STATISTICAL WORKSHOP ON GRADIENT STUDIES

UPPSALA, 23 - 25 OCTOBER 2013

Editor: Ulf Grandin

Authors: Thorsten Balsby, Lena Bergström Jacob Carstensen, Frauke Ecke, Ulf Grandin, Anders Grimvall, Simon Hallstan, Kerstin Holmgren, Maria Kahlert

WATERS Report no. 2013:8

WATERS Report no. 2013:8

Statistical workshop on gradient studies

Uppsala, 23 - 25 October 2013

Editor: Ulf Grandin, Swedish University of Agricultural Sciences

Authors:

Thorsten Balsby, Aarhus University

Lena Bergström, Swedish University of Agricultural Sciences

Jacob Carstensen, Aarhus University

Frauke Ecke, Swedish University of Agricultural Sciences

Ulf Grandin, Swedish University of Agricultural Sciences

Anders Grimvall, Swedish Institute for the Marine Environment

Simon Hallstan, Swedish University of Agricultural Sciences

Kerstin Holmgren, Swedish University of Agricultural Sciences

Maria Kahlert, Swedish University of Agricultural Sciences

WATERS partners:

WATERS: Waterbody Assessment Tools for Ecological Reference conditions and status in Sweden

WATERS Report no. 2013:8.
Title: Statistical workshop on gradient studies, Uppsala, 23-25 October 2013
Publisher: Havsmiljöinstitutet/Swedish Institute for the Marine Environment,
P.O. Box 260, SE-405 30 Göteborg, Sweden
Published: December 2013

Please cite document as:
Gradin, U (ed). Balsby, T., Bergström, L., Carstensen, J., Ecke, F., Grandin, U., Grimvall, A., Hallstan, S., Holmgren, K., Kahlert, M. Statistical workshop on gradient studies, Uppsala, 23-25 October 2013, WATERS Report no. 2013:8. Havsmiljöinstitutet, Sweden.

http://www.waters.gu.se/rapporter

WATERS is a five-year research programme that started in spring 2011. The programme's objective is to develop and improve the assessment criteria used to classify the status of Swedish coastal and inland waters in accordance with the EC Water Framework Directive (WFD). WATERS research focuses on the biological quality elements used in WFD water quality assessments: i.e. macrophytes, benthic invertebrates, phytoplankton and fish; in streams, benthic diatoms are also considered. The research programme will also refine the criteria used for integrated assessments of ecological water status.

This report summarises the third statistical workshop in WATERS that was held in Uppsala 23-25 October.

WATERS is funded by the Swedish Environmental Protection Agency and coordinated by the Swedish Institute for the Marine Environment. WATERS stands for 'Waterbody Assessment Tools for Ecological Reference Conditions and Status in Sweden'. Programme details can be found at: http://www.waters.gu.se

Table of contents

Summary

The third statistical workshop in WATERS was held in Uppsala from 23rd to 25th of October 2013 with the aim of indicator development and uncertainty assessment of indicators. Data analysed at the workshop comprised long-term monitoring data sets and data sampled during the gradient studies in WATERS. A total of 12 persons attended the workshop. Three statistical lectures were given on principles of indicator development and uncertainty assessment. Following the lectures smaller groups were formed combining data providers and statisticians, aiming at analysing the data using appropriate statistical techniques. The outcome of these exercises was reported back to the entire group and discussed, and summarised as separate sections in this report. Although time during the workshop did not allow for an exhaustive examination of the data sets, collaboration between biologists and statisticians was established and these initial analyses will be pursued further in the future. Thus, the workshop was successful in bridging biological and statistical expertise within WATERS.

Svensk sammanfattning

Den tredje workshopen inom WATERS FA2 hölls i Uppsala 23-25 oktober 2013. Syftet var att stödja forskare inom framförallt FA4 (sötvatten) vid utvecklandet av indikatorer för vattenkvalitet. Huvuddelen av seminariet ägnades åt gruppvisa diskussioner kring dataanalys, där deltagarna hade med egna data att diskutera med statistikerna i FA2. Utöver grupparbeten hölls föreläsningar i statistik och kring principer för indikatorutveckling. Resultaten från gruppövningarna återrapporterades i plenum för vidare diskussion kring respektive analys. Resultaten från gruppövningarna samt sammanfattningar av föreläsningarna är sammanställda i detta dokument.

Sammantaget upplevde samtliga deltagare att workshopen var värdefull, och att den lade grunden till ett fortsatt samarbete mellan biologer i FA3 och FA4 och statistiker inom WATERS.

1 Introduction

The third statistical workshop in WATERS was held in Uppsala from 23rd to 25th of October 2013 at the Swedish University of Agricultural Sciences (SLU). The workshop was announced in August 2013. The workshop was attended by 12 participants from WATERS, who brought diverse sets of data. The workshop was mainly intended for participants from scientific focus area 4 (FA4) dealing with freshwater, as the previous statistical workshop on indicator development took place when FA4 people didn't have their data ready for analysis.

The objective of the workshop was to analyse biological monitoring data within WATERS in relation to meteorological data and pressure data with the aim to develop indicators that clearly respond to anthropogenic pressures when other sources of variations have been filtered out. The workshop included one lecture on indicator development, one presentation on data quality and one presentation of the uncertainty framework developed in WP2.2. The focus of the workshop was on analysing data in smaller groups involving both biologists and statisticians.

This summary report contains a short description of the statistical presentations, the outcome of the group work, the agenda for the workshop and a list of participants.

2 Basic concepts of indicator development

Lecture given by Ulf Grandin, Swedish University of Agricultural Sciences.

There are several definitions of what an indicator is. In essence, all definitions state that an indicator is a simple measure related to something more complex of primary interest. Some definitions can include a direction of temporal change, e.g. "A summary measure related to a key issue or phenomenon that can be used to show positive or negative change" (Statistics New Zealand, 2009). Other only focus on trends, e.g. "A statistic or parameter that, tracked over time, provides information on trends in the condition of a phenomenon and has significance extending beyond that associated with the properties of the statistic itself" (OECD, 1994). Some include a relationship between the observed parameter and a societal goal, e.g. "A statistic or measure which facilitates interpretation and judgements about the condition of an element of the world or society in relation to a standard goal" (USEPA, 1972). A last example brings in support for decision-making: "A simple summary of a complex picture, abstracting and presenting in a clear manner the most important features needed to support decision-making" (United Nations, 2007).

When developing an indicator, the first question to ask is what the indicator should indicate. It may be a process, a state or a function. These three concepts are linked together in an ecological hierarchy from the presence or absence of an individual species up to the landscape or region scales (Dale & Beyeler, 2001, Table 2.1).

TABLE 2.1
Different levels of the ecological hierarchy and their associated processes, presented with some suggested indicators and what ecological key characteristic that is indicated (after Dale and Beyeler 2001).

Hierarchy	Process	Suggested indicator	Key characteristic
Organism	Environmental toxicity	Physical deformation	Function
	Mutagenesis	Lesions	Function
		Parasite load	Function
Species	Range expansion or contraction	Range size	Structure
	Extinction	Number of populations	Composition
Population	Abundance fluctuation	Age or size structure	Structure
	Colonisation or extinction	Dispersal behaviour	Multi
Ecosystem	Competitive exclusion	Species richness	Composition
	Predation or parasitism	Species evenness	Composition
	Energy flow	Number of trophic levels	Function
Landscape	Disturbance	Fragmentation	Structure
	Succession	Spatial distr. of communities	Structure
		Persistence of habitats	Function

Indicators may be divided into biological/ecological and societal indicators. The former mostly relate to physical or observed objects, while the latter encompasses more abstract processes such as economic development or legislation. Societal indicators are often divided according to the DPSIR framework (developed by OECD and adopted by i.a. the European Environmental Agency, EEA). The different parts of the framework typically include:

- Driving forces, which are the large scale drivers such as societal, demographic and economic development,
- Pressure and State, which describe causes of environmental chance, e.g. emissions, aliens species or habitat fragmentation, or thousands of other objects or processes that can be measured,
- Impact, which describe changes in environmental conditions, may be both ecological and chemical conditions,
- Response, which is the societal measures to mitigate environmental degradation.

Ecological indicators can be divided in several ways (**TABLE 1**.2). An influential scientific paper by Noss (1990) suggested a division into: *Flagships*, *Umbrella species* and *Keystone species*. This list can be complemented by: *Ecological engineers* and *Link species*.

Irrespective of the type of indicator to be developed, there are some shared characteristics that all indicators should possess. These include:

- Rapid and targeted response to the focal factor
- Low noise:
    - Low natural variability
    - Low sampling variability
- Same signal over whole measured range
- Sufficient span in measured range
- Inexpensive
- Easily measured/sampled
- Fairly common

TABLE 1.2
Different types of ecological indicators.

