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

Deliverable D3.1-3:
Uncertainty in Lake Phytoplankton Metrics

Lead contractor: **Centre for Ecology & Hydrology (CEH)**
Contributors: **Stephen Thackeray, Peeter Nõges, Michael Dunbar, Bernard Dudley, Birger Skjelbred, Giuseppe Morabito, Laurence Carvalho, Geoff Phillips, Ute Mischke**

Due date of deliverable: **Month 27**
Actual submission date: **Month 27**

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

Lake phytoplankton metrics proposed by the EC WISER Project for ecological quality assessment of European lakes are shown to be robust metrics. Latest results from the WISER pan-European field campaign reveal that variability in metric scores is largely due to variability between lakes and is significantly related to differences in eutrophication pressure (total phosphorus concentrations). Differences in locations around a lake, or sampling and analytical variability, only account for a small proportion of the variability in metric scores. These results are especially true for four candidate phytoplankton metrics being considered for Intercalibration: chlorophyll, PTI, MFGI and cyanobacteria abundance, for which >85% of the variability in metric scores was attributed between lakes and total phosphorus concentration was the best single predictor of variation in these metrics. Although, much among-lake metric variation still remained unexplained by the available environmental data, we conclude that these four proposed metrics are sufficiently robust metrics for ecological status assessment and are suitable for adoption in the Intercalibration process.

Techinal summary

A central tenet of the Water Framework Directive (WFD) is that attributes of biological communities are used to make inferences about the ecological quality of rivers, lakes and coastal/transitional waters. For lakes, the phytoplankton are a key biological community, or Biological Quality Element (BQE), to be used for this purpose. Phytoplankton communities are complex, containing many species responding to changes in many different environmental factors. In order to use phytoplankton to indicate ecological quality, it is necessary to develop metrics that describe high-level properties of the phytoplankton community that are sensitive to environmental pressures, such as nutrient enrichment. In the earlier stages of the WISER project, six candidate metrics were proposed as either phytoplankton composition or bloom metrics, to be considered alongside the established metric of chlorophyll *a* (chl), which is used to represent phytoplankton abundance.

If such metrics are to be useful indicators of differences in ecological quality among lakes, it is necessary to know the extent to which they can be affected by sampling and sample processing procedures. Phytoplankton communities vary spatially across lakes and so we could expect phytoplankton metrics to show similar variation among sampling locations. Also, the results of phytoplankton sample processing depend upon procedures used to analyse samples in the laboratory and the expertise of the analyst processing the sample (people vary in their abilities to discriminate between species). If phytoplankton metrics vary more with differences in sampling and sample processing procedures within a lake than they do among lakes of different pressure then these metrics are unlikely to provide a sensitive means of describing differences in the biological impacts of an environmental stressor among lakes. On the contrary, little variation with sampling and sample processing compared to variation among lakes would indicate the potential for relationships between metrics and environmental pressures to be detected after accounting for sample/sample processing "errors". Here we analyse the results of a multi-scale field campaign of 32 European lakes, to resolve the extent to which chl and the six proposed WISER 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.

The seven metrics analysed show some similarities in their variation among countries, lakes and samples, indicating that some of them reveal similar information on changes in phytoplankton communities. However correlations among metrics are not perfect, suggesting that each metric is potentially also showing a unique aspect of phytoplankton community change. For all seven metrics, between 65% and 96% of the variability in metric scores is due to variability among lakes, much higher than variability occurring due to sampling/sample processing. After the among-lake variation, errors associated with sub-sampling of field samples (2-13%) and variation in analysts (2-19%) are more important than spatial variation between locations within a lake (<1–4%). Importantly, among-lake variation in most of the metrics can be related to differences in total phosphorus concentrations, though this effect is often better described when also taking account of some aspects of lake geography (longitude and altitude).

These variables typically explain only a modest proportion of total among-lake metric variability, indicating that metrics are additionally sensitive to unmeasured environmental factors. These positive findings on metric robustness and sensitivity to pressure are especially true for four candidate phytoplankton metrics being considered for Intercalibration: chlorophyll, PTI, MFGI and cyanobacterial blooms, for which > 85% of the variability in metric scores was attributed between lakes, and, total phosphorus concentration was the best single predictor of variation in these metrics. This WISER study, therefore, concludes that these four proposed metrics are robust metrics for ecological status assessment and are suitable for adoption in the Intercalibration process. The study was carried out following clear sampling, counting and identification guidelines and it indicates that rigorous standardisation of procedures helps minimise sampling and analytical variability and makes possible more meaningful comparisons of ecological status among lakes.

1 Introduction

The central tenet of the Water Framework Directive (2000/60/EC) is that key communities, or Biological Quality Elements (BQEs), should be used to assess the water quality of lakes, rivers and coastal/transitional waters. The phytoplankton have been identified as a key BQE to be used in lake ecological quality assessment and are widely used as an important water quality indicator because of its high species differentiation and sensitivity to environmental factors. Murphy et al. (2002) list the following main advantages of using phytoplankton in lake monitoring:

1. As primary producers, algae are directly affected by physical and chemical factors and changes in phytoplankton community status have direct implications for the biointegrity of the lake ecosystem as a whole.

2. Algae generally have rapid reproduction rates and very short life cycles, making them valuable indicators of short-term (scales of days – weeks) impacts.

3. Phytoplankton provide a good indication of lake trophic state, measurable for example as chlorophyll *a* concentration, and respond quickly and predictably to changes in the nutrient status of the lake system. Relatively standard methods exist for evaluation of functional and non-taxonomic structural (biomass, chlorophyll measurements) characteristics of algal communities.

4. Sampling is easy, inexpensive, and creates minimal impact to resident biota. Laboratory chlorophyll *a* analysis is quick and cheap.

5. Algal assemblages are sensitive to some pollutants which may not visibly affect other aquatic assemblages, or may only affect other organisms at higher concentrations (e.g. herbicides).

6. Changes in community composition can provide finer-scale assessment of changes due to ecological impacts.

Because of these features, phytoplankton was included in the Water Framework Directive 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. All these parameters are considered to undergo degradation along the pressure gradient and the extent of this degradation can be “translated” into WFD normative definitions. Phytoplankton is widely recognised as being especially suitable for detecting eutrophication (Carvalho et al., 2006). Phytoplankton community composition and diversity are, however, regulated by a complex interplay of intrinsic and extrinsic drivers such as climate, resource availability, patterns of competition and predation, and dispersal (Reynolds 2006) but they may also act as sensitive indicators of environmental pressures such as changes in nutrient status (Kuemmerlin 1998, Padisák & Reynolds 1998). 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. A number of metrics have been proposed as part of WISER (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 & 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 of the overall assessment of a waterbody (Clarke et al. 2006). This has led to suggestions that results of ecological status classification should be given in terms of probabilities (Hering et al. 2010).

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 & Hering 2006) but there is a need for new data, collected according to a sampling design that allows distinction of different and independent sources of variability in metric scores. Herein, we present the results of a novel analysis of seven established phytoplankton community metrics based on a pan-European field sampling from 32 lakes. Rigorous standardisation of sampling and sample processing procedures, along with a hierarchic 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 1) quantify among-lake and within-lake variation in phytoplankton community metrics, as well as the effect of variation in sample analysts and, 2) examine variation in phytoplankton metrics due to physical, chemical and geographical attributes of lakes, with a specific focus on total phosphorus concentration as an indicator of eutrophication pressure.

2 Methods

2.1 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 3.1). These collectively represented lake types found within Member States comprising the Alpine, Northern, Central/Baltic and Mediterranean Geographical Intercalibration Groups (GIGs) identified within the WFD. All lakes were less than 10 km^2 in surface area and had mean depths ranging between 3.5 and 34 m. Northern lakes were generally low alkalinity (<0.2 m equiv. L^{-1}), while Central/Baltic and Mediterranean lakes were high alkalinity (>1 m equiv. L^{-1}).

Table 3.1. 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.

Lake	Country	GIG	Latitude (°N)	Longitude (°W)	Mean depth (m)	Maximum depth (m)
Nordborgsø	Denmark	CB	55.06	9.76	5.0	8.5
Fussingsø	Denmark	CB	56.47	9.88	12.6	31.0
Saadjärv	Estonia	CB	58.54	26.65	8.0	21.7
Viljandi	Estonia	CB	58.35	25.60	5.5	9.5
Sääksjärvi	Finland	N	62.17	25.73	9.3	15.2
Vuojärvi	Finland	N	62.41	25.94	4.4	10.2
Iso-Jurvo	Finland	N	62.60	25.93	8.6	29.6
Salagou	France	M	43.66	3.40	15.6	49.3
Caramany	France	M	42.74	2.59	14.5	36.0
Glindower See	Germany	CB	52.36	12.92	4.9	14.3
Grienericksee	Germany	CB	53.10	12.89	4.7	11.5
Roofensee	Germany	CB	53.11	13.02	9.0	19.1
Alserio	Italy	AL	45.78	9.21	5.0	8.0
Bidighinzu	Italy	M	40.56	8.66	7.5	21.8
Candia	Italy	AL	45.33	7.92	5.0	7.5
Monate	Italy	AL	45.80	8.66	18.0	34.0
Segrino	Italy	AL	45.83	9.27	3.5	8.0
Nøklevann	Norway	N	59.88	10.88	19.0	31.0
Longumvatnet	Norway	N	58.49	8.76	14.0	35.5
Temse	Norway	N	58.38	8.64	6.0	10.2
Rumian	Poland	CB	53.38	20.00	6.0	14.0
Lidzbarskie	Poland	CB	53.26	19.80	10.0	24.0
Kielpinskie	Poland	CB	53.35	19.79	5.8	10.0
Vencías, Las	Spain	M	41.43	-3.96	8.0	14.8
Vega de Jabalón	Spain	M	38.76	-3.79	6.6	10.8
Arquillo de San Blás	Spain	M	40.36	-1.21	34.0	38.0
Fiolen*	Sweden	N	57.08	14.53	3.8	10.0
Skirösjön*	Sweden	N	57.36	15.38	5.2	8.0
Västra Solsjön*	Sweden	N	59.08	12.29	12.3	40.0
Loweswater	UK	N	54.58	-3.36	8.0	14.8
Grasmere	UK	N	54.45	-3.02	8.4	19.4
Rostherne mere	UK	CB	53.35	-2.39	11.5	29.7

*Only chlorophyll data included

Lakes were all sampled according to the same rigorously standardised protocol. Sampling was conducted according to a hierarchic design that allowed the total variability in chlorophyll *a* concentration (Chl) and phytoplankton community structure, as indicated by a range of metrics, to be decomposed into a series of different components, each indicating a potential source of uncertainty. The sampling design was as follows (Fig. 2.1):

- 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.
- 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.
- 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. 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. This of course is confounded with sub-sample variation in what follows. For Chl 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 replicate to evaluate the sub-sampling error .

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 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.

Water samples were collected from 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). Secchi depth was also recorded at the deepest point of each lake.

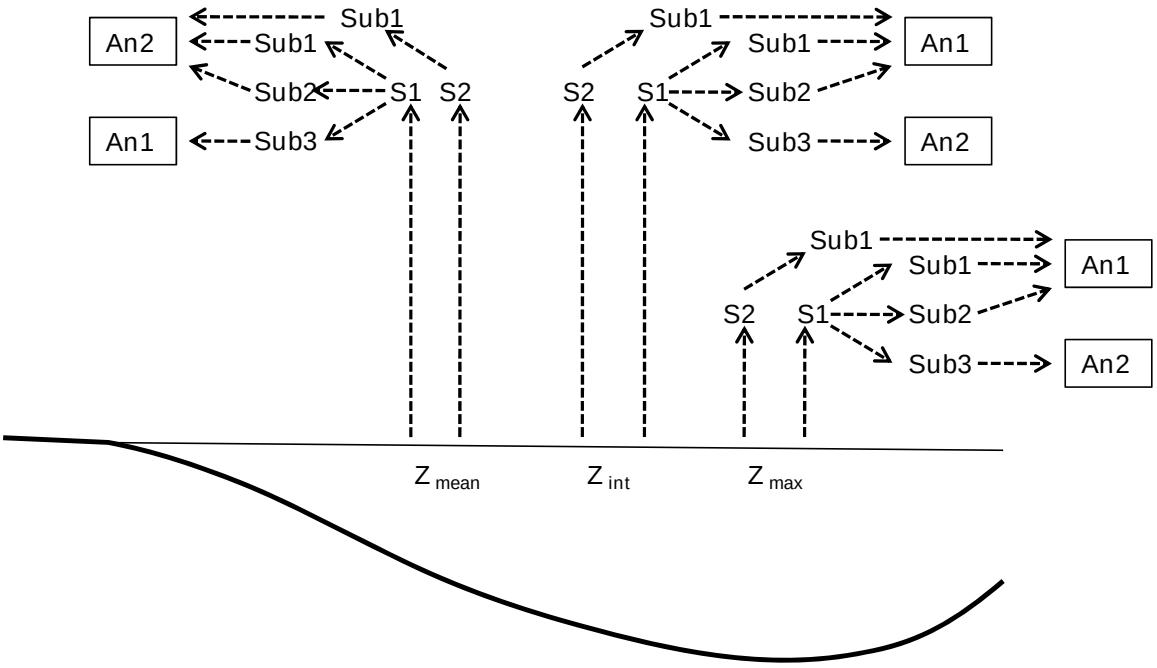

Fig. 2.1. The hierarchic sampling design employed in each lake. Samples were collected from three stations, above the deepest point (z_{max}), the mean depth (z_{mean}) and a depth intermediate between the maximum and mean depths (z_{int}). Two samples (S1, S2) were collected at each station. At each 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).

2.2 Sample processing

2.2.1 Chlorophyll ***a***

Avoiding direct sunlight, samples were immediately filtered through 47-mm GF/F filters using vacuum of less than 100 mm Hg (~0.15 atm). The volume of water filtered was dependent on the amount and type of seston present in the water. A fixed volume was kept for one lake.

The moist filter was placed immediately face inside in a 15-ml centrifuge tube (with a cap) prefilled with 10 ml of 96% ethanol for pigment extraction. The 47 mm GF/F filters retain 0.8 ml of water adjusting the final extraction solution to 89% ethanol and the final extraction volume to 10.8 ml (Knap et al., 1996). The extraction lasted 24 hours in a fridge or cooling bag at 4 °C. No folding, drying, freezing, grinding or sonication of the filter was applied to avoid introduction of additional variation of results.

