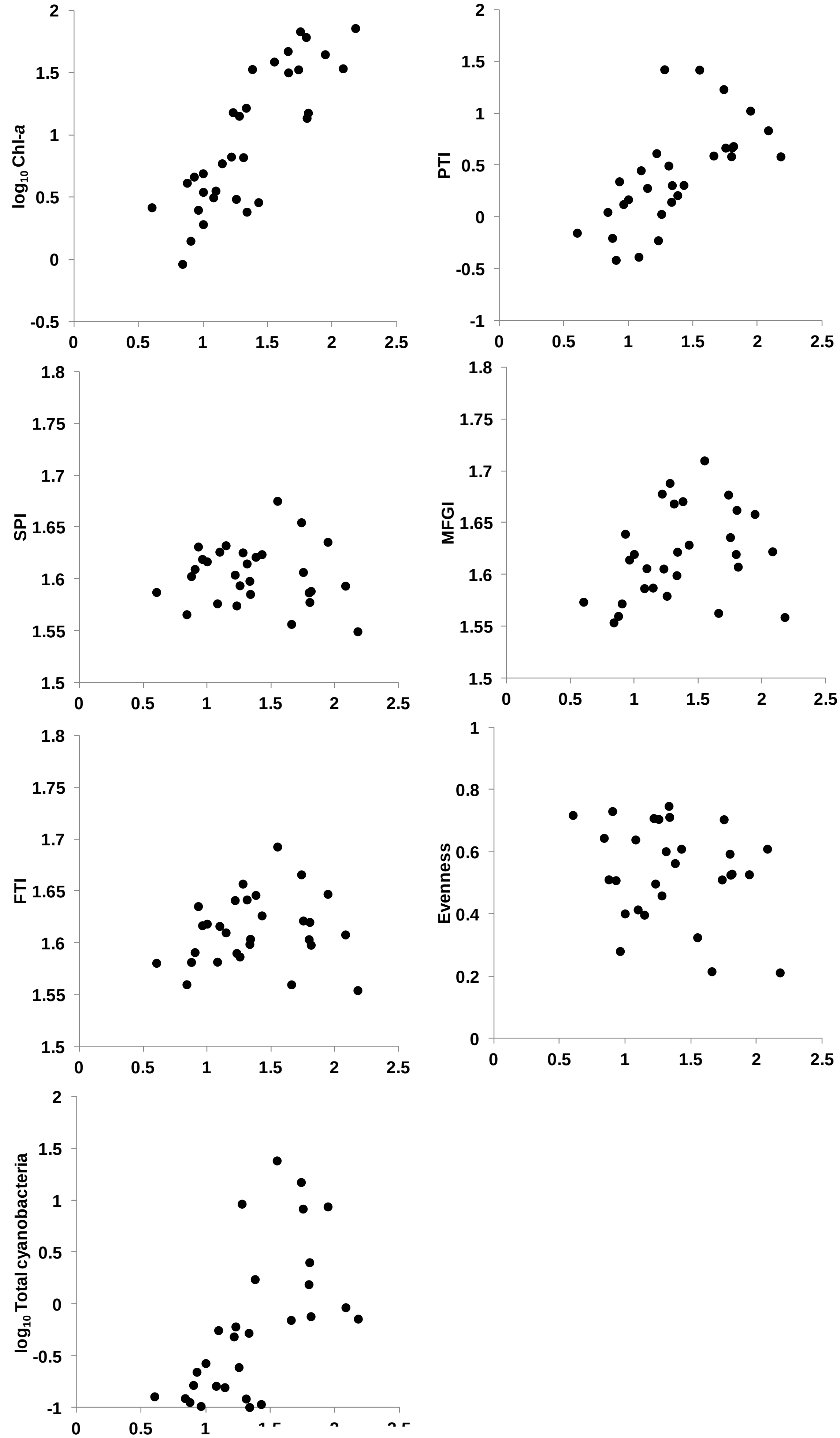

An2
Sub1
Sub1
Sub2
S1
S2
An1
Sub3
Sub1
Sub1
An1
S2
S1
Sub2
Sub3
An2
Sub1
Sub1
An1
S2
S1
Sub2
Sub3
An2
Z mean
Z int
Z max