Indicator	Description	Pros	Cons	Example
Flagships	Often a large, charismatic vertebrate	good symbol	Little use as indicator of diversity; Expensive to preserve	The panda
Umbrella species	Species that need large and varying habitats	Many species gets an indirect protection; Relatively simple	Based on probability calculations; Efforts on the umbrella species disadvantage other species	Northern spotted owls, for old growth forests in northern America White-backed Woodpecker in Sweden
Keystone species	Species that secure the survival of many other species	Focus on one species; Guarantee the survival of many species; Based on knowledge about ecosystems	Difficult to identify key stone species; Unknown how many ecosystems that have key stone species	Star fish, predating on mussels; Elephants, maintaining the African savannah
Ecological engineers	Alters habitats, thereby creating habitats for other species Close to Keystone species	Focus on one species; Guarantee the survival of many species; Based on knowledge about ecosystems	Few good examples; Habitat alternation may lead to conservation conflicts	Beavers Stoneflies
Link species	Important for the transport of matter and energy across trophic levels	Focus on one species; Secures ecosystem functions; Based on knowledge about ecosystems	Based on probabilities; Ecosystems indirectly monitored	Pollinators; Herbivorous pray species

In addition to these general characteristics indicators may also have specific requirements depending on their type. Indicators that in addition to their primary goal also should include the society or a ley public should also comply with the following characteristics:

- Simplicity – will people understand the indicator and find it interesting?
- Ease of communication – can the indicator be communicated and will it be associated with biodiversity?
- Importance and relevance – does the indicator describe an important aspect of the biodiversity issue clearly and unambiguously?
- Measurability – is it easy enough to obtain data?
- Action orientation – will this choice of indicator change the way people behave and think, will it stimulate action and indicate which direction the action should take you?
- Strong people resonance – will the choice of indicator "ring true" to people?

To summarise, there are thousands of indicators and more are developed. When developing an indicator it is important to have several factors in mind. If not, the indicator may indicate different things depending on where or when the indicator is assessed, or in the worst case it may not indicate at all what was intended.

3 Uncertainty framework

Lecture given by Jacob Carstensen, Aarhus University.

In these two combined lectures the uncertainty framework that has been developed in WP2.2 was presented and exemplified with data on eelgrass shoot density from Öresund and BQI from the Skagerrak coast and the Bothnian Sea. The framework partitions variations in monitoring data into temporal, spatial, spatio-temporal and methodological, and the different uncertainty components in the framework was presented and discussed. It was stressed that it is not relevant to consider all uncertainty components for each BQE indicator, as some of these may be considered negligible relative to the other sources of uncertainty. However, the relative importance of the different uncertainty components is specific to the type of data and the sampling procedure. The formulas for calculating the resulting variance on a mean value, assuming this to represent the indicator value, were shown for both a crossed design and a hierarchical design.

Eelgrass shoot density from Öresund has been collected at 13 locations, several of these represented by up to 5 stations along a depth gradient. Six replicates were taken at each sampling occasion. The time series ranged from 1 to 17 years of monitoring, and between 1 and 4 different divers had been involved in the sampling at the different localities. Consequently, the data set was quite heterogeneous with number of observations across localities ranging from 6 to 450. This implied that it was not possible to identify a broad range of uncertainty components at all localities. However, using the entire data set it was

possible to estimate five different uncertainty components, and by modelling the large-scale spatial variation within localities using depth as explanatory variable the estimates of the variance components were reduced substantially. In the presentation it was stressed that a large data set is indeed needed, if several uncertainty components are to be estimated with a reasonable accuracy.

Another example using the benthic quality index (BQI) of benthic invertebrates from the Skagerrak and the Gulf of Bothnia was presented. Data from three years and a total of 24 stations in the Skagerrak and 100 stations in the Gulf of Bothnia were used to estimate spatial and temporal components of variability. The analyses revealed some common patterns among coastal areas, i.e. the large importance of spatial variability among stations (including both static and interactive sources of variability), as well as differences among coastal areas. These included general differences in precision due to differences in overall means and patterns of variability (relative to its mean precision in the Gulf of Bothnia is poorer than in the Skagerrak) and differences in estimation procedures as a consequence of monitoring designs.

The following discussion on the uncertainty framework showed that there was a great need and expectations on further interactions between the cross-cutting work packages developing routines for uncertainty assessment and the work packages dealing with development of individual quality elements. Such interactions will be necessary to develop coherent "uncertainty libraries" and harmonised principles for uncertainty assessments.

4 Data quality

Lecture given by Ulf Grandin, Swedish University of Agricultural Sciences

Before initiating statistical hypotheses testing and modelling, the quality of the data needs to be checked. In this lecture, I presented a number of common mistakes or errors in data sets that need consideration or correction before more detailed statistical analyses are carried out (Zuur et al. 2010). The examples were divided into six different categories:

1. Look at the data!
2. Are there any typos?
3. Are there missing values or missing periods?
4. Are there any outliers, or deviating periods?
5. Are there values below a detection limit?
6. Are there sufficient amount of data?

4.1 Look at the data

An essential part of initial data exploration is looking at the data by different types of plots, especially for univariate analyses. Bivariate scatterplots (Figure 4.1), boxplots and histograms of the distribution are the most common types of plots to visualise data.

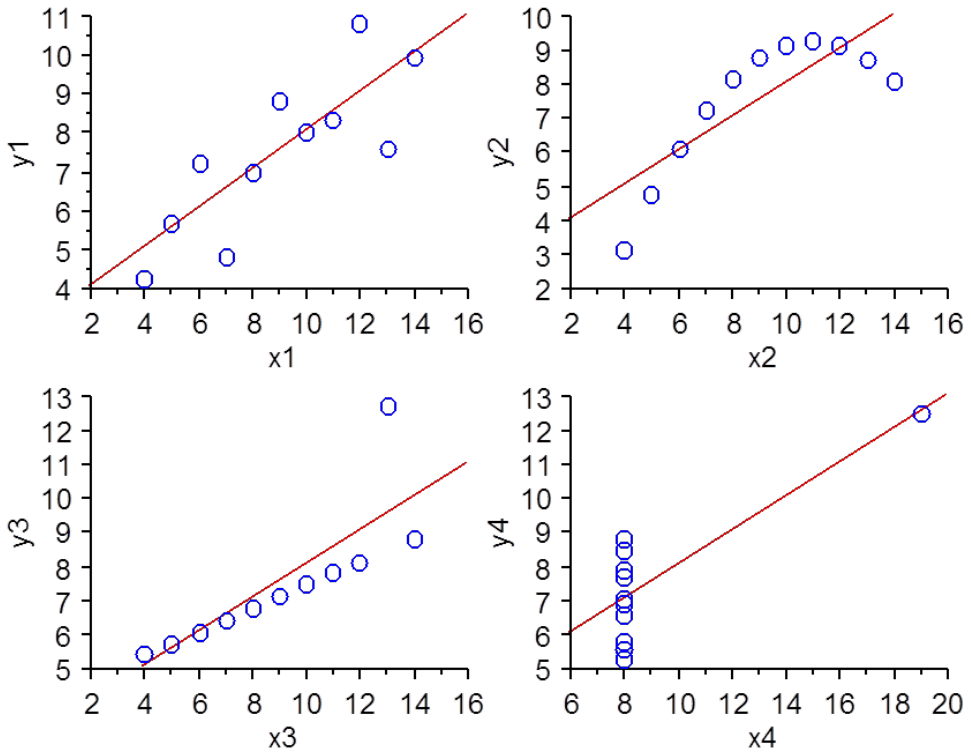


FIGURE 4.1
Scatterplots of four bivariate datasets all described by the relationship y = 3 + 0.5 × x (r^2 = 0.67, p = 0,002). The visualisation illustrates that only one of the data sets (upper left) meets the prerequisites for an ordinary least square linear regression.

Visualisation of univariate data can for example reveal:

- deviating values,
- zeros as observed value or a code for missing value,
- the distribution of the data,
- type of relationship between x and y,
- missing values,
- trends.

4.2 Are there any typos?

The risk of typos is always high when humans have typed in data. An initial check should include check of data type for each variable, e.g. that numeric variables are recognised as numeric by the software, or that integer variables contain only integers. Next step is to look at the maximum, minimum and mean values to ascertain that these are within expected ranges.

However, even if the typing is correct, data coding or automatic formation in software may anyway cause errors that eventually will be seen as typos or cause erroneous results. One example is how missing values are entered. Common coding for missing values includes an empty cell, a period (.), -99 or -9999.

Merging datasets from different sources requires careful checking of the data in each data set. Common pit falls when merging data sets include differences in:

- decimal separator; comma or point,
- symbol used as thousands separator (e.g. 10 357 or 10'357 vs. 10357),
- how censored values are notated, e.g. <0.01or just 0.01,
- units for the same variable between datasets, e.g. $\mu g * l^{-1}$ or $mg * l^{-1}$,
- coding of missing values.

4.3 Missing values

A few missing values in a large data set is generally no problem, given that the missing values are randomly distributed in the data. Problems arise when there are systematic missing values.

No computer program can do magic and replace a missing value with the correct value, but most programs have routines to handle missing values. It is very important to be aware of how missing values are treated by the software used for analysis. The most common ways are:

- automatic removal of the whole column containing a missing value,
- automatic removal of the whole row containing a missing value,
- automatically modelled from non-missing values in the data set.

These automatic procedures may thus remove important data from other row or columns, without notification, or by a general comment in the output.