Following extraction, samples were shaken, filters were pressed to the bottom of the tube with a glass stick and spun down in a centrifuge for 5 minutes to remove filter and cell debris.

The spectrophotometer was allowed to warm up and stabilize prior to use according to the instrument manual. The extract was decantated from the centrifuge tube to 1-cm cuvette. The external surfaces of the cuvette were wiped clean with a soft tissue. The spectrophotometer was zeroed with 90% ethanol.

Readings of spectrophotometric absorption (A) were taken at following wavelengths:

1. Peak in the range of Chl *a* absorption maximum 662-665 nm (A_{peak})
2. 750 nm (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 unacidified 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 chlorophyll-*a* was calculated according to equation given in ISO 10260 (1992):

Chl *a* $[mg/m^3] = 29.6 * (E_{peak} - acE_{peak}) * a/(L * V)$

Where a = final extraction volume (ml)

V = volume of water filtered (L)

L = length of the light path through the cuvette (cm)

2.2.2 Microscopy

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 Brierley *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 were identified to the highest possible taxonomic level. Counts of each taxon were converted to biovolumes by measuring cell/colony dimensions and approximating each taxon to a simple geometric shape (Brierley 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. All data entry and biovolume calculations were carried out using standard WISER data entry sheets and collated on a common master database on the WISER intranet site.

2.3 Phytoplankton metrics

Seven phytoplankton metrics are considered herein, a brief description of which is given below. Full details on each metric are provided in WISER Deliverable D3.1-1: *Report on phytoplankton composition metrics, including a common metric approach for use in intercalibration by all GIGs* (Phillips et al. 2010) and WISER Deliverable D3.1-2: *Report on phytoplankton bloom metrics* (Mischke et al 2010). Metric values were calculated from the WISER field campaign database using the algorithms described below.

2.3.1 Abundance metrics

- **Chlorophyll *a* concentration (Chl).** Chlorophyll *a* concentration (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.

2.3.2 Composition metrics

- **Phytoplankton Trophic Index (PTI).** This is developed 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.
- **Size Phytoplankton Index (SPI)**. 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$0.5µm^3, 0.5-1.0 µm^3, 1.0-2.0 µm^3, 2.0-4.0 µm^3 etc. 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 trophic position of a size class along the trophic spectrum and indicator values estimate the "power" of each size class as a biotic indicator. Analysis of data from the WISER lakes database has shown that 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).
- **Morpho-Functional Group Index (MFGI)**. 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 & 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). Analysis of data from the WISER lakes database has shown that the MFGI tends to increase with increasing lake trophic state, due to an increase in the dominance of colonial cyanobacteria, large diatoms/chlorophytes/coniugatophytes, and unicellular/colonial chlorococcales (Phillips et al. 2010).
- **Functional Traits Index (FTI)**. 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.

2.3.3 Bloom metrics

- **Evenness 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 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).
- **Cyanobacterial abundance**. 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).

2.4 Statistical modelling

Linear mixed effects models (LME) were used to analyse metric scores based upon count data from samples from 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 Chl) as a crossed random effect, representing the fact that it does not naturally nest within the sample hierarchy. As a first step, a null model (i.e. including only the random effects) was run for each metric. The estimated variance parameters for each level in the random effects hierarchy were extracted in order to compare metric variation among lakes with that among samples and sub-samples/analysts. When examining these random effects structures in the resulting models, to determine patterns of among and within-lake metric variability, restricted/residual maximum likelihood (REML) estimation was used during model fitting. This gives unbiased estimates of the random effects.

In order to investigate the extent to which metrics varied with physical and chemical features of the lakes studied, models were then re-run to include certain measured environmental variables as fixed effects. The precise environmental variables measured, and the number of determinations of each, were variable among data sets from different data providers. Metric values were linked to values of environmental data collected at the same location and time. For environmental data, measurements collected from the first sample at the deep station were used in all lakes. Where there were multiple determinations of a variable within this sample, these were averaged. Unfortunately, only total phosphorus was recorded in all lakes during the phytoplankton field campaign. Alkalinity data were missing for some lakes and so representative values were necessarily derived from the data set collected under the macrophyte uncertainty exercise (Dudley 2010). Once a complete data set for total phosphorus and alkalinity was attained, models were fitted with different combinations of these explanatory variables. As proxies for broad climatic gradients that might impact upon phytoplankton communities, latitude, longitude and altitude were also considered as fixed effects. Secchi depth was omitted since the direction of causality between this variable and the phytoplankton community is equivocal. Maximum likelihood (ML) estimation was used during model selection, and models

were compared by using the Akaike Information Criterion (AIC) to find the model with the most optimal combination of environmental predictor variables i.e. lowest AIC value (Burnham & Anderson 2002, Zuur et al 2009).

Models with each of the explanatory variables as the sole fixed effect were fitted first, in order to find the variable which was the "best" single predictor of the modelled metric. After this, additional explanatory variables were included and those that caused the greatest reduction in the AIC were retained. During the model fitting exercise it was necessary to simplify the random effects structure by omitting the "Country" random effect. Preliminary analyses revealed that the inclusion of this effect resulted in convergence errors, due to a high level of model complexity. All analyses except for Chl were conducted using the base and lme4 packages of R version 2.12.1 (Pinheiro & Bates 2000, R Development Core Team 2009). The partitioning of Chl variance among the different levels of sampling design was analysed using the Variance Estimation and Precision (VEPAC) package of STATISTICA 8.0 (StatSoft. Inc. 1984-2007). Similarly to R, REML estimation was used during model fitting and the AIC criterion was used to determine the strength of single predictors among environmental variables and select the best model when multiple fixed effects were included.

3 Results

3.1 Data exploration: metrics

Frequency histograms of the metrics studied (Fig. 3.1) showed that the chlorophyll *a* and cyanobacterial biovolume metric were positively skewed, while the evenness metric was negatively skewed. This skewness was most pronounced in the case of the chlorophyll *a* and cyanobacterial biovolume metrics 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. This improved the spread of the data, however, the transformed dataset for cyanobacterial abundances was still dominated by large numbers of low values. Evenness values, though also skewed, were not transformed prior to model fitting. Due to these skews, residuals from models based upon all metrics were examined for evidence of departure from a Normal distribution. It is the residuals, rather than the original data, that are required to be Normally distributed.

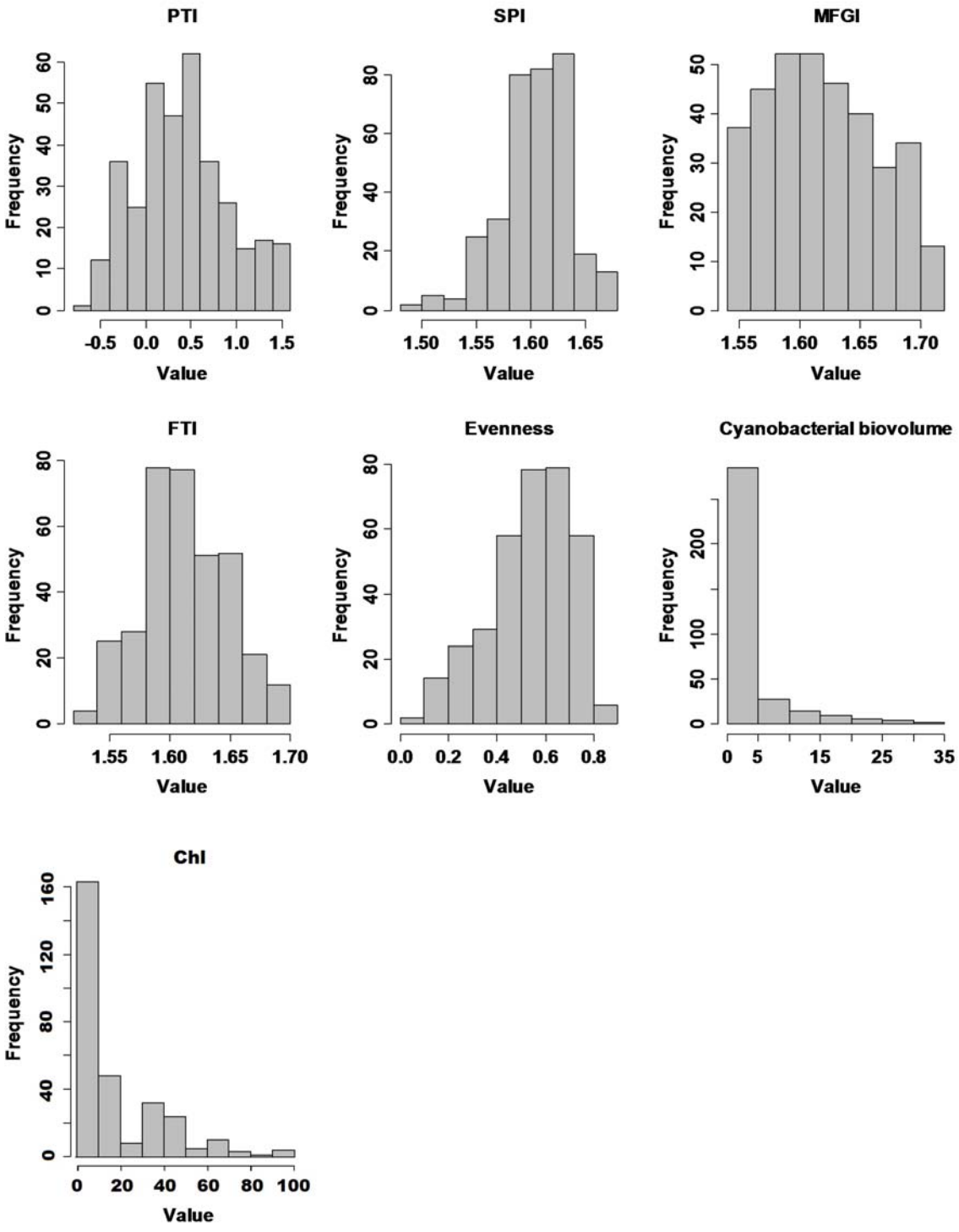

Fig. 3.1. Frequency histograms of the seven phytoplankton metrics studied. PTI = Phytoplankton Trophic Index, MFGI = Morpho-Functional Group Index, FTI = Functional Traits Index, Chl = Chlorophyll a concentration (spectrophotometric).

There are clear geographical variations in both average metric values, and their within-country variability (Fig. 3.2). However, patterns of variability among countries differ between metrics.

One commonality is that Norwegian lakes have comparatively low median values of the PTI, SPI, MFGI and FTI metrics. German and Danish lakes also tend to demonstrate low median scores for these metrics (excepting PTI), as do Finnish lakes (excepting SPI). In contrast, French lakes are perhaps most exceptional in terms of median chlorophyll *a* concentration and total cyanobacterial biovolume. Median evenness was lowest for the German and Danish lakes. Interestingly, within-country metric variability is itself very different when comparing the different countries. Higher levels of within-country variability can be seen for Germany, Spain and Finland for all metrics but evenness. These differences in within-country variation may occur 1) due to differences in the trophic gradient represented by the lakes sampled within each country or 2) by chance due to the small sample size (number of lakes) per country.

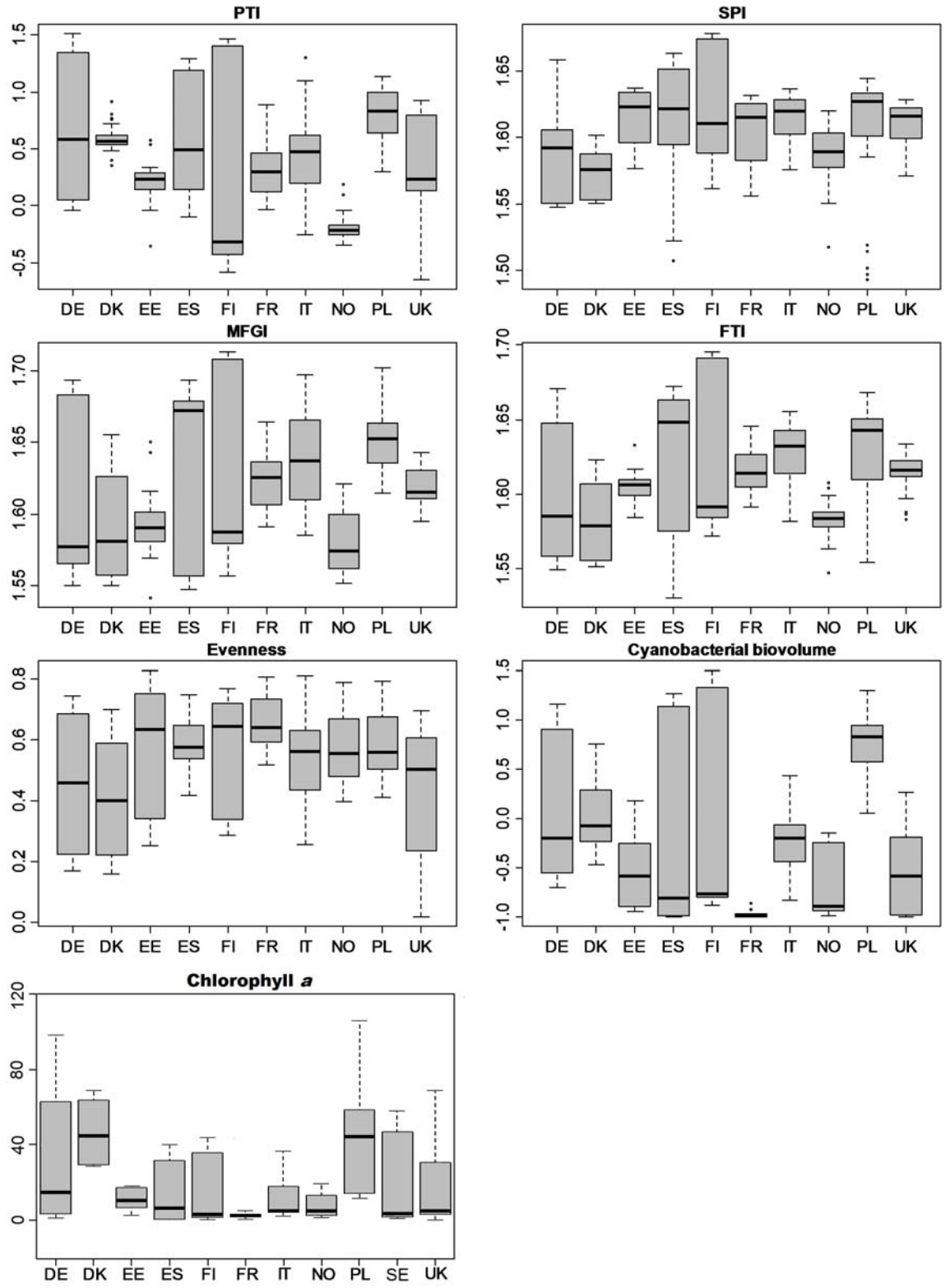

Fig. 3.2. Boxplots of metric scores, categorised by country. Bold horizontal lines indicate median values. Metrics are PTI = Phytoplankton Trophic Index, MFGI = Morpho-Functional Group Index, FTI = Functional Traits Index. Countries: DE = Germany, DK = Denmark, EE = Estonia, ES = Spain, FI = Finland, FR = France, IT = Italy, NO = Norway, PL = Poland, SE = Sweden, UK = United Kingdom.

