WATERS

STATISTICAL WORKSHOP ON GRADIENT STUDIES

TJÄRNÖ, 30 JANUARY -1 FEBRUARY 2013

Jacob Carstensen, Ulf Grandin, Thorsten Balsby, Anders Grimvall, Kerstin Holmgren, Maria Kahlert, Bengt Karlson, Martin Karlsson, Dorte Krause-Jensen, Ragnar Lagergren, Mats Lindegarth, Leif Pihl, Sofia Wikström

WATERS Report no. 2013:4

WATERS Report no. 2013:4
Deliverable 2.4-2

Statistical workshop on gradient studies Tjärnö, 31 January - 1 February 2013

Jacob Carstensen, Aarhus University

Ulf Grandin, Swedish University of Agricultural Sciences

Thorsten Balsby, Aarhus University

Anders Grimvall, Swedish Institute for the Marine Environment

Kerstin Holmgren, Swedish University of Agricultural Sciences

Maria Kahlert, Swedish University of Agricultural Sciences

Bengt Karlson, Swedish Meteorological and Hydrological Institute

Martin Karlsson, Swedish University of Agricultural Sciences

Dorte Krause-Jensen, Aarhus University

Ragnar Lagergren, Länstyrelsen in Västra Götalands Län

Mats Lindegarth, University of Gothenburg

Leif Pihl, University of Gothenburg

Sofia Wikström, AquaBiota

WATERS partners:

GÖTEBORGS UNIVERSITET

Stockholms universitet

UMEÅ UNIVERSITET

AARHUS UNIVERSITY

SVERIGES VATTENEKOLOGER

Marine Monitoring RESEARCH & CONSULTING

SLU

SMHI

TOXICON AB

Hafok AB

AquaBiota WATER RESEARCH

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

WATERS Report no. 2013:4. Deliverable 2.4-2
Title: Statistical workshop on gradient studies, Tjärnö, 31 January - 1 February 2013
Publisher: Havsmiljöinstitutet/Swedish Institute for the Marine Environment,
P.O. Box 260, SE-405 30 Göteborg, Sweden
Published: July 2013

Please cite document as:
Carstensen, J., Grandin, U., Balsby, T., Grimvall, A., Holmgren, K., Kahlert, M., Karlson, Krause-Jensen, D., Karlsson, M., Lagergren, R., Lindegarth, M., Pihl, P., Wikström, S.Statistical workshop on gradient studies, Tjärnö, 31 January - 1 February 2013, Deliverable 2.4-2, WATERS Report no. 2013:4. 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 is a deliverable of one statistical workshops held in WATERS every year with participants from the research programme and representatives from the Swedish Country Administrative Boards.

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

In order to facilitate collaboration and to ensure that analyses and tools are based on adequate statistical procedures, WATERS organises a series of statistical workshops that are open to all participants and relevant authorities. The second statistical workshop in WATERS was held at the Sven Lovén Centre for Marine Research Tjärnö from 30th January to 1st of February 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 14 persons attended the workshop. Four statistical lectures were given on principles of indicator development, general linear models, generalised additive models, 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, statisticians and authorities 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.

WATERS is coordinated by:

Havsmiljöinstitutet
Swedish Institute for the Marine Environment

WATERS is financed by:

NATUR VÅRDS VERKET
SWEDISH ENVIRONMENTAL PROTECTION AGENCY

Svensk sammanfattning

För att underlätta samarbete och för att se till att alla analyser och verktyg baseras på sunda statistiska rutiner ordnar WATERS statistiska workshops för alla programmets deltagare och för berörda myndigheter. WATERS andra statistiska workshop hölls på Sven Lovén Centrum för Marina Vetenskaper på Tjärnö från den 30:e januari till den 1:a februari 2013. Syftet var att fokusera på indikatorutveckling och osäkerhetsbedömning. Vid mötet analyserades och diskuterades data från långa tidsserier från miljöövervakning och från WATERS' pågående gradientstudier. Totalt deltog fjorton personer vid workshopen. Fyra föreläsningar om principer för indikatorutveckling, generella linjära modeller, generella additiva modeller och osäkerhetshantering gavs av projektets statistiska experter. I anslutning till föreläsningarna diskuterades och analyserades data i mindre grupper bestående av datainnehavarna och de statistiska experterna. Resultaten av dessa ansträngingar rapporterades sedan till och diskuterades bland alla deltagare samt sammanfattades i denna rapport. Även om workshopen inte tillät en fullständig hantering och analys av data, etablerades samarbeten mellan biologer, statistiker och myndigheter. Dessa analyser kommer att utvecklas mer i framtiden. Mötet lyckades alltså med målsättningen att skapa länkar och samarbeten mellan olika typer av kompetenser inom WATERS.