4.4 Deviating values or outliers

Outliers are a large problem in some statistical analyses, while they cause no problem in others. Generally, outliers pose no problems in methods based on ranking of the original data. There is no generally accepted definition of what an outlier is. It is up to the scientist to define what an outlier is in a specific dataset and under a certain hypothesis. However, there are a number of statistical tests to help identify outliers. One example is Grubb's test. A simple way to detect outliers is to draw histograms and boxplots, or a Cleveland plot where observed values are plotted against row number in the data set (Figure 4.2).

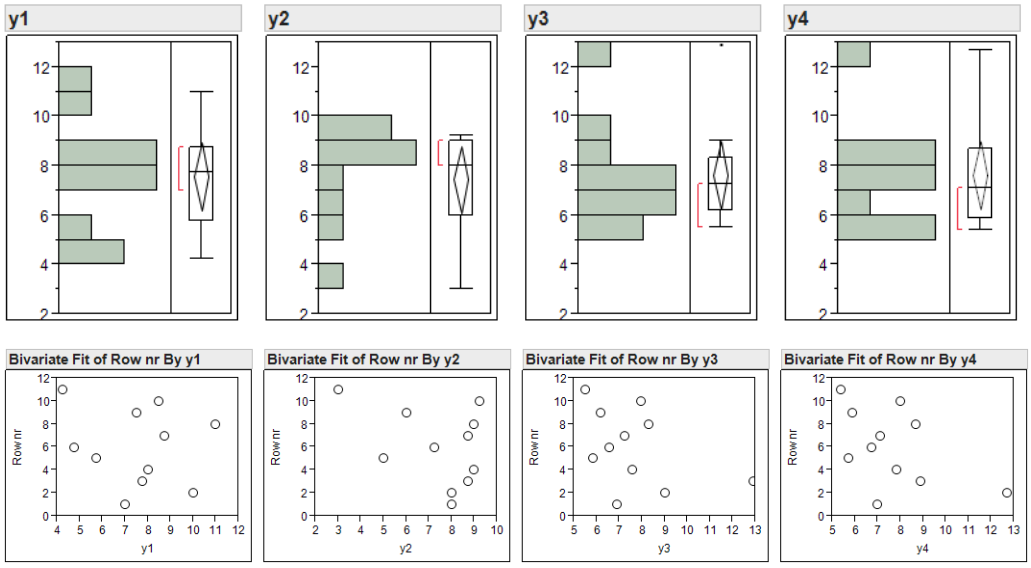


FIGURE 4.2
Histogram, boxplot and Cleveland plot of the data in fig 1, used to detect outliers.

Removal of deviating values should be done with caution. Obvious typos or measurement errors should (of cause) be corrected or removed. In other cases, removal of data should only be done if it can be justified by the theory and hypotheses behind the sampled data.

4.5 Values below a detection limit

There are many ways to deal with values that are reported as being below detection limit. See e.g. miljostatistik.se for a summary of the most common methods (in Swedish).

4.6 Are there sufficient amount of data?

Design of studies and a priori analysis of statistical power are important to ascertain that the sampled data can answer the research question at hand. Most statistical analyses are principally related to the question: "-Which risk am I willing to take that all my data are a collection of extreme values not representative for the population I'm studying". Increased number of samples decreases this risk, but at the cost of more expensive data acquisition. The cost for increasing sample size should be balanced against the cost (economically, ecologically and in decreased credibility) of:

a) miss a true negative impact on the environment, or
b) erroneously say that there is an environmental impact.

The theory of statistical power calculation is too wide to be summarised in this document. There are plenty of literature on statistical power and design of experiment. Readers understanding Swedish could start by looking at miljostatistik.se.

5 Fish data from the gradient study in lakes

Exercise summarised by Kerstin Holmgren, Swedish University of Agricultural Sciences and Thorsten Balsby, Aarhus University

One of the gradient studies of WATERS consisted of ten lakes in Uppland, with 4-30 % agricultural land and 0.5-13 % urban land in the catchments. Biological samples except fish were taken in summer and autumn 2012, and water samples in August 2012 revealed a total phosphorous (TP) gradient from 7-131 □ g/L. The fish communities of each lake were sampled once during 2008-2013.

In the present assessment criteria for fish in Swedish lakes, the fish fauna is assessed by a multi-metric index. The fish index EQR8 is based on eight whole-lake metrics (see below). We knew in advance that EQR8 responds only weakly to increasing TP. Five lake characteristics, e.g. altitude and maximum depth, were originally used to estimate site-specific reference values of fish metrics. Unfortunately, low altitude and shallow lakes were not well represented in the least disturbed lakes used to calibrate multiple regression models. In the gradient study, the lowest TP was found in a relatively deep lake with reference value for fish biomass estimated to 1274 g/gillnet. The lake with highest TP was much shallower, with a higher reference value for fish biomass (3476 g/gillnet). Total fish biomass is expected to increase with increased eutrophication. A high covariance between lake morphometry and nutrient pressure will, however, confound the response of whole-lake fish metrics in the eutrophication gradient.

Before just comparing the fish index EQR8 (as used in the present Swedish assessment criteria) with other biological quality elements in the gradient study, we wanted to explore variation in alternative fish metrics, e.g.;

1) The EQR8 metrics calculated from littoral catch (0-3 m depth) instead of pooled for the whole lake
2) Length of perch after each of its first to seventh completed growth seasons

5.1 Description of data and calculated fish metrics

The fish communities of each lake were sampled in July or August, using multi-mesh Nordic gillnets, according to a European standard method (EN 14757). The recommended or default sampling effort (number of benthic gillnets) increases with area and maximum depth of the lake, corresponding to 16-40 gillnets per lake in the present lakes. The gillnets were set randomly within fixed depth strata, covering available depths of 0-3 m, 3-6 m, 6-12 m, 12-20 m, 20-35 m, and 35-50 m. The catch was recorded as number of individuals and biomass (g), for each fish species caught in each gillnet. Additionally each individual was measured for total length (mm).

No. 2-8 of the following EQR8 metrics were calculated for each net set in the shallowest depth stratum (0-3 m depth, Table 5.1);

1. **Number of native fish species**.
2. **Simpson's Dn** (**SDn**, diversity index based on number of individuals): calculated as 1 / (□ P_i^2), where P_i = numerical proportion of species i, and the sum is taken for all species in the catch.
3. **Simpson's Dw** (**SDw**, diversity index based on biomass): calculated as 1 / (□ P_i^2), where P_i = biomass proportion of species i, and the sum is taken for all species in the catch.
4. **Relative biomass of native fish species (BMtot)**: total biomass (g) of all native species, divided by number of nets.
5. **Relative abundance of native fish species (Ntot)**: total number of individuals of all native species, divided by number of nets.
6. **Mean mass (MeanW)**: biomass of all species (g) divided by the number of individuals.
7. **Proportion of piscivorous percids** (**PropPisc**, based on biomass in the total catch): The proportion of potentially piscivorous perch is 0 at fish length less than 120 mm and 1 at length above 180 mm. At intermediate length the proportion is calculated as 1 – ((180 – length) / 60). Individual mass of perch (g) is estimated as $a \cdot \text{length (mm)}^{b}$, where $a = 3.377 \cdot 10^{-6}$, and b = 3.205. Each individual mass is multiplied with the length-specific proportion piscivorous perch. The sum of the products is the biomass of piscivorous perch, which is then added to any biomass of pikeperch. Finally, the total sum of piscivorous percids is divided by the total biomass of all species in the catch.
8. **Ratio perch / cyprinids** (**P/C-ratio**, based on biomass): total biomass of perch divided with total biomass of all native cyprinids.

TABLE 5.1
Data used in the analysis. Nets is the total number of gillnets set in the depth stratum 0-3 m. Perch samples were taken in 2013, or in 2008 for Lake Tärnan. Age is the number of aged perch. Min-max is the range of ages observed. L_{0+} -L_{6+} is the number of estimates of perch length after the first to seventh years of growth.

Lake	Years	Net	Age	Min-max	L0+	L1+	L2+	L3+	L4+	L5+	L6+
Bottenfjärden	2013	8	73	0+ - 10+	60	49	46	19	15	12	2
Largen	2008	8									
Lejondalssjön	2010	12									
Lilla Ullfjärden	2007, 2013	13	71	0+ -10+	64	57	38	32	26	16	13
Lommaren	2009	6									
Långsjön	2013	8	76	0+ - 10+	73	56	43	25	23	17	10
Sparren	2013	9	70	0+ - 11+	68	60	49	33	30	21	19
Syningen	2009	16									
Tärnan	1996, 1998, 2001, 2004, 2008	40	71	0+ - 13+	65	53	45	33	28	22	14
Ullnasjön	2010	12									

In five lakes, ca 70 perch (*Perca fluviatilis*) were sub-sampled for age determination (Table 5.1), using both sagittal otoliths and operculum bones. Distances between annuli on the operculum bones were used to back-calculated length after each completed year of growth. We primarily intended to use back-calculated length of the first growth season (L_{0+}, mm), i.e. one or more years before the fish was caught in the gillnet. Another consequence is that for fish sampled in one single year, first year growth can be estimated and compared between fish born in successive years. After some consideration, we also included back-calculated length after the second to sixth growth seasons (L_{1+} to L_{6+}) in preliminary tests.