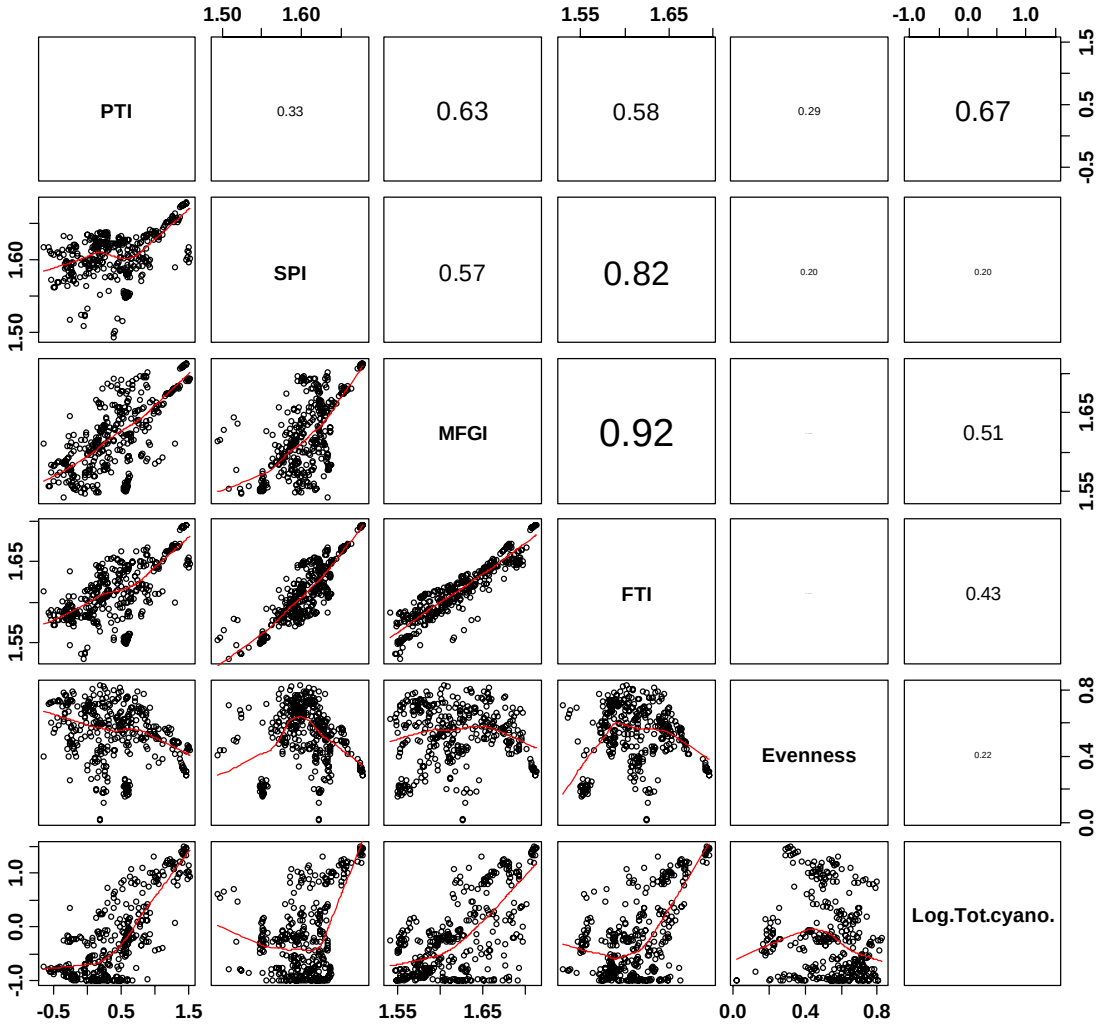

Fig. 3.3. Spearman rank correlation matrix of metric scores at the sub-sample scale. Below diagonal – scatterplots with smoothers. Above diagonal – Spearman rank correlation coefficients.

As a next step, the different metrics were correlated with each other in order to assess the extent to which they yielded similar "information" on variations in phytoplankton community structure. This was done using the original data, resolved at sub-sample level (Fig. 3.3), and also on lake-level averages of the different metrics (Fig. 3.4). As data for chlorophyll *a* was analysed separately, this metric was not included the figures. All correlations were based upon Spearman rank coefficients due to the non-Normal distributions of at least some metrics. Patterns of among metric correlation were essentially similar at both the sub-sample and lake scales. As would be expected due to its method of calculation, the FTI index was very strongly correlated with both SPI and MFGI. The PTI index was positively correlated with MFGI, FTI, log total cyanobacterial biovolume and SPI, though only weakly in the latter case. Log cyanobacterial biovolume was also weakly correlated with MFGI and FTI, while evenness was poorly correlated with any other metric. Chl a (not shown) had significant correlations with Log

cyanobacterial biovolume (Spearman r=0.744), PTI (r=0.618), and MFGI (r=0.393) whereas those with SPI and evenness were weak and non-significant.

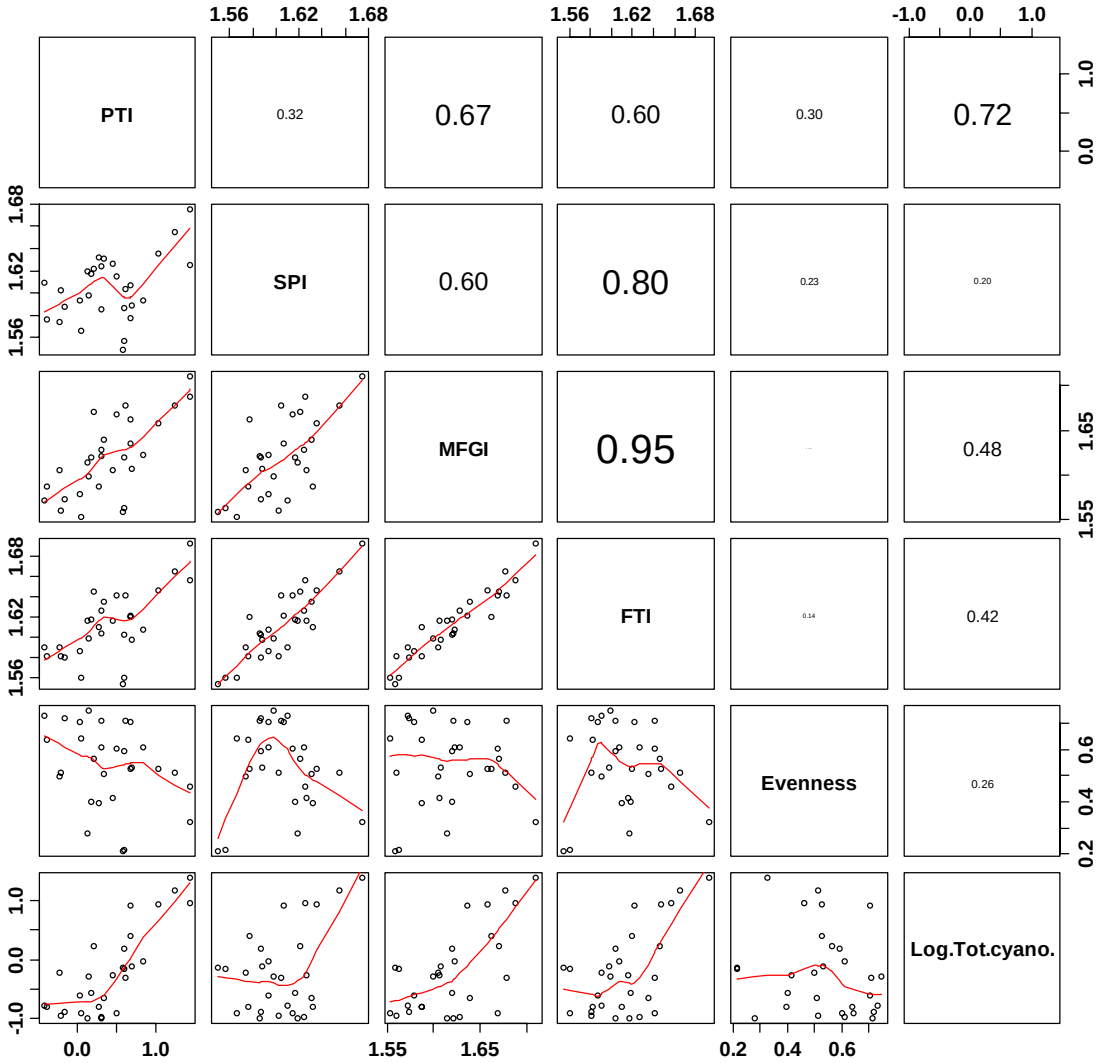

Fig. 3.4. Spearman rank correlation matrix of metric scores at the lake-averaged scale. Below diagonal – scatterplots with smoothers. Above diagonal – Spearman rank correlation coefficients.

3.2 Data exploration: environmental data

Data on total phosphorus indicated that the 29 lakes (32 lakes for Chl analysis) spanned a wide eutrophication pressure gradient (Fig. 3.5). However, lakes were not evenly dispersed across this gradient with many lakes yielding low phosphorus concentrations during the field campaign and only a few high concentrations. Lakes fell within two major groups with respect to alkalinity, with most falling within either the 0-1 meq/l or 2-3 meq/l range. The majority of lakes were found at comparatively low altitude but were more evenly dispersed with respect to latitude and longitude. Due to the positive skew in the total phosphorus, mean depth, and altitude data, we

log transformed these variables prior to further analysis in order to avoid the minority of high total phosphorus/altitude lakes having an excessive influence on model fitting.

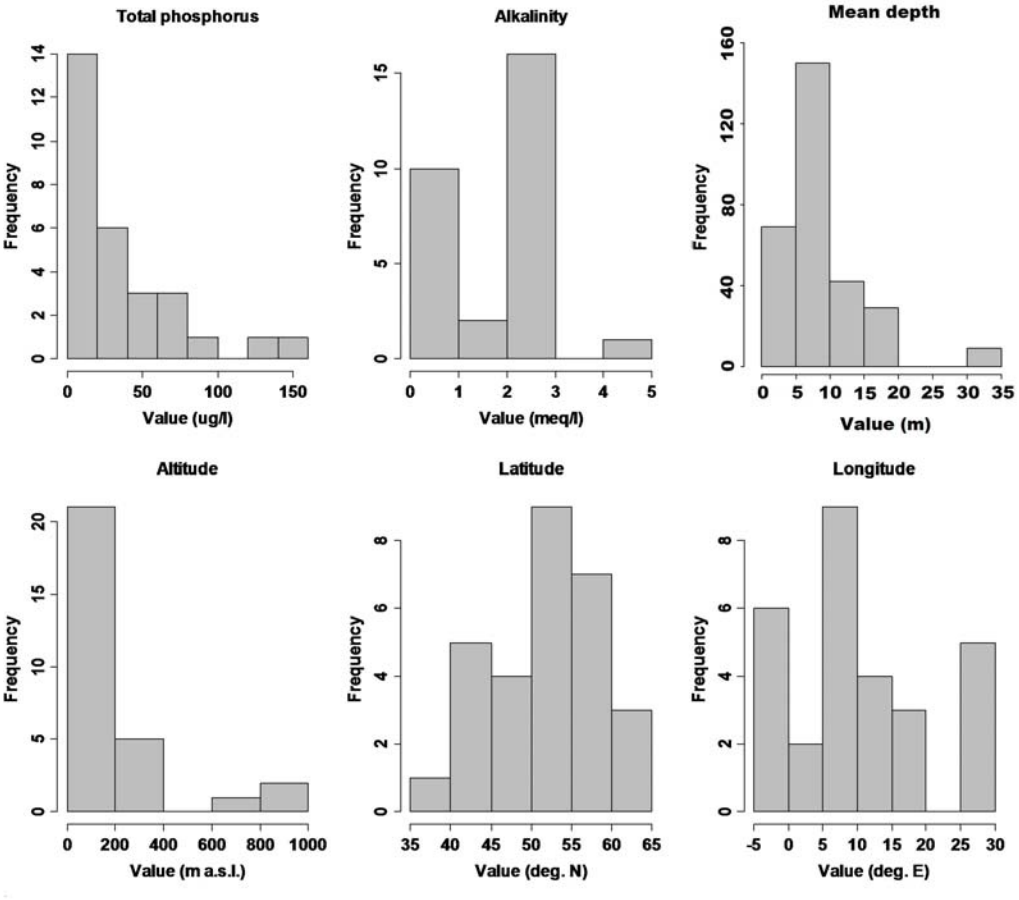

Fig. 3.5. Frequency histograms of the six environmental variables.

Relationships between predictor variables were examined in order to determine whether they were strongly inter-related; an issue which could complicate subsequent model fitting (Fig. 3.6). This was again done using Spearman rank correlation, due to the non-Normal distributions of the data (Fig. 3.5). Some correlations were apparent but were not judged sufficiently strong to omit any of the variables from subsequent analyses. Alkalinity increased non-linearly with log total phosphorus concentration. The mean depth of lakes increased towards lower longitudes and deeper lakes tended to have lower phosphorus concentrations. Geographic variables were also modestly inter-related, with higher altitude lakes being found typically at low latitude (the Spanish lakes, all with altitudes in excess of 600m a.s.l.). Latitude and longitude were also correlated since lakes were distributed along a SW-NE axis.