Fig. 1



2
1.5
1
0.5
0
-0.5
log$_{10}$ Chl-a
0
0.5
1
1.5
2
2.5
2
1.5
1
0.5
0
-0.5
-1
PTI
0
0.5
1
1.5
2
2.5
1.8
1.75
1.7
1.65
1.6
1.55
1.5
SPI
0
0.5
1
1.5
2
2.5
1.8
1.75
1.7
1.65
1.6
1.55
1.5
MFGI
0
0.5
1
1.5
2
2.5
1.8
1.75
1.7
1.65
1.6
1.55
1.5
FTI
0
0.5
1
1.5
2
2.5
1
0.8
0.6
0.4
0.2
0
Evenness
0
0.5
1
1.5
2
2.5
2
1.5
1
0.5
0
-0.5
-1
log$_{10}$ Total cyanobacteria
0
0.5
1
1.5
2
2.5
Log$_{10}$ Total phosphorus concentration

Fig. 2.

1.0

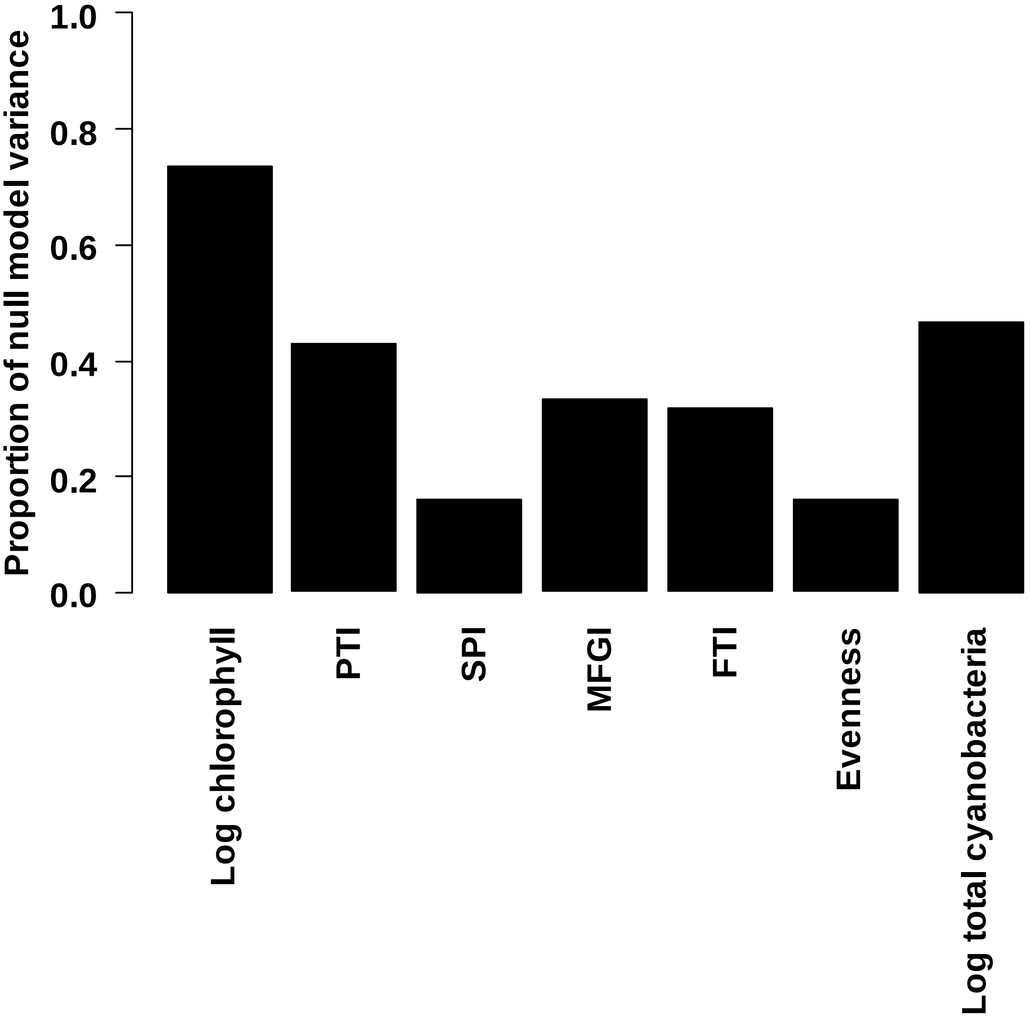

Proportion of null model variance
1.0
0.8
0.6
0.4
0.2
0.0
Log chlorophyll
PTI
SPI
MFGI
FTI
Evenness
Log total cyanobacteria

Fig. 3.