5.2 Analysis

The relationships between seven of the EQR8 metrics and eutrophication were tested with a mixed model with the EQR8 metrics as dependent variables and total N, total P and proportion of agriculture were independent variables and lake identity was the random variables. Differences in perch growth (L_{0+} to L_{6+}) between lakes and years were analyzed with a general linear model and the effect of eutrophication on growth were all analyzed with general linear models for each age-class (0+ - 5+).

All analyses were made with SAS 9.3 (SAS Institute, Cary, NC). The analyses used either Proc Mixed or Proc GLM.

5.2.1 EQR8 metrics and eutrophication

Only one of the 7 EQR8 metrics showed a significant response to the eutrophication measures of the lakes (Table 5.2). The total biomass (BMtot) caught per net increased with an increase in total N (Table 5.2, Figure 5.1).

TABLE 5.2
Estimates and test of effect of TP, TN and proportion of agriculture on seven EQR8 metrics.

	TP				TN				Agriproc			
	Estimate	F	df	p	Estimate	F	df	p	Estimate	F	df	p
BMtot	12.4	0.43	1.122	0.51	9.2	11.9	1.12	**0.0008**	-85.7	1.68	1.122	0.2
Ntot	1.5	3	1.122	0.08	0.12	0.91	1.12	0.34	-0.62	0.00	1.122	0.84
SDw	-0.0049	1.53	1.122	0.22	0.0007	1.9	1.12	0.17	0.013	1.01	1.122	0.31
SDn	-0.008	1.18	1.122	0.28	0.0003	0.08	1.12	0.77	0.0044	0.02	1.122	0.88
MeanW	-0.18	0.67	1.122	0.42	0.008	0.07	1.12	0.8	-0.16	0.04	1.122	0.84
P/C-ratio	0.009	1.22	1.122	0.27	-0.0056	0.25	1.12	0.62	-0.005	0.04	1.122	0.85
PropPisc	-0.0001	0.01	1.122	0.93	0.00001	0	1.12	0.96	-0.0018	0.12	1.122	0.73

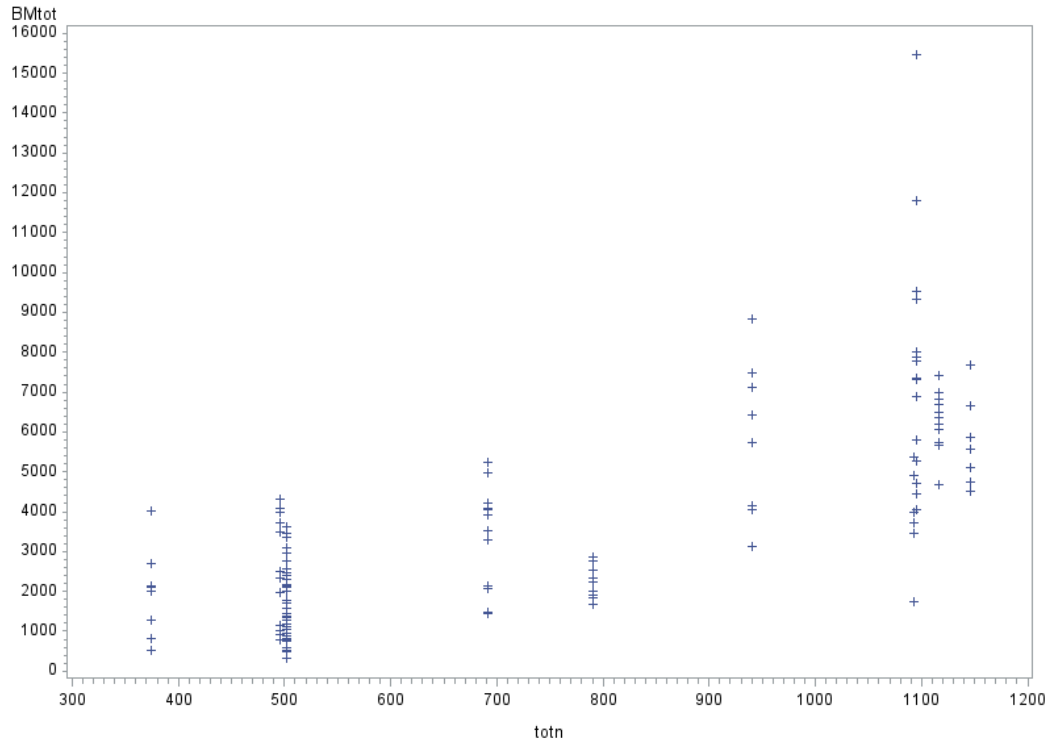


FIGURE 5.1
Total biomass versus total N for each net set in the 0-3 m depth stratum in the lakes surveyed.

5.2.2 Growth differences between lakes

The growth differed significantly between the five lakes for age classes 0 to 4, but not for age class 5 and 6 (Table 5.3). The year of growth only had a significant effect for fish in their first year (i.e. age-class=0). We have not analysed post hoc pairwise differences because they were not of interest.

TABLE 5.3
General linear model test of effect of year of growth and lake on the growth for each age class.

	Year of LX		Lake	
Age	**F**	**p**	**F**	**p**
0	4.04	<0.001	5.5	0.0003
1	1.45	0.13	10.9	<0.0001
2	0.99	0.47	12.99	<0.0001
3	0.5	0.91	9.58	<0.0001
4	0.44	0.93	3.58	0.0089
5	0.87	0.56	2.06	0.095
6	1.06	0.41	1.34	0.27

5.2.3 Eutrophication affects growth

Growth, or attained length, for age classes 3, 4, and 5 were lower at lakes with higher proportion of agriculture, as indicated by the significant negative estimates (Table 5.4). For younger age-classes there was no significant effect of proportion of agriculture.

Total nitrogen appeared to have an effect on growth in age-classes 0, 3, 4, 5. The estimates for the slopes were positive, which suggest that higher concentration of N resulted in higher growth. However, it should be mentioned that N had small influence relative to the proportion of agriculture in age-classes 3, 4, 5.

This analysis did not detect any effect of total phosphorous on the growth.

TABLE 5.4
Test of the effect of eutrophication on growth for each age class.

Age-class	Effect	NumDF	DenDF	Estimate	F Value	Prob F
0	totn	1	253	0.018444	6.52308	**0.011235**
0	totp	1	253	-0.03659	1.828115	0.177558
0	AgricProc	1	253	-0.10781	0.375327	0.540665
1	totn	1	198	0.017933	0.116708	0.732995
1	totp	1	198	-0.04289	0.041468	0.838846
1	AgricProc	1	198	-0.16652	0.016603	0.897606
2	totn	1	145	1.318588	1.62E-10	0.99999
2	totp	1	145	-2.19996	9.89E-11	0.999992
2	AgricProc	1	145	-30.9438	1.66E-10	0.99999
3	totn	1	78	0.22045	18.88518	**4.14E-05**
3	totp	1	78	-0.2302	1.823532	0.180799
3	AgricProc	1	78	-4.96174	19.98237	**2.62E-05**
4	totn	1	63	0.232445	6.747497	**0.011673**
4	totp	1	63	-0.37453	1.456615	0.231982
4	AgricProc	1	63	-4.39947	4.993921	**0.028991**
5	totn	1	35	0.320466	5.216286	**0.028556**
5	totp	1	35	-0.46253	0.932068	0.34095
5	AgricProc	1	35	-7.15171	4.669135	**0.037636**

6 Coastal fish data set

Exercise summarised by Lena Bergström, Swedish University of Agricultural Sciences.

Data on coastal fish communities in the Baltic Sea were analyzed in order to identify main patterns in temporal and spatial variation. The data set included 7303 data points (stations), distributed over the years 2002-2012 and 42 sites (Figure 6.1). On average, 30-45 stations were sampled at each site and year. These were sampled by depth stratification, so that an equal number of stations were sampled in the depth intervals 0-3, 3-6 and 6-10 meters (usually 10-15 stations per stratum). Additionally, a smaller number of stations were sampled at 10-20 m depth in some of the areas (usually 5 stations). Only some of the sites were sampled in more than one year. Eleven of the sites were sampled annually within the environmental monitoring programs. Sampling was performed using Nordic Coastal multimesh gill nets, which catches fish from about 10 cm length.

The data set was analyzed using generalized linear models. In order to obtain a comparable sampling effort among geographical areas, the analyses were narrowed down to data collected at depths 0-10 meters in August (6128 data points, 36 sites). Analyses with respect to total number of fish were performed assuming a Poisson distribution. Analyses with respect to indicators (ratio number of large fish to total number of fish; ratio number of large piscivores tot total number of piscivores) were performed assuming a binomial distribution.

There was a large variation among sites, whereas differences among years were generally smaller. There was no consistent pattern in differences among years, when comparing all sites at the same time. Variation within sites could be attributed to depth, but the direction of response varied among sites. Mean values for each site were plotted in GIS for explorative purposes. This was also done for the slope in relation to depth, based on analyses were depth was included as a continuous variable.

The primary question to address further is to explain the observed spatial differences. Potential explanatory variables were listed, and further analyses will aim at identifying their relationship to the abundance of key species and different functional groups/size groups. The potential explanatory variables represent natural as well as anthropogenic variables (eg temperature, water chemistry, nutrient loading, coastal morphology, water transparency, depth conditions, fishing pressure).