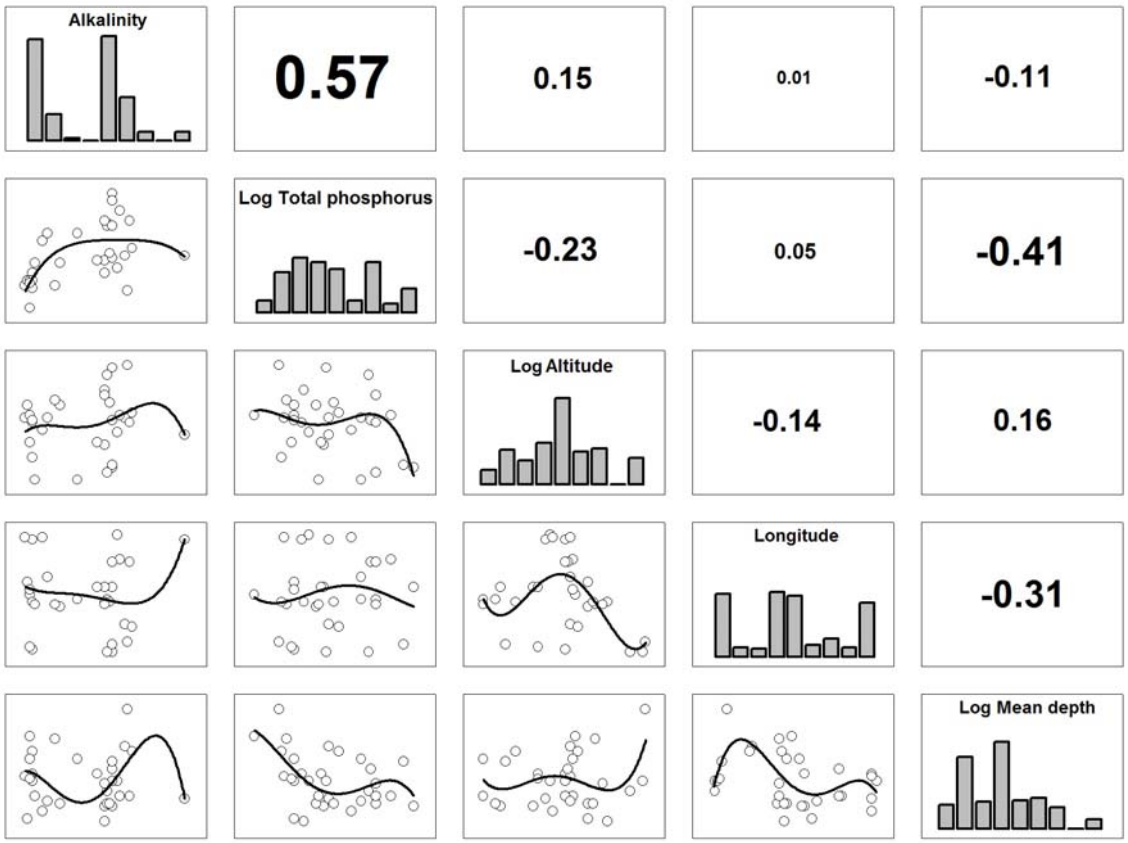

Fig. 3.6. Spearman rank correlation matrix of lake-level environmental variables. Diagonal – frequency distributions of the variables. Below diagonal – scatterplots with polynomial smoothers. Above diagonal – Spearman rank correlation coefficients.

3.3 Mixed-effects modelling

Results from null models of all seven metrics (Table 3.2) suggest that the majority of metric variance occurred between waterbodies. The Country and Waterbody random effects accounted for between 65% and 96% of the total metric variance. It is noteworthy that the Analyst and Error (sub-sample level) variance components were the major contributors to the within-waterbody component, suggesting that analyst differences and sub-sampling errors may account for more variation in the phytoplankton composition metrics than differences between samples and stations within a waterbody.

Table 3.2. Proportions of metric variance at different levels in the sampling hierarchy, for null models of the six different metrics. Total between = Country + Waterbody, Total within = Station + Sample + Analyst + Error(sub-sample).

Metric	Country	Waterbody	Station	Sample	Analyst	Error (sub-sample)	Total within	Total between
Chl	0.00	0.96	0.01	0.01	-	0.02	0.04	0.96
PTI	0.00	0.88	<0.01	0.00	0.04	0.07	0.12	0.88
SPI	0.00	0.65	0.03	0.00	0.19	0.13	0.35	0.65
MFGI	0.00	0.86	0.02	<0.01	0.05	0.08	0.14	0.86
FTI	0.00	0.81	0.02	0.00	0.09	0.08	0.19	0.81
Evenness	0.00	0.69	0.04	0.00	0.17	0.10	0.31	0.69
Log total cyanobacteria	0.09	0.86	0.01	0.00	0.02	0.03	0.06	0.94

For subsequent analyses of metric scores with respect to environmental predictor variables, the random effects structure was simplified by removing the "Country" level of the hierarchy. For each metric, the AIC values of the null model (i.e. no predictors) and models containing different combinations of predictor variables were then compared in order to determine the best single predictor, and best combination of predictors (i.e. which predictors resulted in the model with the lowest AIC). For Chl, cyanobacterial biovolume, PTI and MFGI, total phosphorus concentration was the best single predictor of metric values (Table 3.3). All metrics increased at higher phosphorus concentrations. However, there was much scatter around these relationships (Figs. 3.7, 3.9, 3.13, 3.19). For PTI and MFGI, once the total phosphorus effect was modelled metric scores also appeared to vary systematically with altitude, while for cyanobacterial biovolume there appeared to be some effect of longitude. The effect of these geographic variables was strongest in the case of MFGI, and modest for the other two metrics (AIC comparison of TP models including and excluding altitude/longitude, Table 3.3). For SPI, FTI and evenness altitude was the best single predictor of metric scores though, in all cases, only a small reduction in AIC compared to the null model was achieved by including this variable in the model; indicative of a weak effect (Table 3.3). This can be seen from the scatter in the bivariate relationships between these variables (Figs 3.11, 3.15, 3.17). For Chl where also the effect of lake morphometry was studied, the mean depth of the lake was the second strongest predictor after TP.

Models of SPI, FTI and evenness that included predictor variables as fixed effects were only a modest "improvement" upon null models (difference in AIC between the null model and models with predictors = 0.4-1.3, Table 3.3). For the remaining metrics, differences in AIC between null models and models including predictor variables ranged from 2.8 (MFGI) to 51.7 (Chl). This would indicate the metric scores for MFGI, PTI, total Cyanobacterial biovolume, and especially Chl are better predicted by the measured environmental variables.

Comparisons of standard deviations in metric scores in null models and models with the most optimal combination of predictor variables (and therefore lowest AIC) show that among-lake variability is reduced more by the inclusion of environmental variables for Chl, total

Cyanobacteria and PTI (Figs. 3.8, 3.10) than for SPI, MFGI, FTI and evenness (Figs. 3.12, 3.14, 3.16, 3.18). For all seven phytoplankton community metrics, it is clear that there is still considerable among-lake variation in metric scores that is not explained by the environmental variables considered i.e. the non-zero standard deviations associated with the "lake" random effect in the models with fitted predictor variables (labelled "top model" in Figs. 3.8, 3.10, 3.12, 3.14, 3.16, 3.18). The fact that the "lake" level standard deviation is still substantial after inclusion of predictor variables indicates the presence of "unexplained" among-lake variability in metric scores. This unaccounted variability between lakes could, however, still include other responses to eutrophication pressures that are not represented by total phosphorus (nitrogen, turbidity, etc.). A model for Chl including TP and Alkalinity as significant positive fixed effects and Mean depth and Altitude as significant negative fixed effects, showed the greatest explanatory power, with more than 3.5 times lower variance at the lake level compared with the null model (Fig. 3.20).

Table 3.3. AIC comparison of models of the phytoplankton metrics. For each metric, a model with a fixed effect of total phosphorus concentration is compared with the corresponding null model.If inclusion of additional environmental variables as fixed effects improved model fit, then these are included. pos = positive relationship between metric and selected variable, neg = negative relationship. Models in bold are those where the difference in AIC from the null model is >2.

Metric	Predictors	AIC
Chl	Null	-347.2
	Log total phosphorus (pos)	**-385.2**
	Log total phosphorus (pos)+Log Mean depth (neg)	**-391.1**
	Log total phosphorus (pos)+Log Mean depth (neg)+Alkalinity (pos)	**-394.0**
	Log total phosphorus (pos)+Log Mean depth (neg)+Alkalinity (pos)+Log Altitude (neg)	**-397.0**
	Log total phosphorus (pos)+Log Mean depth (neg)+Alkalinity (pos)+Log Altitude (neg)+Longitude (pos)	-398.9
PTI	Null	-195.6
	Log total phosphorus (pos)	**-208.0**
	Log total phosphorus (pos)+Log Altitude (pos)	**-208.6**
SPI	Null	-1814.0
	Log Altitude (pos)	-1814.4
MFGI	Null	-1830.6
	Log total phosphorus (pos)	-1831.0
	Log total phosphorus (pos)+Log Altitude (pos)	**-1833.4**
FTI	Null	-1974.0
	Log Altitude (pos)	-1974.8
	Log Altitude (pos)+Log total phosphorus (pos)	-1974.9
Evenness	Null	-784.9
	Log Altitude (pos)	-786.2
Log total cyanobacteria	Null	-217.4
	Log total phosphorus (pos)	**-229.3**
	Log total phosphorus (pos)+Longitude (pos)	**-229.7**

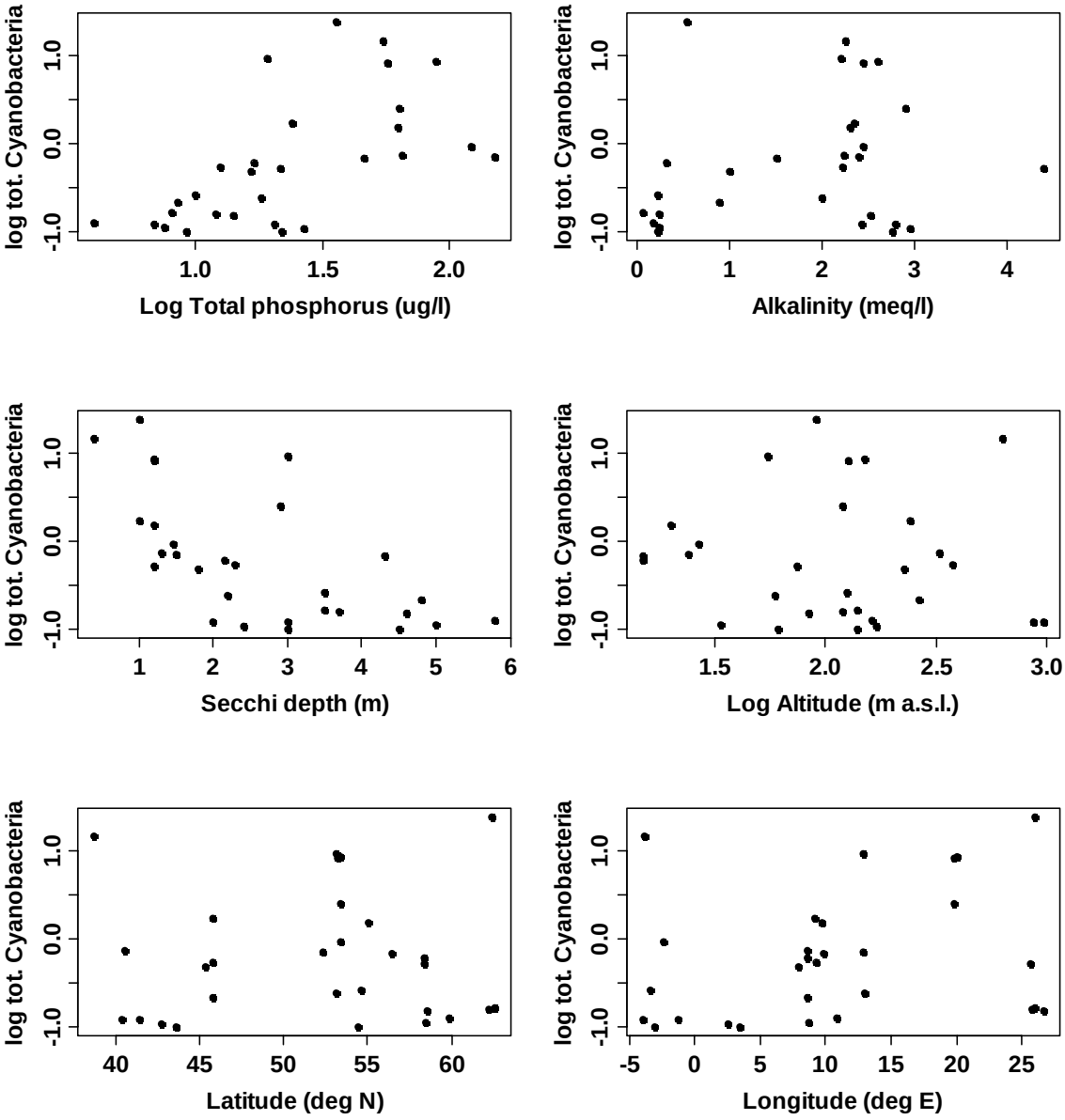

Fig. 3.7. Scatterplots of log total cyanobacterial biovolume against the predictor variables included in the analysis.