Introduction

The second statistical workshop in WATERS was held at Tjärnö from 30th of January to 1st of February 2013 at the Sven Lovén Centre for Marine Science, which is a marine infrastructure organisation under Gothenburg University. The workshop was announced in September 2012 and final plans including the agenda were circulated within the WATERS consortium and authorities represented in WATERS reference group on November 2nd 2012. The workshop was attended by 14 people, 13 from WATERS and 1 from the County Administrative Board of Västra Götaland (Länsstyrelsen), who brought with them diverse sets of data. The majority of participants were from scientific focus area 3 (FA3), because data from the gradient studies in FA4 were not yet ready for analysis. For this reason it was decided to organise a similar workshop for the freshwater scientists during summer 2013.

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 three statistical lectures and two presentations 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.

Basic concepts of indicator development

This lecture was given by Ulf Grandin from the Swedish University of Agricultural Sciences.

There are several definitions of what an indicator is. In essence, all definitions say 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 1).

TABLE 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 Mutagenesis	Physical deformation Lesions Parasite load	Function Function Function
Species	Range expansion or contraction Extinction	Range size Number of populations	Structure Composition
Population	Abundance fluctuation Colonisation or extinction	Age or size structure Dispersal behaviour	Structure Multi
Ecosystem	Competitive exclusion Predation or parasitism Energy flow	Species richness Species evenness Number of trophic levels	Composition Composition Function
Landscape	Disturbance Succession	Fragmentation Spatial distr. of communities Persistence of habitats	Structure Structure 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 the European Environmental Agency, EEA). The different parts of the DPSIR 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 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*.

TABLE 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

Irrespectively of what type of indicator that should be developed, there are some shared characteristics that all indicators should have. 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

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 not at all indicate what was intended.

General Linear Models

This lecture was given by Thorsten Balsby from Aarhus University.

General linear model (GLM) can analyse datasets with both discreet and continuous predictor variables. Procedures for GLM are available in most major statistical programs. The GLM is based in multiple regressions where categorical variables can be included in the analyses as dummy variables. Besides ANOVA and regressions the GLM can analyse repeated measures design, analysis of covariance, and many others. It is also possible to handle random effect models, which enable analysis of mixed models. In models with multiple variables one may desire to illustrate interaction effects but note that this is done differently depending on whether the interaction involves categorical and / or continuous factors. For two categorical variables: plot means for each combination of categorical variables; interactions between a categorical and a continuous variable: draw lines for the continuous variable for each category; two continuous variables: standardize variables and draw lines for one of the variables for selected values of the other variable. The assumptions for GLM are that residuals should follow normal distribution and homogeneity of variance, and fixed variables should be measured without error or at least with smaller error than the dependent variable. If data cannot be transformed to fulfil assumptions on normality, generalised linear models (also denoted GLM) can be used if a suitable distribution can be found.

Generalised Additive Models

This lecture was given by Anders Grimvall from Havsmiljöinstituttet.

Generalized additive models (GAMs) constitute a widely applicable class of models that can be used to describe statistical relationships between a single response variable and one or more explanatory variables. For example, GAMs can be used to fit a so called spline function to a scatter-plot of XY-data. In this case, the horizontal axis is split into subintervals in which cubic polynomials are fitted to data so that they together form a response function with two continuous derivatives. Another useful application is to estimate a response function that has a common nonlinear component but exhibits different average levels of the response in different subsets of data. More generally, GAMs can be employed to fit models in which the expected response to several inputs can be written as a sum of terms in which each term is a linear or nonlinear function of single explanatory variable. In spite of the name of the model class, GAMs can also be used to examine non-additive effects of arbitrary pairs of variables. Such model components (or response surfaces) are usually called thin plate splines. When the error terms are normally distributed, GAM shall be read general additive models. However, response variables with other distributions, e.g. binary, Poisson, exponential or gamma, can be handled within the same theoretical framework, and GAM is then read generalized additive models. In the workshop, GAMs were employed to analyse catches of fish at different depths in different areas. Both SAS and R have user-friendly and reliable software procedures or packages named GAM.

Uncertainty framework

This lecture was given by Jacob Carstensen from Aarhus University and Mats Lindegarth from Gothenburg 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.

Gradients of marine hydrochemistry data

Bengt Karlson (SMHI, Oceanography) presented results from the gradient study on the Swedish west coast carried out in summer 2012. This study was funded by the Swedish Agency for Marine and Water Management. A co-operation with the sampling program of the Water Quality Association of the Bohus Coast (BVVF) made high frequent sampling possible. Sampling was made approximately every two weeks at 12 stations (Figure 1). Station Byfjorden was only sampled once a month, standard in the BVVF program.