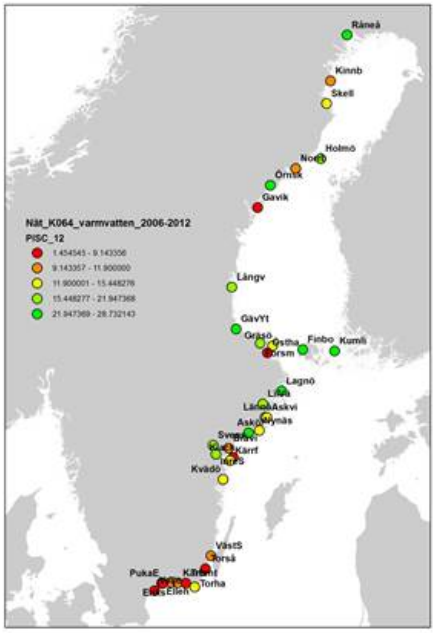


FIGURE 6.1
Sites included in the coastal fish data set. The color of the symbols show values for the indicator *abundance of piscivores*, classified by quintiles.

7 Diatoms in lakes and streams

Exercise summarised by Maria Kahlert, Swedish University of Agricultural Sciences.

The objective of this study was to further develop reference conditions, i.e. diatom reference communities, for Swedish streams and lakes. Today's phytobenthos method is based on traditional indices (IPS assessing eutrophication and organic pollution, supplement indices TDI & %PT, acidity index acid) calculated after Zelinka & Marvan. These indices work well, but give no answer on which reference diatom communities actually are typical for Swedish unimpacted streams and lakes, and deviations from those communities, a question that is required by the WFD to be answered. One problem of the IPS is that it usually returns "high ecological status" only if the stream is nutrient poor, even if it could be assumed that there also must be naturally nutrient rich streams in Sweden. The question now was if it would at all be possible to sort out the diatom communities between eutrophicated streams and natural nutrient rich streams – one hypothesis was that both might actually be the same, and the antropogenic impact would not be visible as the diatoms only rely on the nutrient level of a stream. The alternative hypothesis would be that there would be differences, largely because the anthropogenic impact usually not only leads to increased nutrient levels, but also to other changes of the water chemistry such as for example an increased amount of organic pollutants which should give advantage to heterotrophic living diatoms. The present exercise was done to study if there are clear differences between natural nutrient rich and anthropogenic eutrophicated streams, including an assessment of uncertainty of using any reference communities.

7.1 Data and analysis

From the complete collection of "all" Swedish stream diatom and environmental background data (1222 sites, national and regional monitoring programs and research projects) a set of 67 streams was selected representing nutrient rich streams (filter: total phosphorus > 50 µg/l) with either a calculated clear eutrophication impact (strongest eutrophication impact class "poor") or with a clear natural high nutrient content (zero eutrophication impact, class "high"). The method used to calculate eutrophication was the classification of status of total phosphorus from the assessment criteria for lakes and watercourses. The formula used was the simplified one using absorbance because data for base cations were missing. As most of the nutrient rich streams had more than 10% agricultural area in the catchment ref-Pjo was calculated using Pjo from the model calculations from SMED 2007 (PLC5).

The following environmental variables were used for analysis: latitude, longitude, catchment area, amount of lakes, wetlands, agriculture, meadows, forest, clear cuttings and

urbanisation, pH (annual mean), TP, TN, N-NH4, N:P (molar ratio) and conductivity. Eutrophicated streams and natural nutrient rich streams were not distributed evenly over Sweden, instead, most eutrophicated streams were in Skåne and on Gotland, and most natural rich streams in Uppland (Figure 7.1). Differences in the chemical and other background data of the eutrophicated and the naturally nutrient rich streams were tested with ANOVA and Tukey post-hoc tests. Eutrophicated streams were significantly richer in TN and TP, the nutrients were about the double in the eutrophicated streams (4000 µg TN/l vs. 2000 µg TN/l and 150 µg TP/l vs. 72 µg TP/l). Additionally, the eutrophicated streams had only half of the clear cutting area than the natural streams (4 vs. 7%), and their catchment area was on average significantly smaller (100 km² vs. 375 km²). The other environmental variables were similar.

A CA was run to see if the diatom communities of the two waterbody types would differ from each other. A CCA analysis was done to find which of the environmental variables taken into account would steer the diatom community at the most. An IndVAL analysis and a SIMPER analysis were done to see which taxa were typical for the eutrophicated respective naturally rich streams. PCOrd was used for calculations. 338 diatom taxa were used in the analyses, the relative abundance data were arcsin squareroot transformed prior to analyses.

7.2 Results

Significant differences between the diatom communities of eutrophicated and naturally nutrient rich streams were found in both IndVal and SIMPER analyses, even if the CA (Figure 7.1) and the CCA did not show a clear difference of the two water types in general.
The CCA (Table 7.1) showed that the main factors steering those communities were pH, conductivity and TN, together with the land use factors wetlands, agriculture and (negatively) forest on the first axis. The impact of lakes was important on the second axis, latitude on the third axis.

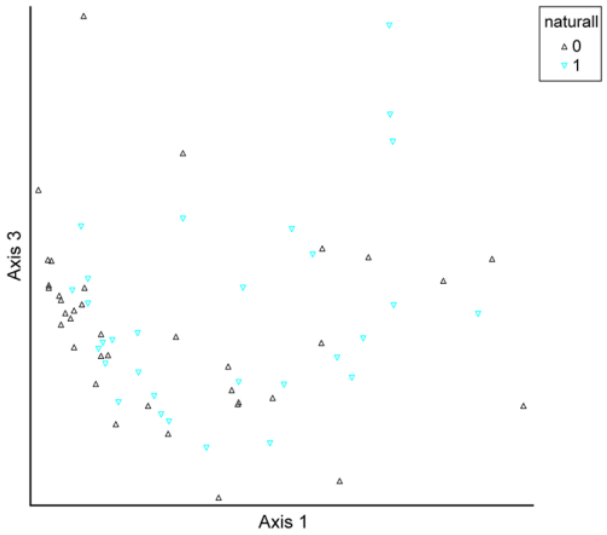


FIGURE 7.1
CA showing no differences of the diatom communities in eutrophicated (black) and natural nutrient rich (blue) streams.

TABLE 7.1
CCA Interset correlations with the environmental variables

	Axis 1	Axis 2	Axis 3
latitude	-0.437	-0.036	0.626
longitude	-0.113	0.293	0.291
catch_km2	-0.014	0.456	0.02
lake	-0.533	0.695	0.097
wetland	-0.697	-0.189	-0.547
agricult.	0.734	-0.031	-0.095
meadow	0.404	0.154	-0.23
forest	-0.66	-0.012	0.176
clearcut	-0.501	-0.06	0.341
urban	0.152	-0.048	0.082
pH	0.725	0.294	-0.069
TP	0.138	-0.107	0.074
TN	0.61	0.011	-0.403
N-NH4	0.099	-0.07	0.024
N:P (mol)	0.417	0.132	-0.371
Conduct.	0.718	-0.091	-0.153

However, both IndVal and SIMPER analyses pointed out diatom taxa that were typical for either eutrophicated or nutrient-rich streams, and a future reference community

should make use of those taxa to assess if the diatom community of a stream actually is deviating from a natural reference community typical for nutrient rich streams. Four diatom taxa (*Amphora pediculus* (Kützing) Grunow, *Navicula reichardtiana* Lange-Bertalot, *Navicula veneta* Kützing and *Reimeria sinuata* (Gregory) Kociolek & Stoermer) got high (> 20) significant indicator values and were thus typical for eutrophicated streams, whereas ten taxa got vice versa high values for naturally nutrient rich streams (*Achnanthidium minutissimum* group II (mean width 2,2-2,8µm), *Cocconeis placentula* incl. varieties Ehrenberg, *Staurosira venter* (Ehrenberg) Cleve & Moeller, *Stephanodiscus parvus* Stoermer & Håkansson, *Navicula cryptotenella* Lange-Bertalot, *Staurosira construens* var. *construens* Ehrenberg, *Meridion circulare* var. *circulare* (Greville) C.A. Agardh, *Meridion circulare* var. *constrictum* (Ralfs) Van Heurck, *Diatoma tenuis* Agardh and *Nitzschia media* Hantzsch). SIMPER pointed out *A. pediculus* as the most important taxa separating the two stream groups, with a mean abundance of 22% in eutrophicated, but of only 8% in naturally nutrient rich streams. Likewise, *A. minutissimum* and *C. placentula* were important taxa in the SIMPER analysis, with mean abundances of 16 resp. 17% in naturally nutrient rich streams, but only half that much in eutrophicated streams.

7.3 Conclusions

Diatom communities in eutrophicated and natural nutrient rich streams are not separated that easily, but there are still soma taxa which are more typically found in one of those stream types, so there is a possibility to develop an indicator telling how much a diatom community would deviate from a natural reference community, even for streams above 50 µg TP/l.. TP is not steering the diatom communities at this high nutrient level, but TN does which is expected when P is no longer limiting. However, several factors restrict the use of this analysis to a pilot one: The streams did only cover middle and southern Sweden, and natural nutrient rich and eutrophicated streams were not evenly distributed. There could be local diatom communities depending on geographical region only, but the amount of data is too poor to test this hypothesis. It would be possible to actually develop a pilot index from here to test on more streams, but it is necessary to assess for each of them if the P level is natural or enhanced, which requires a lot of data, and more work.