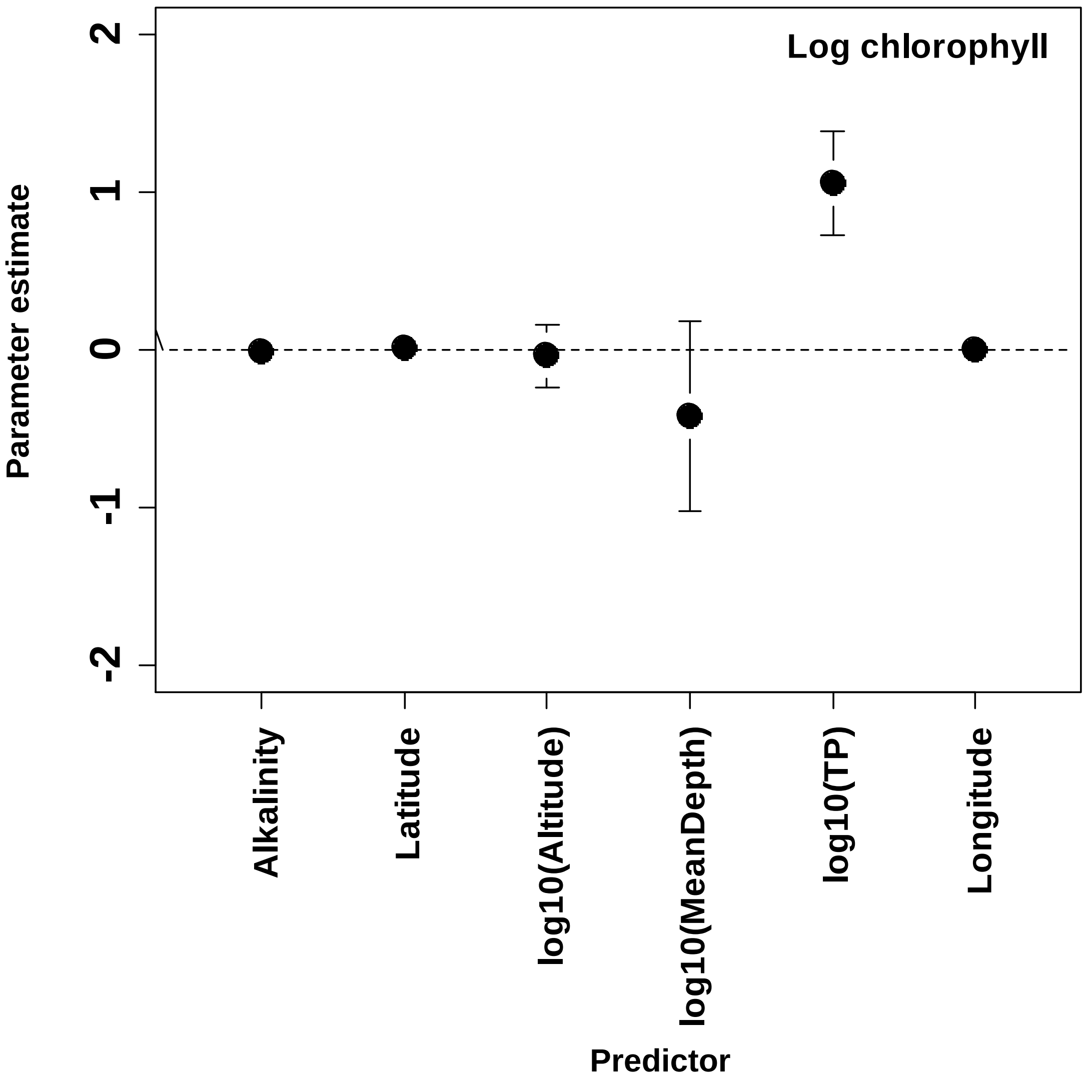

Log chlorophyll
Parameter estimate
2
1
0
-1
-2
Alkalinity
Latitude
log10(Altitude)
log10(MeanDepth)
log10(TP)
Longitude
Predictor