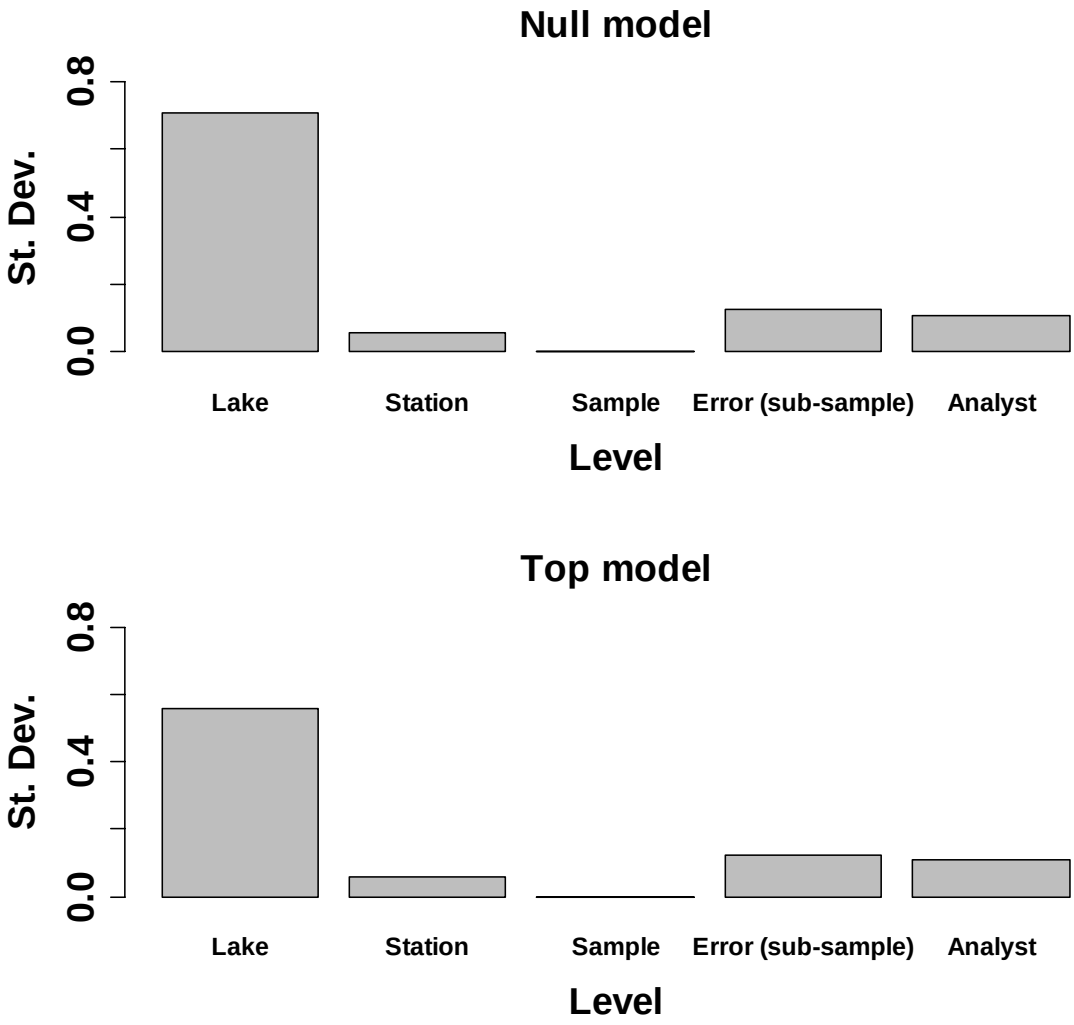

Fig. 3.8 Standard deviations of log total Cyanobacterial biovolume at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC.

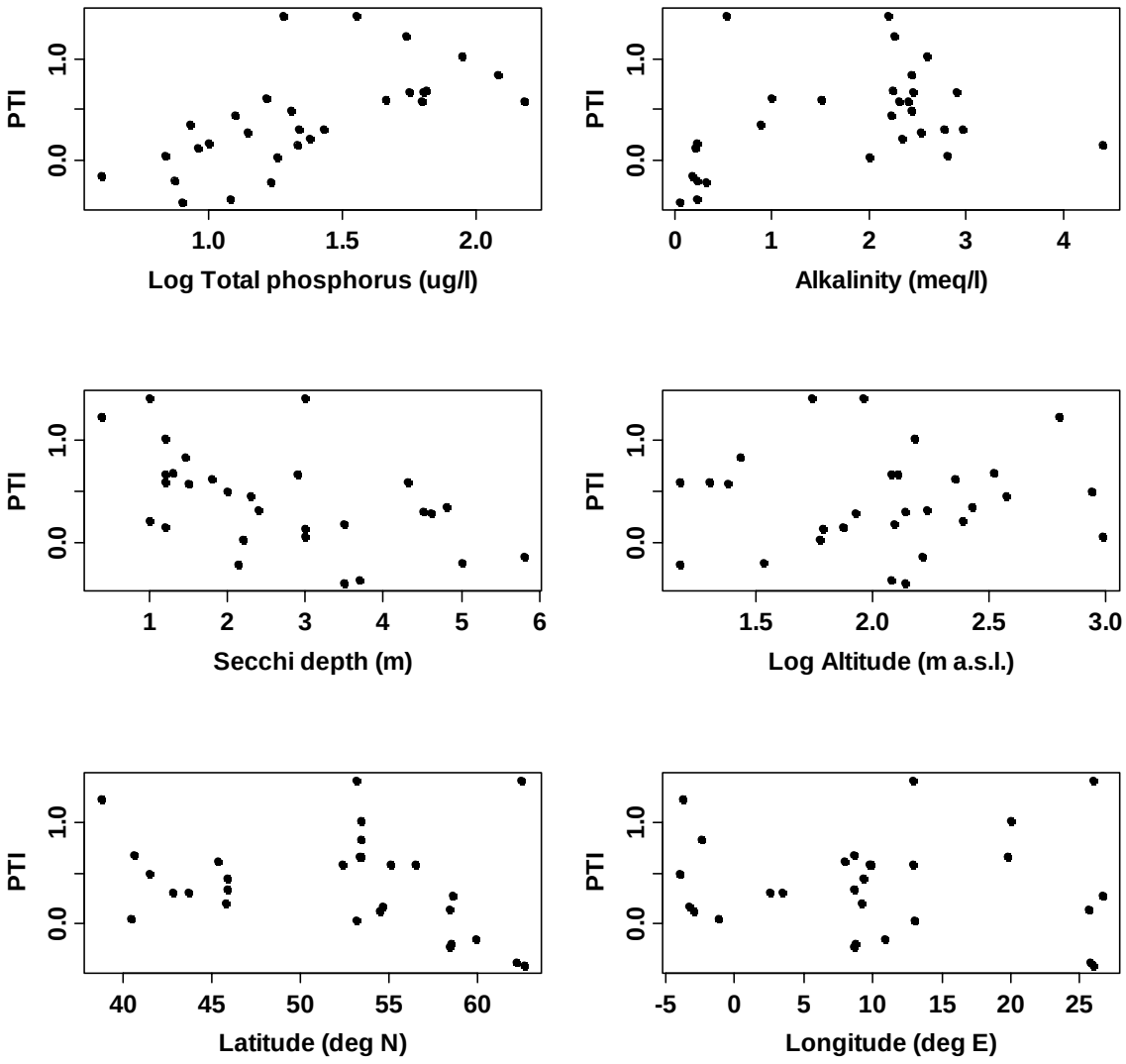

Fig. 3.9. Scatterplots of PTI against the predictor variables included in the analysis.

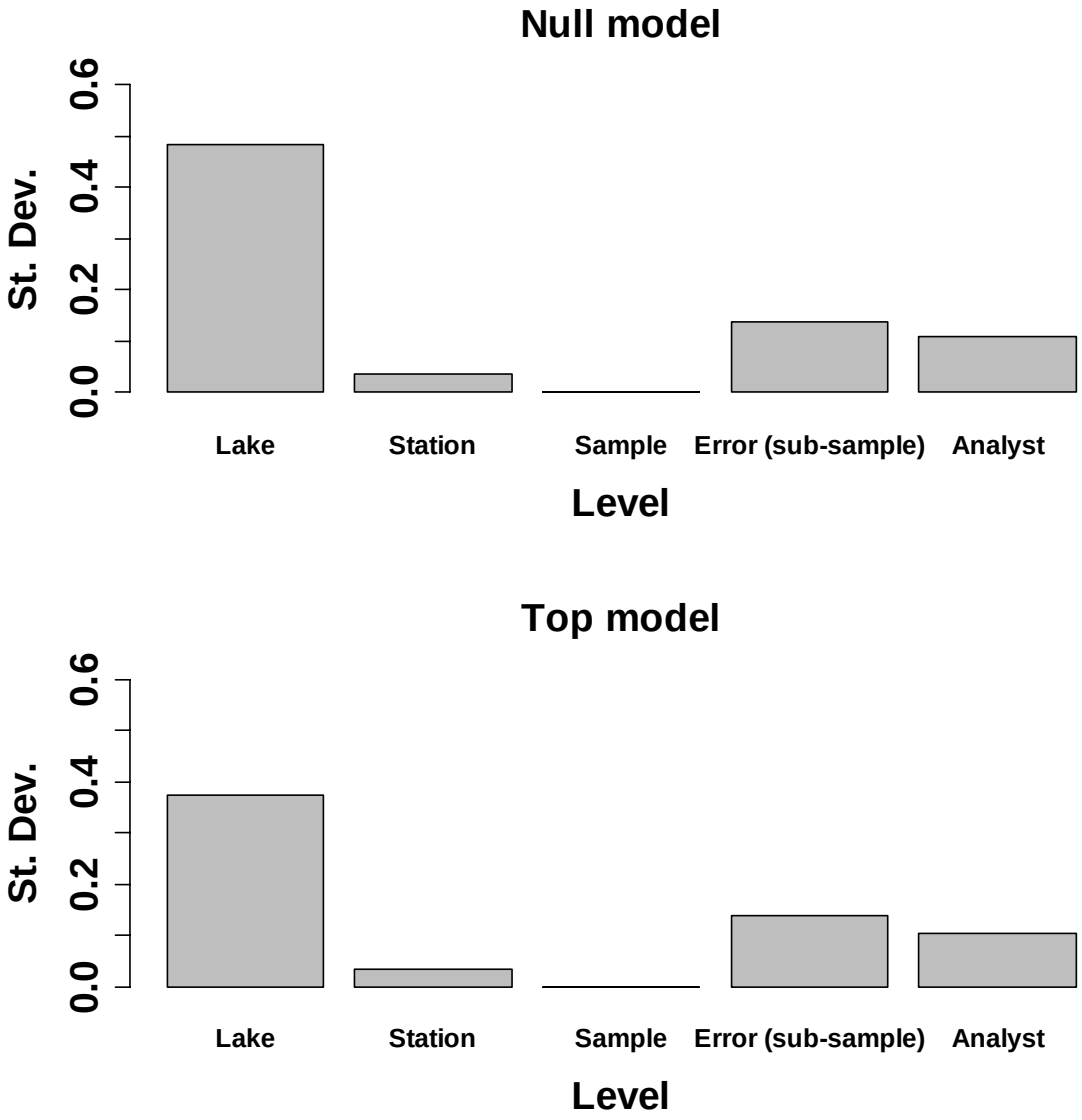

Fig. 3.10 Standard deviations of PTI metric at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC.

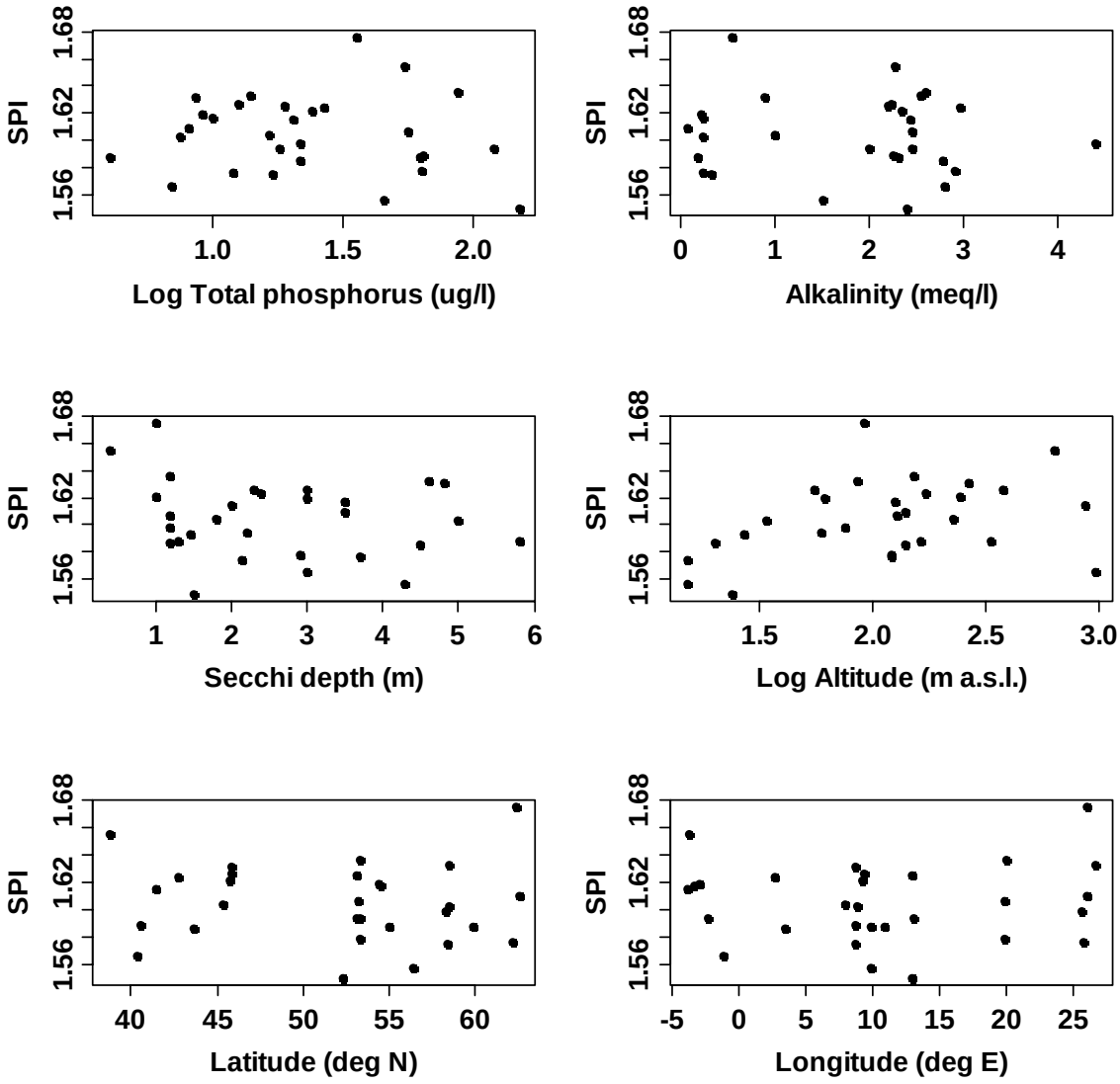

Fig. 3.11. Scatterplots of SPI against the predictor variables included in the analysis.