8 Macroalgae along the Swedish coast

Exercise summarised by Jacob Carstensen, Aarhus University and Mats Blomqvist, Hafok

We have a large dataset of macrophyte data from the entire Swedish coast, collected in different surveys and monitoring programs over the period from 2000-2012. We want to use this data to address the following broad questions, relevant for indicator development:

1) Do vegetation variables identified as potential indicators (total cumulative cover and cover of certain functional groups) show a statistical relationship with anthropogenic disturbance (eutrophication)?
2) Do these same variables show a statistical relationship with natural gradients (e.g. salinity, wave exposure, seabed substrate, slope)

The aim of the work in this group was to solve a few issues with the data, set up appropriate statistical models and run models for one or a few vegetation variables.

8.1 Data

The macrophyte data consists of diving transects, perpendicular to the shoreline. The cover of all taxa is recorded in more or less homogenous sections of the transect, which can be seen to describe different "belts" or depth zones with different species composition or dominating species. Here, we include only segments with homogenous substrate cover (>= 75% cover of soft sediment or hard substrate).

We also discarded observations from shallow depths, where physical exposure is believe to be dominant. The physical exposure can be described by means of wave exposure calculated from a hydrodynamical model and it is expressed in 7 classes ranging from ultra-sheltered to very exposed. Plots of cumulative cover versus depth for the different exposure classes were used to determine a cut-off depth, the observations above which were not used in this study. Cut-off values ranged from 0.5 m in the very sheltered areas to 7 m in the highly exposed areas.

Data on N and P concentrations and salinity are taken from the Coast Model, SMHI, which has values modelled for each coastal water body. A total of 297 of the water bodies have been investigated with at least one diving transect. The survey intensity differs strongly between water bodies, both in terms of the number of study sites and the number of years that are investigated.

Data on seabed substrate is present for each transect segment and data on wave exposure for each site.

8.2 Analysis of total cumulative cover

For each typology we ran the model

$$\log(\text{cum cover macroalgae}) = \text{area} + \text{depth} + \text{year} + \text{month}$$

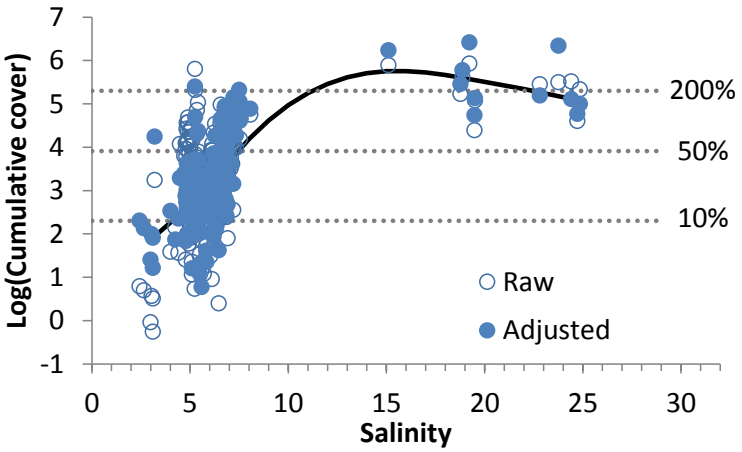
where 'area' is the waterbody-specific intercept and 'depth' describes the exponential decline in cumulative cover with depth. The factors 'year' and 'month' describe interannual variation and seasonal variation within typologies. The waterbody-specific intercepts and slopes parameters were extracted from the model and combined with

salinity, chlorophyll and nutrient levels from the Coast Model. Thus, the waterbody-specific intercepts describe the overall cumulative cover for a standard depth common to all areas, taking into account different in the months and years of sampling. The expected cumulative cover at 7 m was calculated for all areas and combined with area-specific mean salinity, chlorophyll and nutrient levels for the study period.

Expected cumulative cover (at 7 m; log-transformed) and mean salinity and nutrient levels were related using non-parametric Generalised Additive Models (GAM) that in addition to a linear component tests for higher-order relationships (smoothers). If the higher-order relationship was not significant, the relationship with the explanatory variable was reduced to a linear relationship and tested again. Through this backward elimination procedure non-significant non-parametric smoothers and linear relationships were iteratively excluded until all factors included in the model were significant.

Using the approach above it was found that the expected cumulative covers across areas could be described as a combination of a smooth relationship to summer total nitrogen (log-transformed), a linear relationship to summer chlorophyll (log-transformed) and a smooth relationship to salinity. This model explained 63 % of the total deviance.

There was an increasing relationship with salinity flattening out at salinities above 10, suggesting that there was a lower cumulative cover in the brackish Baltic Sea coastal areas, compared to the west coast, when differences in chlorophyll and total nitrogen were accounted for (Figure 8.1). The cumulative cover decreased strongly with chlorophyll a, suggesting a cumulative cover around 100 % at 7 m for areas with summer chlorophyll levels less than 1 µg/l and decreasing to less than 10 % for areas with summer chlorophyll levels above 10 µg/l. Summer chlorophyll levels were good in explaining decreasing cumulative cover along a eutrophication gradient. This gradient was also observed in the raw cumulative cover means versus total nitrogen, but when adjusted for variations in salinity and chlorophyll only a weak slightly decreasing relationship with total nitrogen remained.



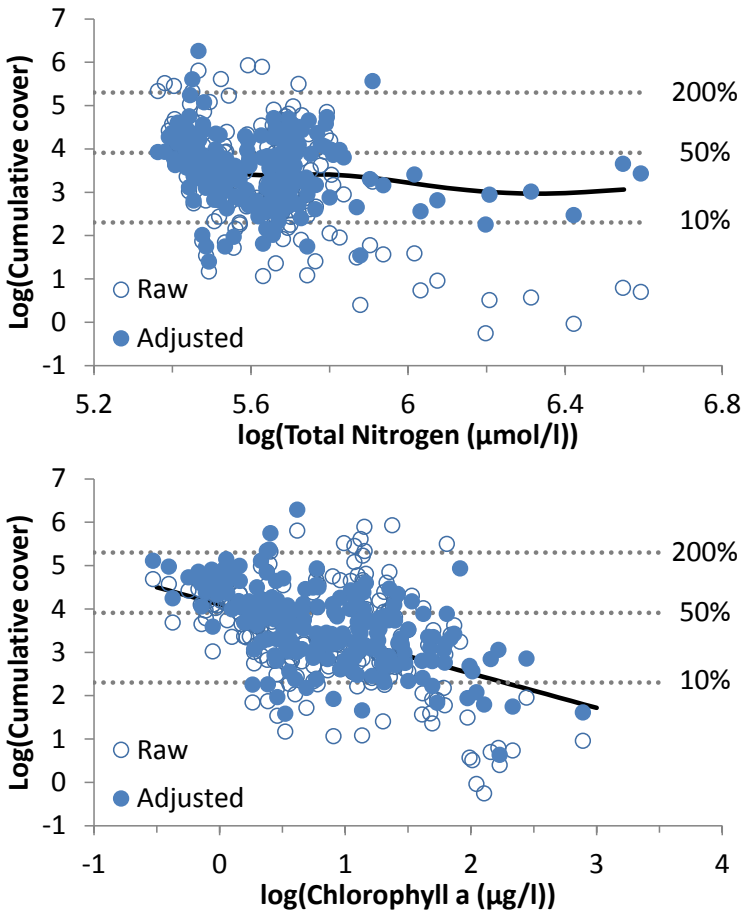


FIGURE 8.1
GAM relationships between cumulative cover (log-transformed) and salinity, chlorophyll a (log-transformed) and total nitrogen (log-transformed). Open symbols show the area-specific means, whereas filled symbols have been adjusted for variation explained by the two other factors in the model. Dotted lines indicate levels of cumulative cover on the original scale.

9 Benthic invertebrates in lakes

Exercise summarized by Simon Hallstan, Swedish University of Agricultural Sciences and Anders Grimvall, Havsmiljöinstituiet

The benthic invertebrate index MILA (Multimetric Index for Lake Acidification) is one of the indices used for assessment of acidification in Swedish lakes. MILA consists of six variables, namely the number of Gastropoda taxa, the number of Diptera taxa, the proportion Ephemeroptera, the proportion Diptera, the proportion predators, and the

acidity index AWIC. The indices are standardized and the MILA value is the mean value of the standardized values.

The purpose of this exercise was to examine the response of the six biological variables to a pH gradient, and to assess the differences in response between three ecoregions (Illies 1978).

9.1 Data and analysis

The data set used contains benthic invertebrate autumn samples (littoral kick samples) and water chemistry from monitoring programs. In total 129 lakes were sampled 1–19 times during 1990–2011. Yearly mean values of pH ranged from 4.37 to 8.46.

The analyses was made using generalized additive model (GAM) and (i) data from samples taken in 1995 (105 lakes), and (ii) mean values for the response variables for all years available.

The GAM models describe how the expected values of the response variable (E) is affected by the explanatory variables X_1, X_2, ... For this exercise, the X_1 was observed pH values (from 1995 or mean for several years), and X_2, X_3 and X_4 were indicators (binary) for the three ecoregions.