Fig. 4

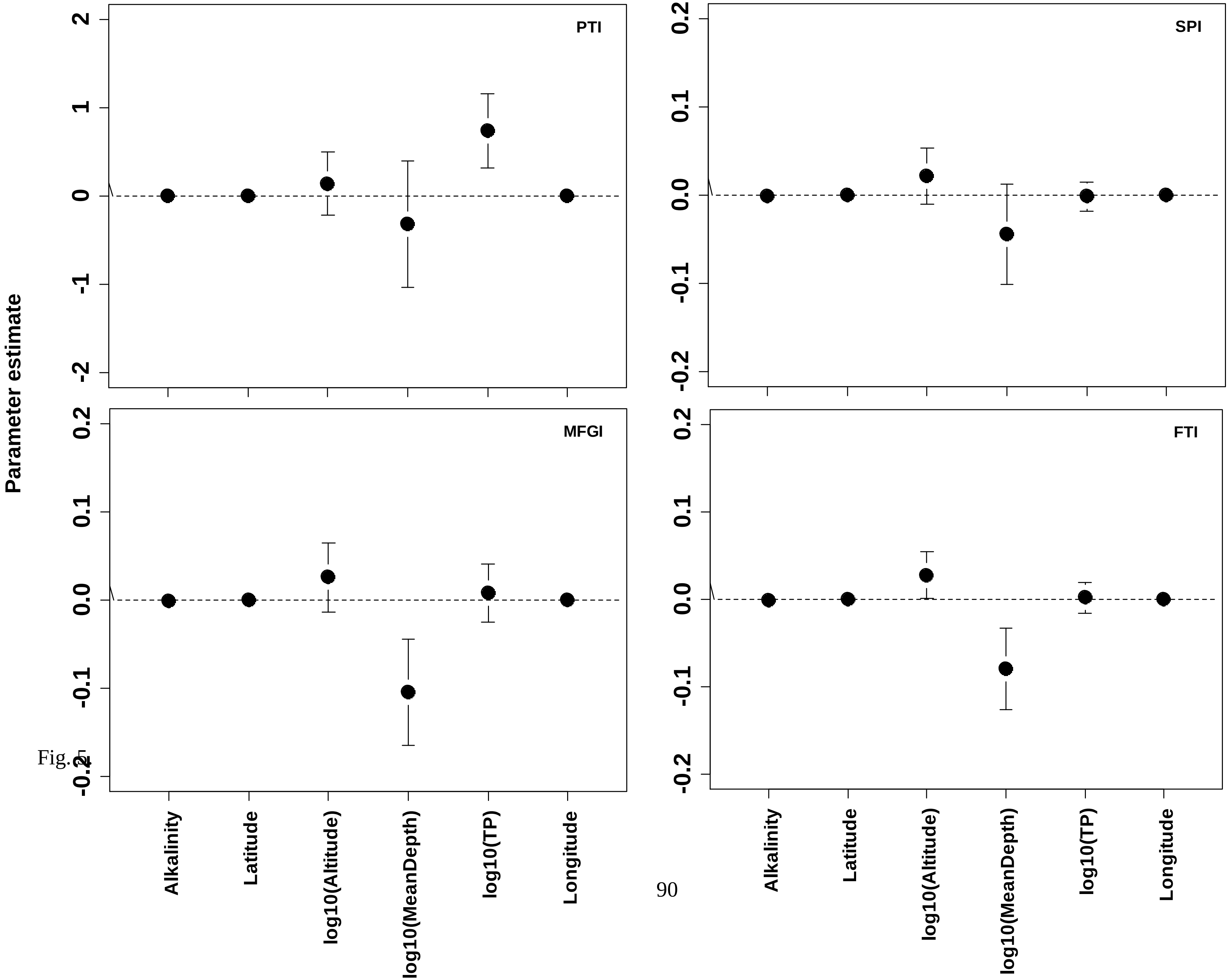

Parameter estimate
PTI
2
1
0
-1
-2
SPI
0.2
0.1
0.0
-0.1
-0.2
MFGI
0.2
0.1
0.0
-0.1
-0.2
FTI
0.2
0.1
0.0
-0.1
-0.2
Alkalinity
Latitude
log10(Altitude)
log10(MeanDepth)
log10(TP)
Longitude
Alkalinity
Latitude
log10(Altitude)
log10(MeanDepth)
log10(TP)
Longitude