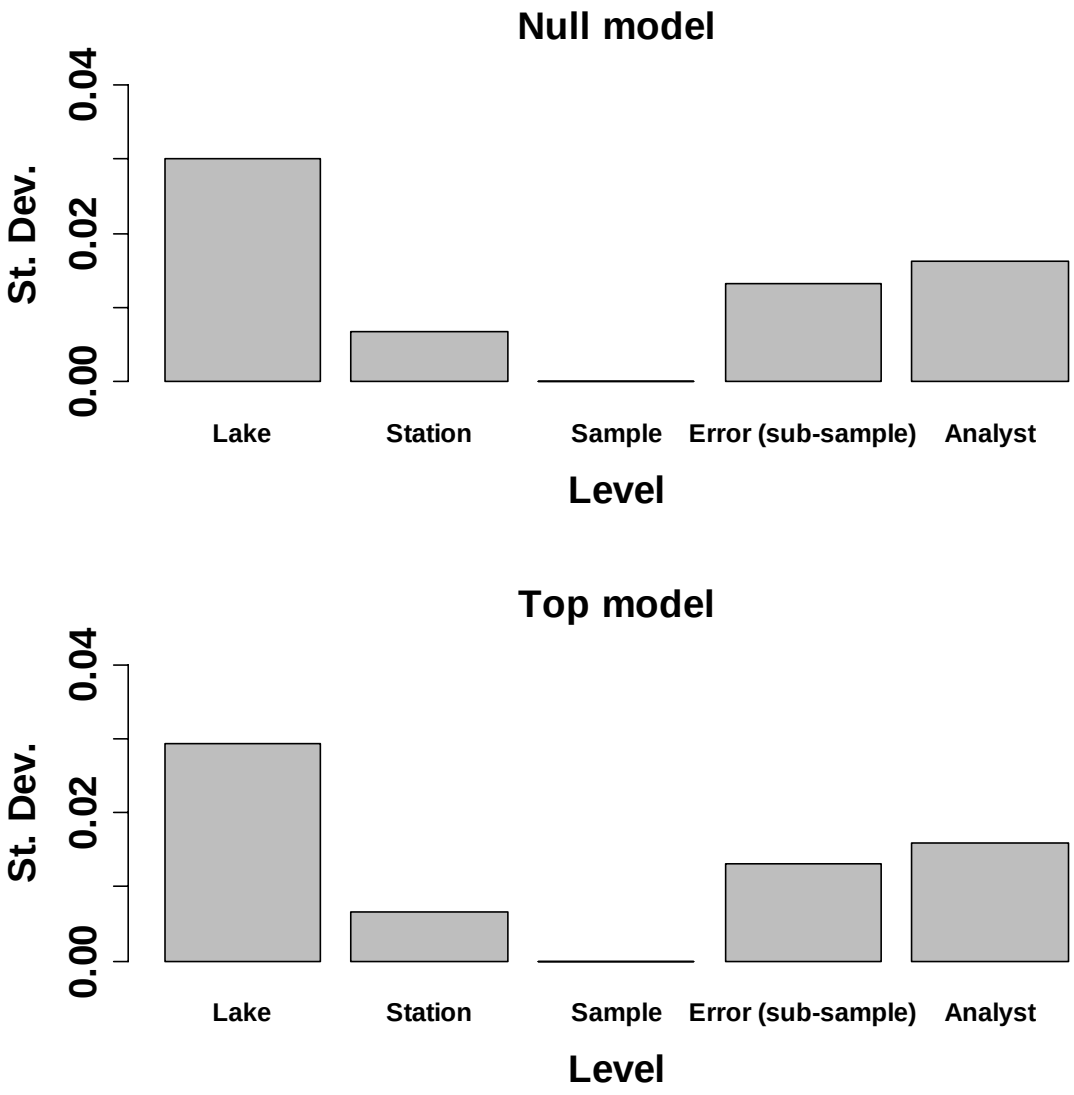

Fig. 3.12 Standard deviations of SPI metric at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC.

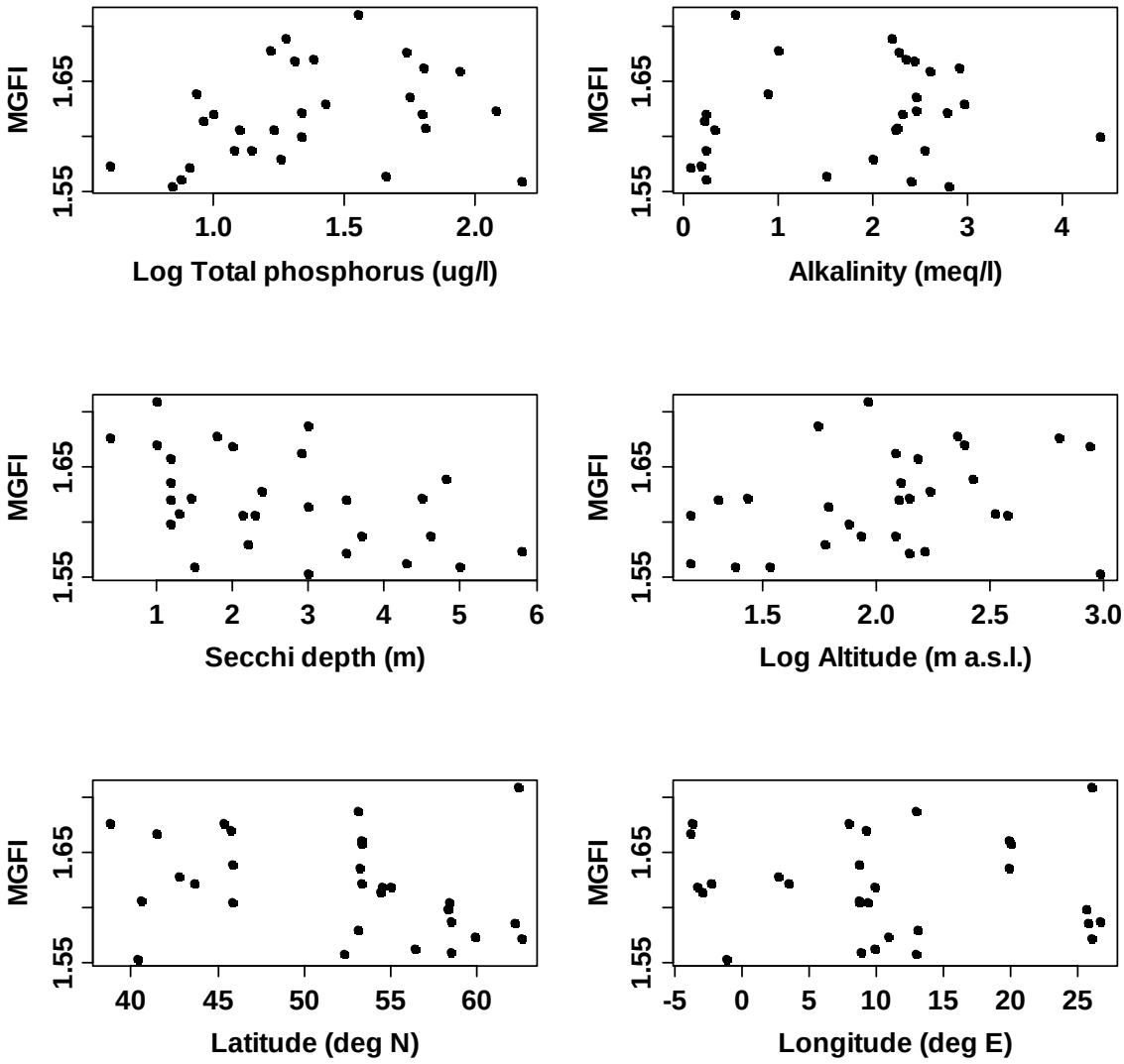

Fig. 3.13. Scatterplots of MFGI against the predictor variables included in the analysis.

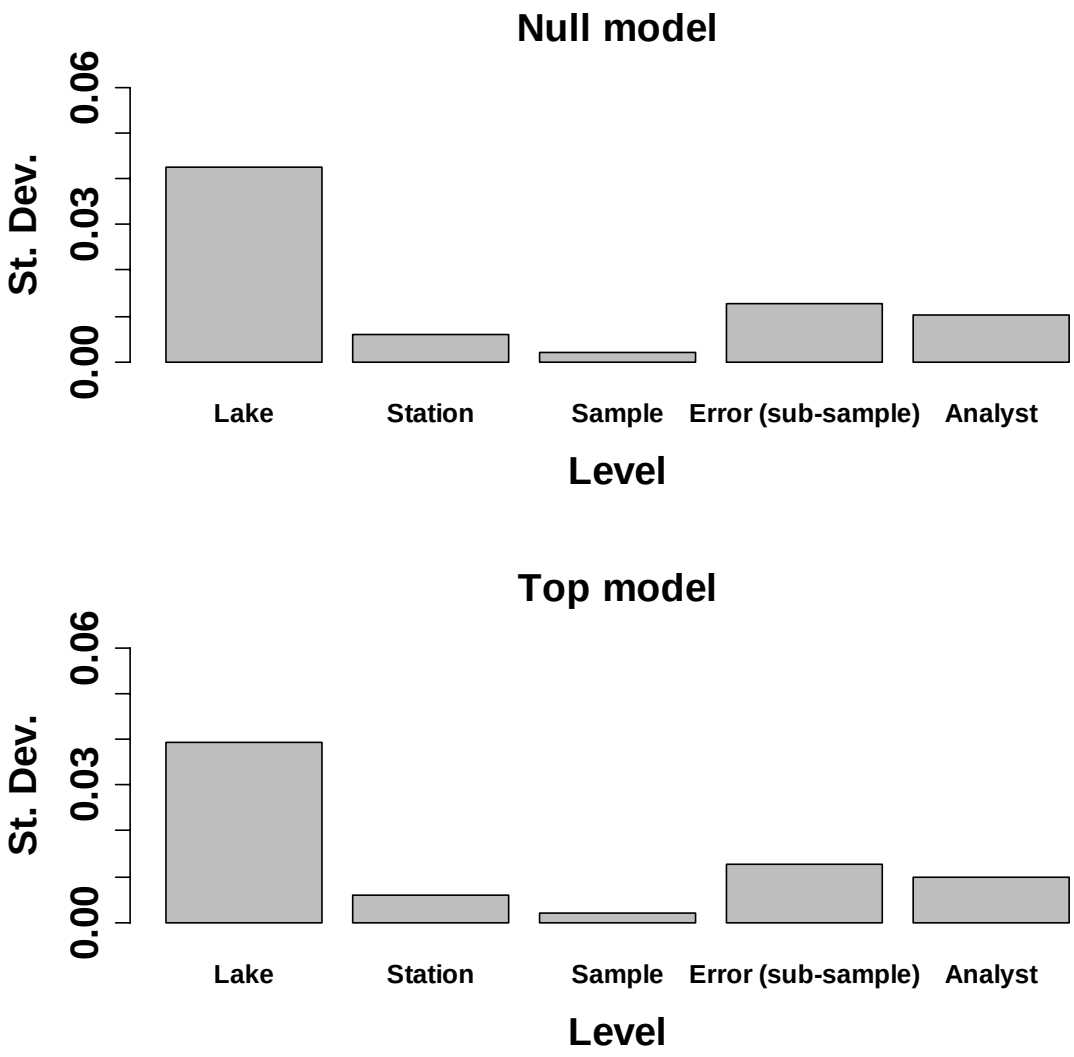

Fig. 3.14 Standard deviations of MFGI metric at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC.

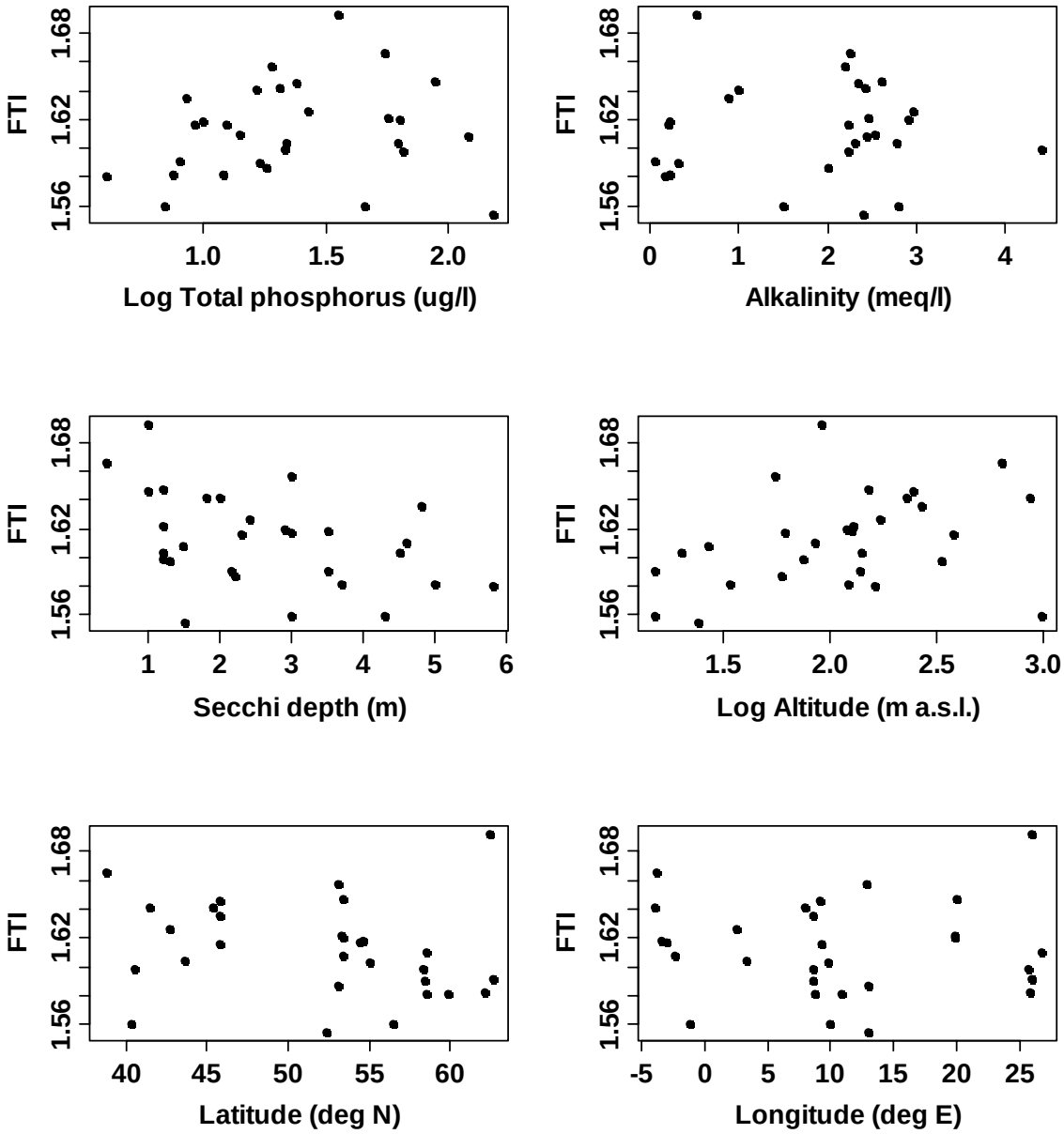

Fig. 3.15. Scatterplots of FTI against the predictor variables included in the analysis.