The following models were tested with the data from 1995:

(i) Response with normal distribution E(Ephemeroptera_prop) =param(X_1, X_2, X_3) + spline(pH, df=2)

(ii) Response with normal distribution E(Diptera_prop)= param(X_1, X_2, X_3) + spline(pH, df=2)

(iii) Response with Poisson distribution log(E(Gastropoda)) = param(X_1, X_2, X_3) + spline(pH, df=2)

(iv) Response with Poisson distribution log(E(Ephemeroptera)) = param(X_1, X_2, X_3) + spline(pH, df=2)

(v) Response with normal distribution E(AWIC) = param(X_1, X_2, X_3) + spline(pH, df=2)

(vi) Response with normal distribution E(Predator_prop) = param(X_1, X_2, X_3) + spline(pH, df=2)

The following models were tested with inter-annual mean values for pH and the response variables:

(vii) Response with normal distribution E(Ephemeroptera_prop) =param(X_1, X_2, X_3) + spline(pH, df=2)

(viii) Response with normal distribution E(Diptera_prop)= param(X_1, X_2, X_3) + spline(pH, df=2)

(ix) Response with Gamma distribution log(E(Gastropoda)) = param(X_1, X_2, X_3) + spline(pH, df=3)

(x) Response with Gamma distribution log(E(Ephemeroptera)) = param(X_1, X_2, X_3) + spline(pH, df=2)

(xi) Response with normal distribution E(AWIC) = param(X_1, X_2, X_3) + spline(pH, df=2)

(xii) Response with normal distribution E(Predator_prop) = param(X_1, X_2, X_3) + spline(pH, df=2)

9.2 Results

The most important results are summarized in figures 9.1-9.6 and are summarized in the following points:

- The strongest response was found for proportion Diptera, number of taxa of Ephemeroptera and number of taxa Gastropoda.
- Proportion of Ephemeroptera did not change over the pH gradient in 1995.
- The lower number of species generally found in the alpine region (#14 green line in figures) was evident in most models.
- The AWIC index might not be suitable for Swedish alpine lakes.

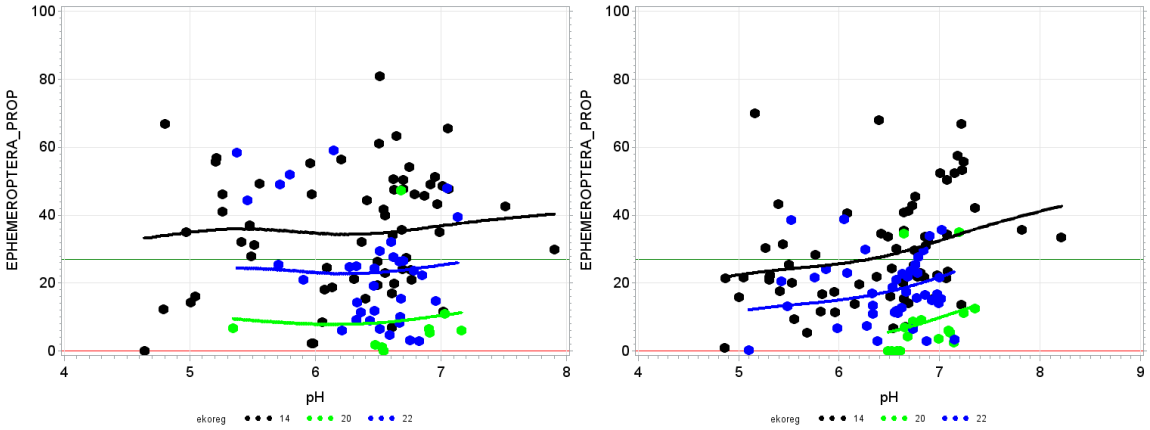


FIGURE 9.1
Proportion Ephemeroptera. 1995 (left) and mean 1990–2011 (right).

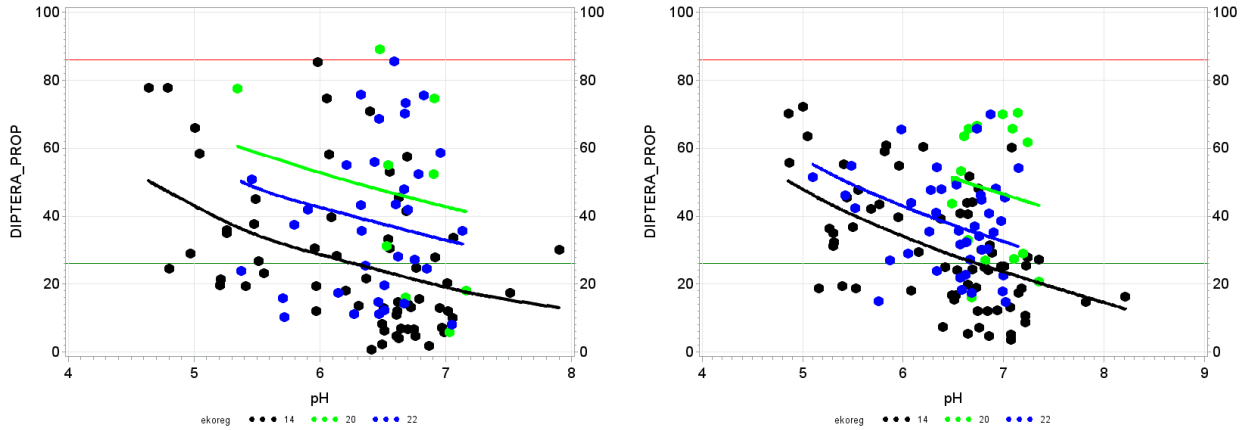


FIGURE 9.2
Proportion Diptera. 1995 (left) and mean 1990–2011 (right).

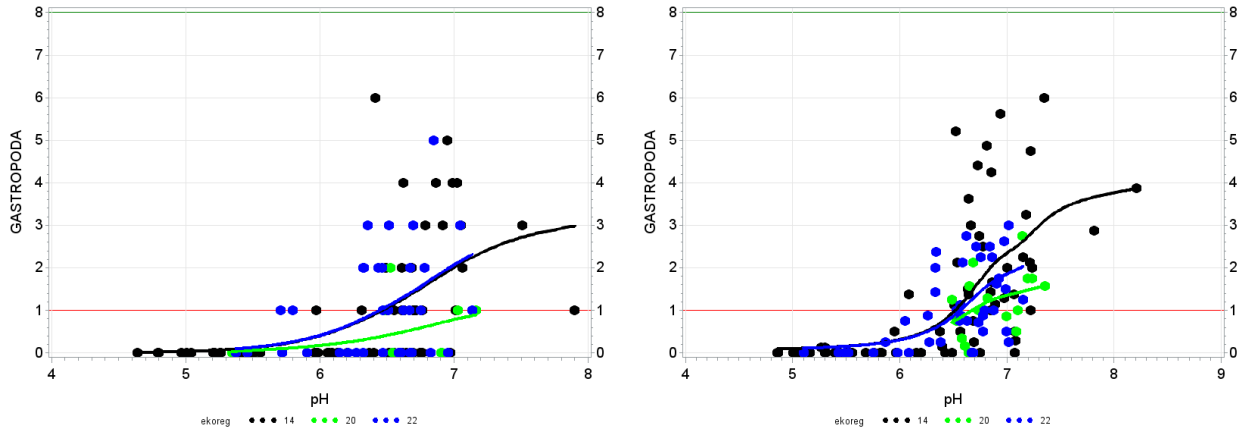


FIGURE 9.3
Number of Gastropoda taxa. 1995 (left) and mean 1990–2011 (right).

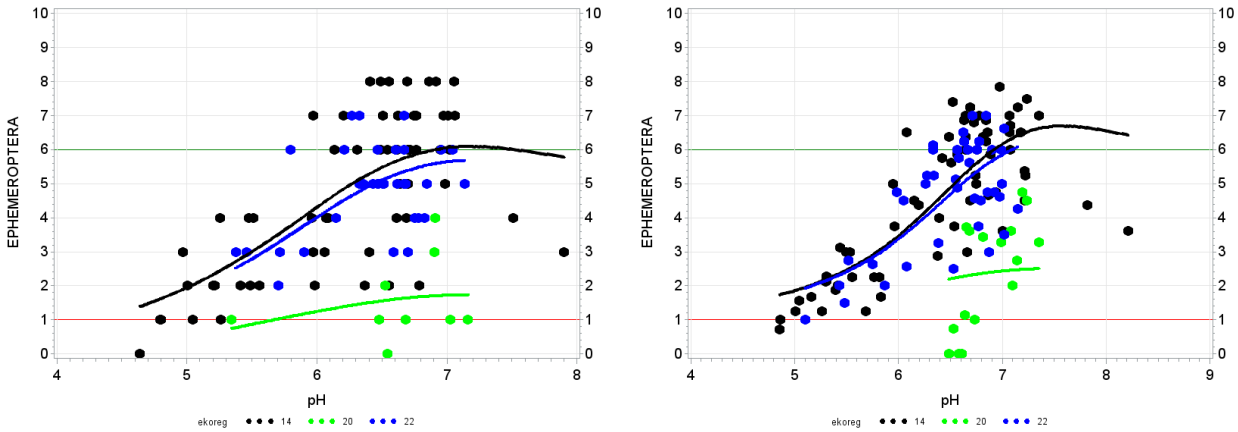


FIGURE 9.4
Number of Ephemeroptera taxa. 1995 (left) and mean 1990–2011 (right).