Fig. 5

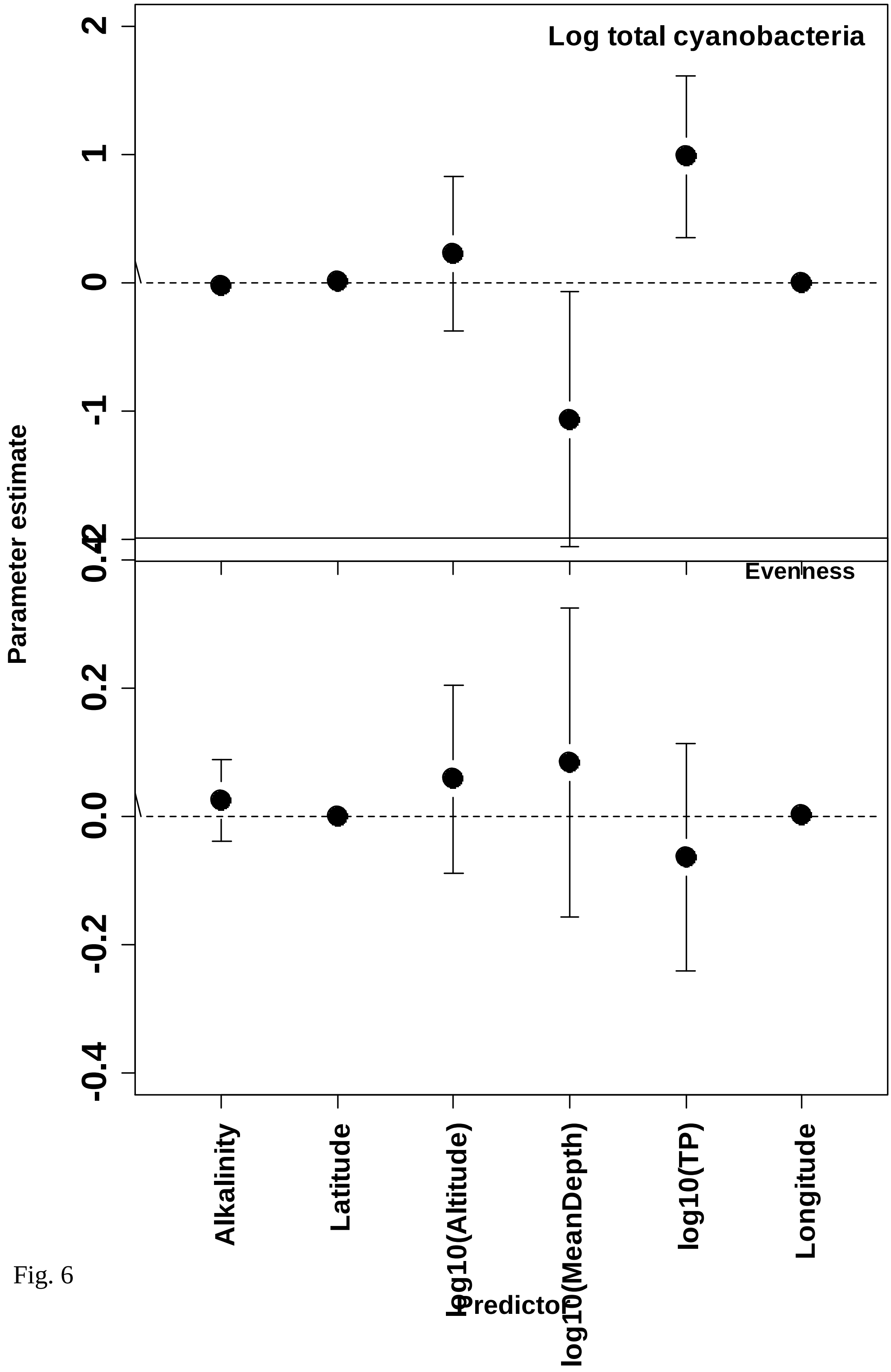

Log total cyanobacteria
Evenness
Parameter estimate
2
1
0
-1
-2
0.4
0.2
0.0
-0.2
-0.4
Alkalinity
Latitude
log10(Altitude)
log10(MeanDepth)
log10(TP)
Longitude
Predictor

Fig. 6