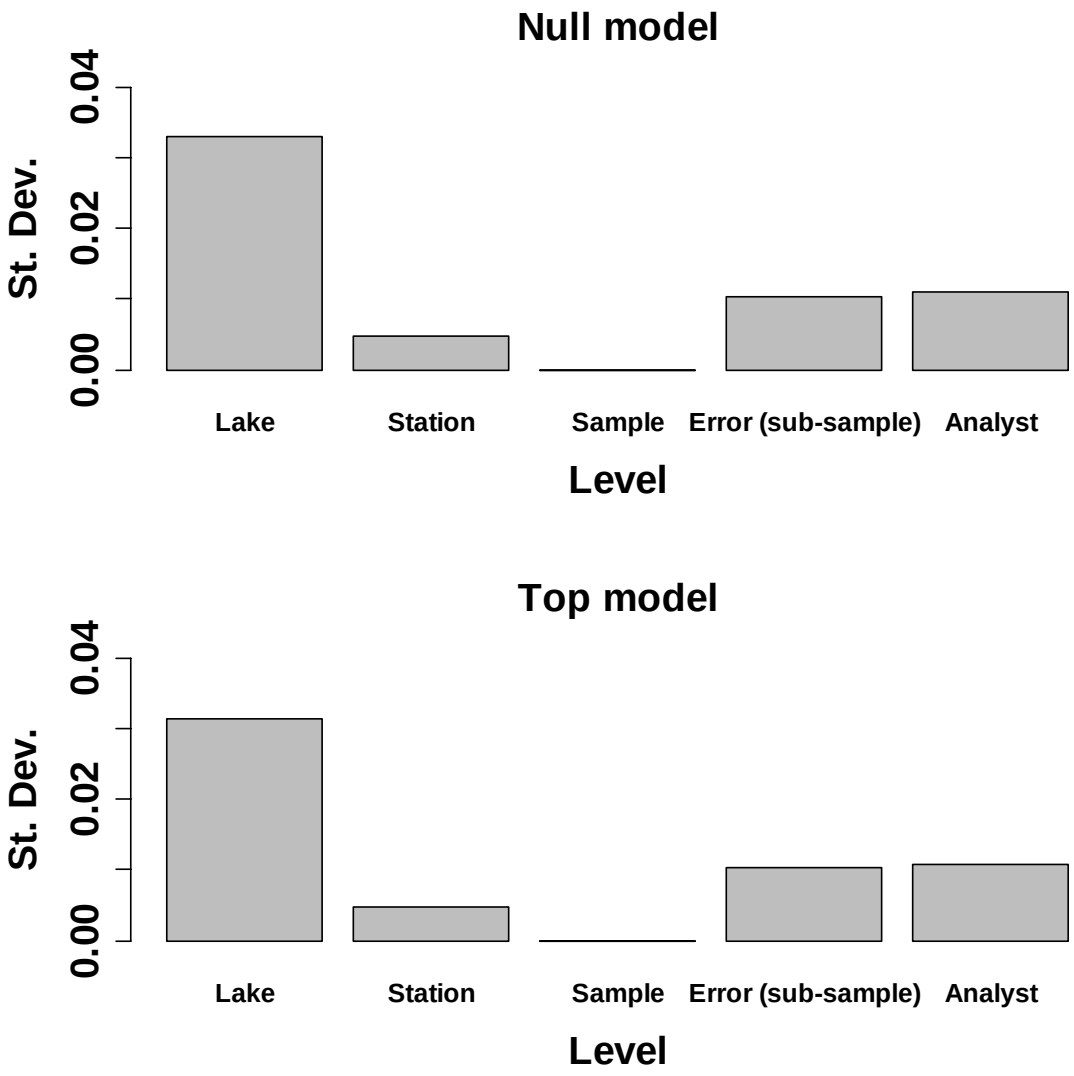

Fig. 3.16 Standard deviations of FTI metric at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC

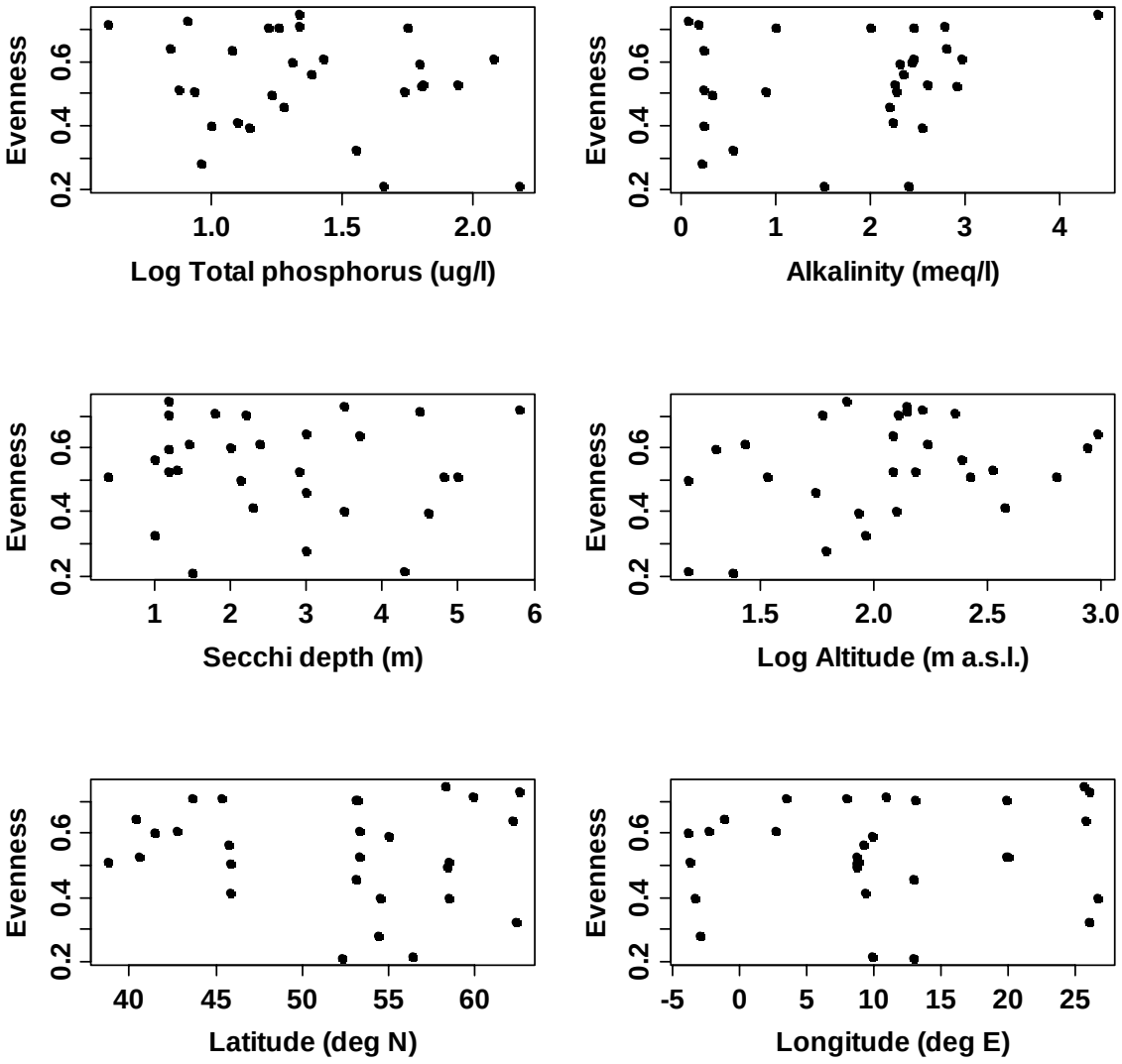

Fig. 3.17. Scatterplots of Evenness against the predictor variables included in the analysis.

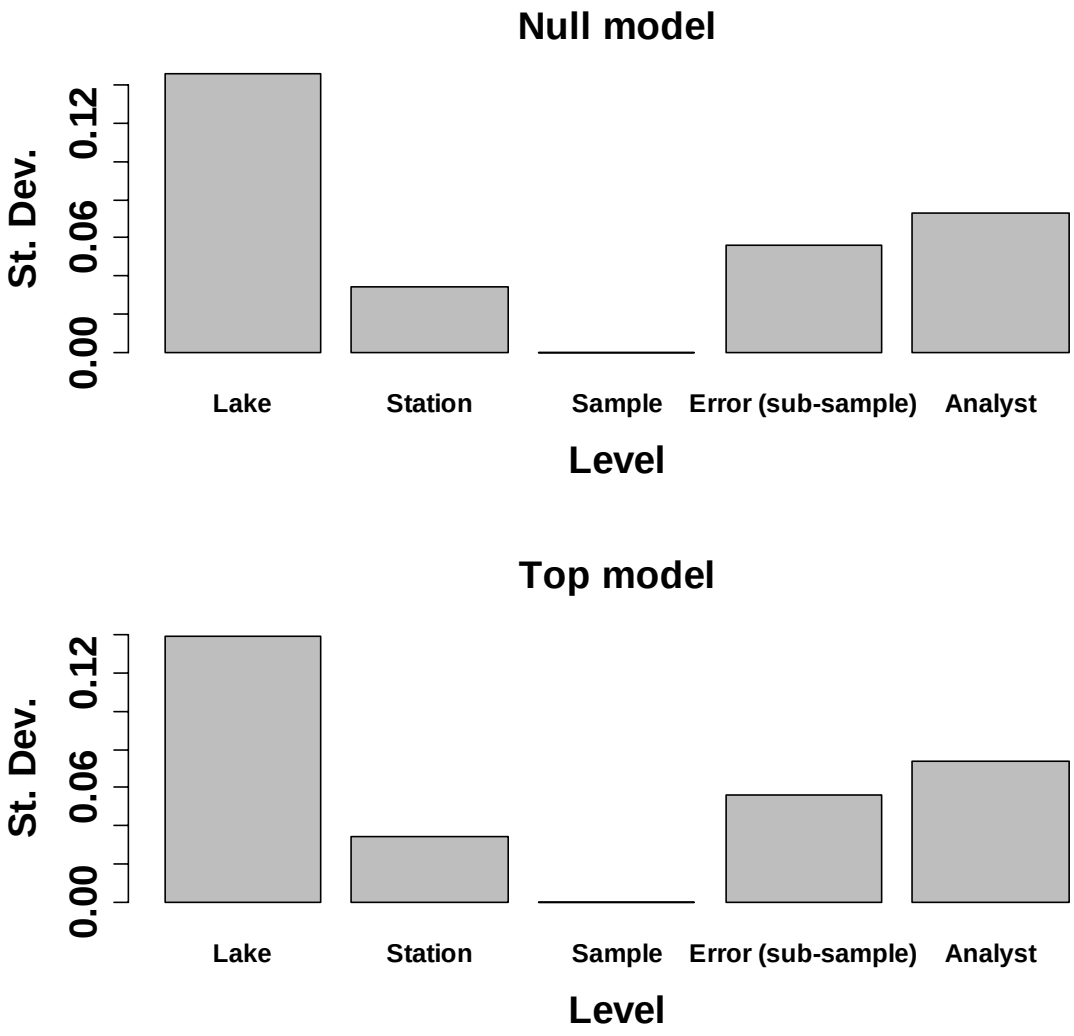

Fig. 3.18 Standard deviations of Evenness at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC

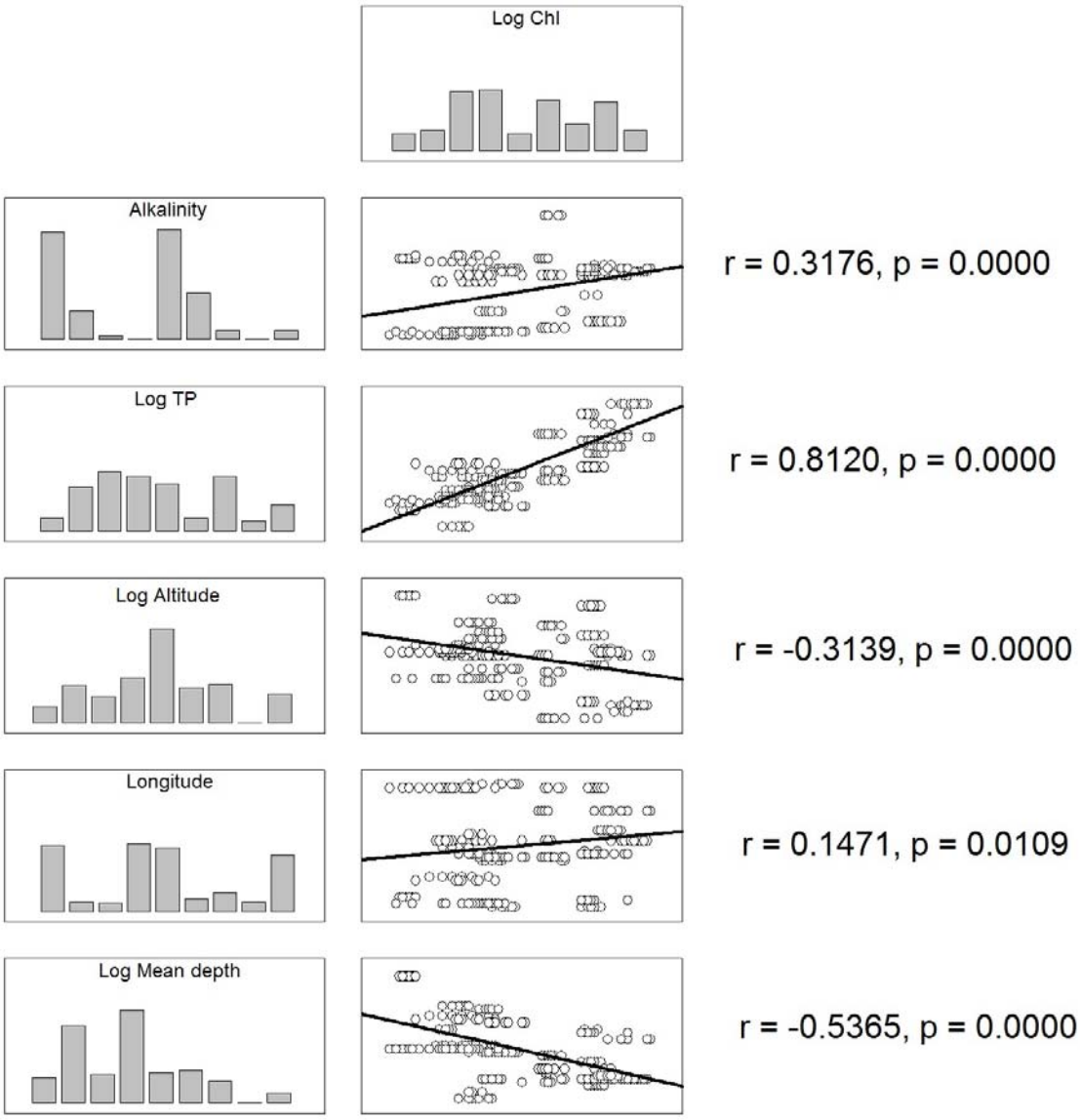

Fig. 3.19. Scatterplots of chlorophyll a *concentration against the predictor variables included in the analysis.*