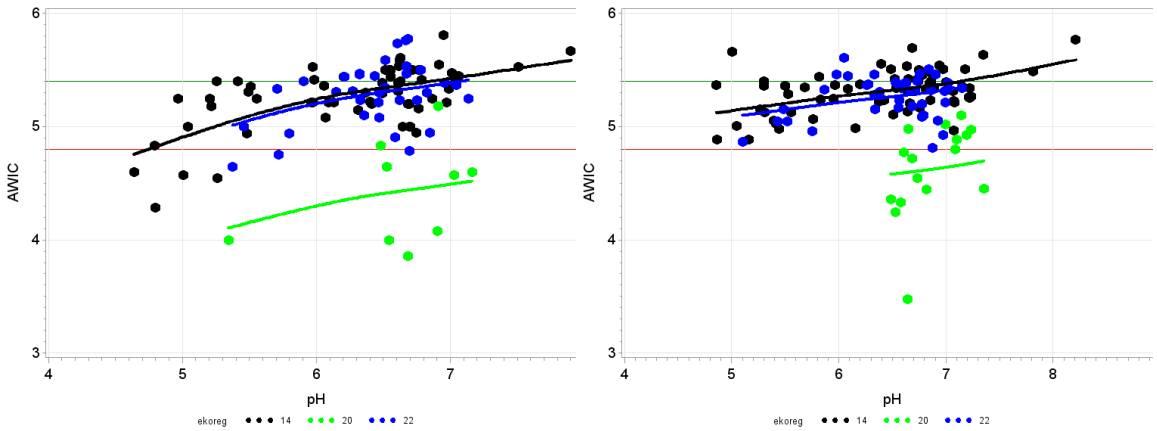


FIGURE 9.5
AWIC. 1995 (left) and mean 1990–2011 (right).

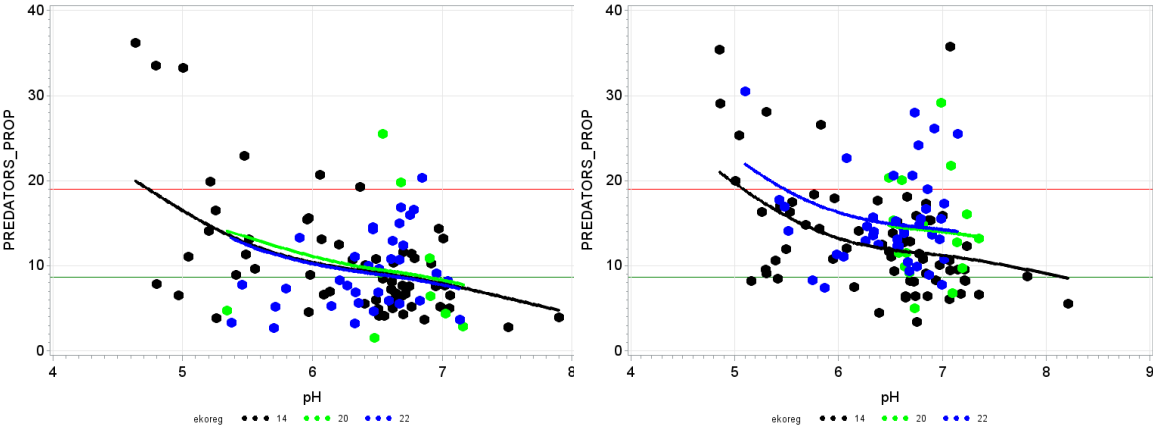


FIGURE 9.6
Proportion predators. 1995 (left) and mean 1990–2011 (right).

10 Macrophyte data cleaning

A summary by Frauke Ecke of experienced problems with macrophyte data sets

The macrophyte data that were provided by the "dataförsörjningsprogrammet" were unfortunately of low quality. This was mainly due to mistakes done by consultants that compiled the individual datasets for the respective lakes. The mistakes were first detected when evaluating the data. After identifying the most common mistakes, the entire dataset

was screened for these mistakes and the respective counties were asked to provide revised and correct data. The most common mistakes include:

1. Wrong substrate code (eight codes allowed, >20 were provided)
2. Wrong taxon code according to Dyntaxa. This mistake was most likely caused in Excel when dragging codes from one cell to the next instead of copy and paste. Dragging causes the addition of "1" to the taxon codes and results in wrong species names.
3. Wrong Swedish species names which resulted in non-matching with Dyntaxa
4. Provision of wrong X- and Y-coordinates for sampled transects
5. Missing values for distance from shore
6. Non-reporting of species that were found outside of the sampled transects

One example for wrong taxon codes is the following list of Dyntaxa codes all reported for *Phragmites australis*:

219733
219734
219735
219736
219737
219738
219739
219740
219741

However, only 219733 represent this species. All other codes are an artifact from the "dragging" mistake in Excel.

F. Ecke spend most of the workshop identifying such mistakes in the data and contacting responsible persons and/or the responsible counties and data providers.

References

Dale V.H., Beyeler S.C. (2001). Challenges in the development and use of ecological indicators. Ecological Indicators, 1, 3-10.

Noss R.F. (1990). Indicators for Monitoring Biodiversity: A Hierarchical Approach. Blackwell Synergy.

OECD (1994), Environmental Indicators: Core Set Paris, OECD.

Statistics New Zealand (2009). http://www.stats.govt.nz/surveys_and_methods/methods/indicator-guidelines/definitions.aspx.

United Nations (2007). Indicators to Measure Violence against Women. Expert Group Meeting organized by UN Division for the Advancement of Women, Geneva, United Nations, UN Economic Commission for Europe, UN Statistical Division.

USEPA (1972). Quality of life indicators: a review of state-of-the-art and guidelines derived to assist in developing environmental indicators, United States Environmental Protection Agency. Environmental Studies Division. US Government Printing Office.

Zuur, A.F., Ieno, E.N., Elphick, C.S. (2010). A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution, 1, 3-14.

Agenda for the workshop

Program for WP2.4 statistical workshop Uppsala October 2013

Wednesday 23 October 2013		
morning	Travel to Uppsala	
13:00	Welcome and practical details	Ulf G./Jacob C.
13:15	Introductory lecture about indicator development, basic concepts, and different approaches.	Ulf G.
13:45	Presentation of data and problem – solicit cases among participants	All participants
	What type of data do you have? Which are your questions or hypotheses?	
14:15	Break out groups to work on data	
14:45	Coffee break	
15:00	Continued work in break out groups	
15:30	Presentation of progress so far	
16:00	Continued work in break out groups, etc.	

Thursday 24 October 2013		
09:00	General information, short feed-back on status from groups	
09:10	Uncertainty framework (concepts, eelgrass example)	Jacob
09:40	Lecture about Data quality and data management	Ulf G.
10:00	Coffee break + Break out groups to work on data	
12:30	Lunch	
13:30	Check if groups need rearrangements. Old/new break out groups to work on data	
14:30	Coffee break	
14:45	Continued work in break out groups	
19:00	Dinner, central Uppsala	

Friday 25 October 2013 – may be subjected to changes depending on progress in BOG:s	
09:00	General information, feed-back on status from groups
09:10	Break out groups to work on data
10:00	Coffee break
12:00	Lunch
13:00	Presentations from break out groups
	Expected outcome: draft indicator and outline of report/paper
15:00	Coffee and departure

List of participants

Namn	Organisation	e-post
Ulf Grandin	SLU	ulf.grandin@slu.se
Jacob Carstensen	DMU	jac@dmu.dk
Thorsten Balsby	DMU	thba@dmu.dk
Anders Grimvall	Linköpings universitet / Havsmiljöinstitutet	anders.grimvall@liu.se
Kerstin Holmgren	SLU	kerstin.holmgren@slu.se
Lena Bergström	SLU	lena.bergstrom@slu.se
Stina Drakare	SLU	Stina.Drakare@slu.se
Maria Kahlert	SLU	maria.kahlert@slu.se
Frauke Ecke	SLU	frauke.ecke@slu.se
Richard Johnson	SLU	richard.johnson@slu.se
Simon Hallstan	SLU	simon.hallstan@slu.se
Mats Blomqvist	Hafok	mb@hafok.se

Statistical workshop on gradient studies

The third statistical workshop in WATERS was held in Uppsala from 23rd to 25th of October 2013 with the aim of indicator development and uncertainty assessment of indicators. A total of 12 persons attended the workshop. Three statistical lectures were given on principles of indicator development and uncertainty assessment. Data analysed at the workshop comprised long-term monitoring data sets and data sampled during the gradient studies in WATERS. The lectures and outcome of the data analysis are summarised in this report. The workshop was successful in bridging biological and statistical expertise, forming a solid foundation for further collaboration between disciplines within WATERS.

WATERS is coordinated by:

Havsmiljöinstitutet
Swedish Institute for the Marine Environment

WATERS is financed by:

NATUR VÅRDS VERKET
SWEDISH ENVIRONMENTAL PROTECTION AGENCY