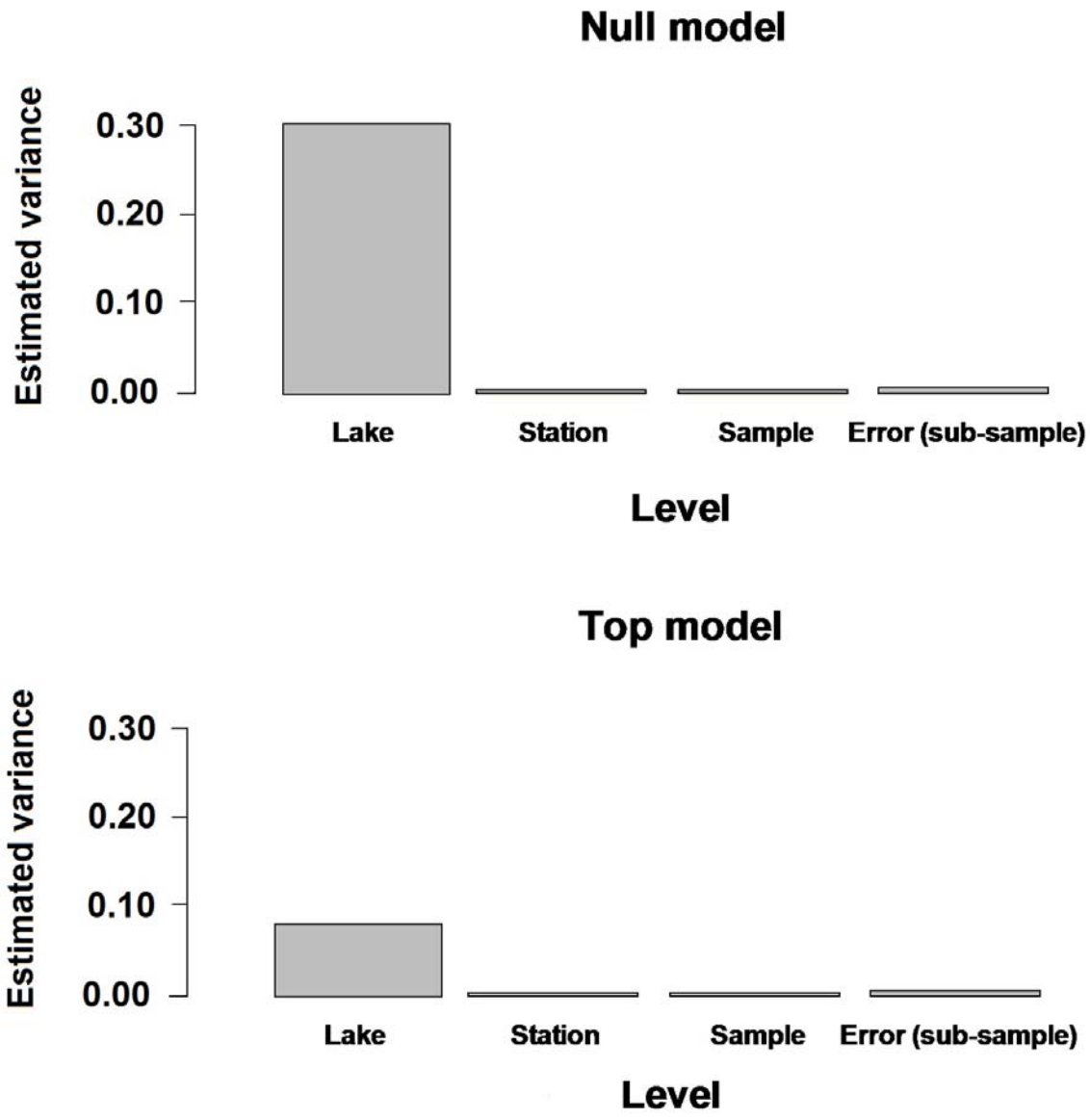

Fig. 3.20 Estimated variance of Chlorophyll a at different levels in the sampling hierarchy in the null model and the model with the most optimal combination of predictor variables, selected according to AIC

4 Discussion and conclusions

Phytoplankton communities are taxonomically complex and are affected by a wide variety of intrinsic and extrinsic environmental drivers. Their responsiveness to such drivers is a desirable property if we are to use them as biological indicators of water quality. However, this mutliplicity of influential factors and species responses makes the effects of single environmental drivers challenging to resolve. During the earlier stages of WISER, six candidate lake phytoplankton composition metrics were proposed, alongside the established metric of Chl, as biological indicators of the primary pressure gradient of eutrophication. Full details of these six metrics are given in WISER Deliverable D3.1-1: *Report on phytoplankton composition metrics, including a common metric approach for use in intercalibration by all GIGs* (Phillips et al. 2010) and WISER Deliverable D3.1-2: *Report on phytoplankton bloom metrics* (Mischke et al 2010). Herein, we have analysed the results of a unique hierarchic field sampling campaign to resolve sources of variation in these seven metrics and the relationships between metrics and total phosphorus (as a proxy for eutrophication), as well as other environmental factors.

Exploratory analysis of the field exercise results showed that some of the seven metrics considered were correlated with each other (e.g. PTI with Chl, MFGI, FTI and total cyanobacterial biovolume) indicating that, to some extent, they yield similar "information" on the biological differences between sub-samples. However, these correlations were far from perfect, indicating that each metric was also capturing unique aspects of phytoplankton community change. At the extreme, the evenness metric appeared quite unrelated to any of the remaining metrics, suggesting that this captured changes in the phytoplankton community to which the other metrics were less sensitive. This may suggest that assessments based upon multiple metrics will provide a more sensitive indication of biological responses to pressure gradients, and further work on defining a minimal suite of metrics is a high priority.

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 variability 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 variability in metric scores was attributed between lakes. Of the remainder variability, between-analyst and between sub-sample variation accounted for most of within-lake variation, with little attributable to variation between lake stations and multiple samples collected from each station. This is despite the fact that we have treated lake stations as "random" despite them being selected: which ought

to over-estimate the station-to-station variability. Lake stations were selected to represent water columns of mean depth or greater 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. For Chl and cyanobacterial abundance the within-lake spatial variation was very small (4% and 6% respectively) and nearly equal for the station, sample, sub-sample and analyst levels.

The highest contribution of analyst and sub-sample variation was apparent for the SPI and Evenness metrics (32% & 27% respectively) even though these two metrics do not rely so explicitly on consistency in taxonomic identification. The more taxonomically explicit metrics of PTI and cyanobacteria abundance had relatively low contribution of analyst and sub-sample variance (11% & 5% respectively). One possible reason for this was that 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). The results indicate that rigorous standardisation of sample mixing and sedimentation protocols, as well as of taxonomic procedures, 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 is currently unquantifiable. It is, therefore, difficult to truly isolate the effect of analyst variation upon metric scores and future studies targeting sources of variation arising from sampling processing and analyst variation alone would be welcomed.

For four of the seven metrics, total phosphorus concentration was the best single predictor of metric variation. However model fits were frequently improved, if only modestly, by the inclusion of geographic variables such as longitude and altitude. Alongside these variables, total phosphorus concentration was included in the most optimal models of five of the seven metrics. This would suggest that many of the proposed metrics are indeed responsive to the eutrophication pressure gradient apparent across these lakes. However, bivariate scatterplots and comparisons of among-lake metric variability between models showed that much among-lake metric variation remained unexplained by the available environmental data. This is indicative of the existence of important unmeasured drivers of phytoplankton community structure. The geographic variables included will act 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, all these variables were not recorded consistently enough to resolve their effects in the current analysis. The

variability in the Chl metric was better explained by the environmental parameters, particularly total phosphorus. Since the classical work by Dillon & Rigler (1974), the phosphorus – chlorophyll *a* relationship has been the most exploited chemistry-biology relationship in limnological models. Our analysis shows that this strong relationship is, however, modified along several independent natural gradients: mean depth, alkalinity and altitude (in decreasing importance). The significant relationships observed in this study for Chl were the same as those found in regression models predicting reference Chl concentrations in European lakes (Carvalho et al., 2010), with the exception of colour that was not investigated in this study.

Variability around the relationship between each metric and contemporaneously measured environmental variables 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.

In summary, lake phytoplankton metrics proposed by the EC WISER Project for ecological quality assessment (Phillips et al., 2010; Mischke et al., 2010) are generally shown to be robust metrics. Variability in metric scores is largely due to variability between lakes; differences in locations around a lake, or sampling and analytical variability, only account for a small proportion of the variability in metric scores. This is especially true for four candidate phytoplankton metrics being considered for Intercalibration: chlorophyll, PTI, MFGI and cyanobacterial blooms, for which >85% of the variability in metric scores was attributed between lakes, and, total phosphorus concentration was the best single predictor of this variation. Although, much among-lake metric variation still remained unexplained by the available environmental data, we conclude that these four proposed metrics are sufficiently robust metrics for ecological status assessment and are suitable for adoption in the Intercalibration process.

5 Acknowledgements

We are indebted to all who processed the samples from the field campaign: Caridad de Hoyos (CEDEX), Christophe Laplace (CEMAGREF), Martina Austoni (CNR), Pierisa Panzani (CNR), Tomasa Virdis (University of Sassari), Maria Antonietta Mariani (University of Sassari), Aimar Rakko (EMU), Kairi Maileht (EMU), Agnieszka Pasztaleniec (IEP), Małgorzata Poniewozik (IEP), Ute Mischke (IGB), Barbara Meinck (IGB), Andreas Nicklisch (IGB), Annette Tworeck (LBH Freiburg), Laurence Carvalho (NERC-CEH), Mitzi De Ville (NERC-CEH), Rene Groben (NERC-CEH), Sarah Pritchard (Beacon Biological), Bill Brierley (EA), Birger Skjelbred (NIVA), Stina Drakare (SLU) & Marko Jarvinen (SYKE).

6 References

Brierley, B.,Carvalho, L., Davies, S. and Krokowski, J. 2007. Guidance on the quantitative analysis of phytoplankton in freshwater samples. Report to SNIFFER (Project WFD80), 24pp.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York. 488pp.

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

Carvalho, L., Dudley, B., Dodkins, I., Clarke, R., Jones, I., Thackeray, S. J. & Maberly, S. C. (2007). Phytoplankton classification tool (Phase 2). Final report to SNIFFER (Project WFD80), pp 94.

Carvalho, L., Solimini, A., Phillips, G., Pietiläinen, O-P., Moe, J., Cardoso, A.C., Lyche, A., Ott, I., Søndergaard, M., Tartari, G. & Rekolainen, S. 2009. Site-specific Chlorophyll Reference Conditions for European Lakes. Hydrobiologia 633: 59-66.

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

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

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

Dudley, B., Dunbar, M., Penning, E., Kolada, A., Hellsten, S. and Kanninen, A. 2010. WISER Deliverable D3.2-2: Report on uncertainty in macrophyte metrics, 29pp.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C. K., Heiskanen, A.-S., Johnson, R. K., Moe, J., Pont, D., Solheim, A. L. and van de Bund, W. 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, Int. Org. Standard., Geneva, 1st edn. 1992–07–15, 6 pp.

Knap, A., Michaels, A., Close, A., Ducklow, H. and Dickson, A. (eds.). 1996. Protocols for the Joint Global Ocean Flux Study (JGOFS) Core Measurements. JGOFS Report Nr. 19, 170 pp. Reprint of the IOC Manuals and Guides No. 29, UNESCO 1994. http://ijgofs.whoi.edu/Publications/Report_Series/JGOFS_19.pdf

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.

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

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

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

Murphy, K. J., Kennedy, M. P., McCarthy, V., Ó'Hare, M. T., Irvine, K. and Adams, C., 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 Reynolds, C. S. 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.

Pielou, E.C. 1969. An introduction to mathematical ecology. Wiley-Interscience, New York. 294pp.

Pielou, E.C. 1975. Ecological diversity. Wiley Interscience, New York. 165pp.

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

Pinel-Alloul, B., and Ghadouani, A. 2007. Spatial heterogeneity of planktonic microorganisms in aquatic systems: mutiscale patterns and processes, pp 203-307. In: The Spatial Distribution of Microbes in the Environment. R. Franklin and A. Mills (Editors). Springer.

Pinheiro, J.C. and Bates, D.M., 2000. Mixed-effects Models in S and S-Plus. Springer, New York, 528pp.

R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reynolds, C. S. 2006 Ecology of phytoplankton. Ecology, Biodiversity and Conservation Series. Cambridge University Press, Cambridge, 564 pp.

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

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

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

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. and Smith, G. M. 2009. Mixed effects models and extensions in ecology with R. Springer, New York. 574pp.