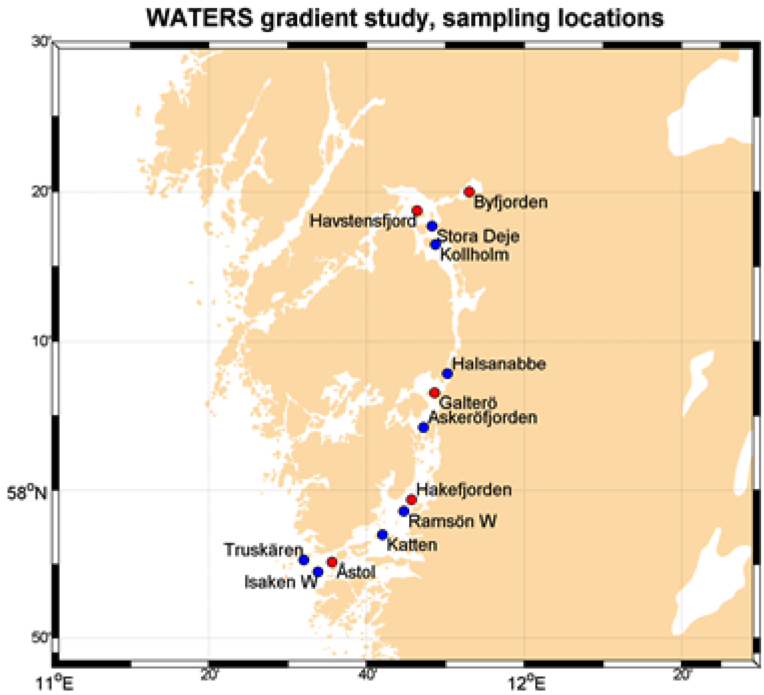


Figure 1: Map shows sampling locations for the gradient study 2012. Red dots represent standard sampling locations for the sampling programme of the Water Quality Association of the Bohus Coast and blue dots extra stations sampled in the WATERS gradient study.

One aim of the study was to investigate if a gradient in eutrophication related parameters (

Table 3) can be observed in the area. Ideally a gradient in salinity or other structuring parameters should not be present. The potential gradient in nutrient related parameters in the water mass is to be compared with data on fish, benthic macrophytes etc. Another aim is to verify that methods used is appropriate for the study.

Table 3: Parameters measured in the hydrographic and phytoplankton part of the gradient study summer 2012.

From the surface

Secchi depth

Depth profiles measured using CTDFO

Temperature

Salinity

Chlorophyll fluorescence

Oxygen

Water samples

Several depths at the main stations (red dots)

Near surface at the other stations (blue dots)

Oxygen

Phosphate

Total phosphorus

Nitrite

Nitrate

Ammonium

Total nitrogen

Silicate

Coloured Dissolved Organic Matter

Suspended Particulate Matter

Suspended Inorganic Particulate Matter

Tubes 0-10 m at main stations

Phytoplankton biomass

Phytoplankton species composition

Some preliminary results

Results from the gradient study is planned to be published in a scientific journal. Here some preliminary results are presented. Results are also presented in a report in Swedish to the Swedish Agency for Marine and Water Management.

Figure 2 shows a comparison of oxygen data from sensor on CTD vs. oxygen from Winkler method and chlorophyll fluorescence vs. chlorophyll a from water samples. Results indicate that the data from the CTDFO are really useful.

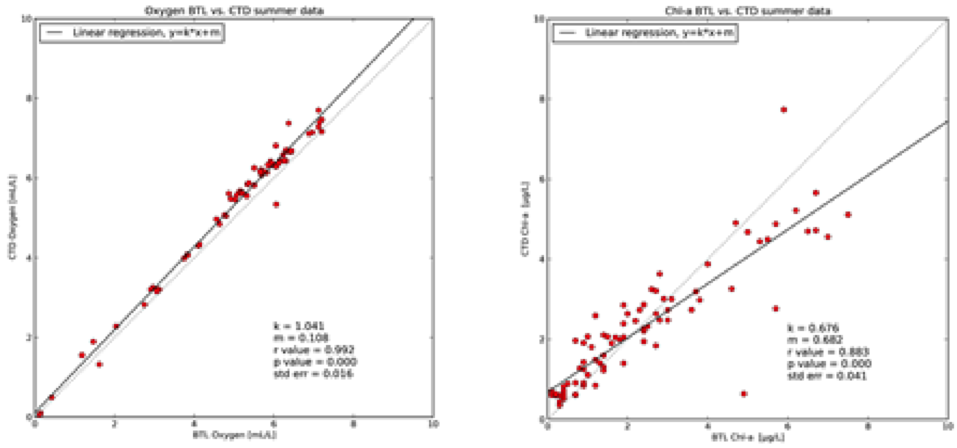
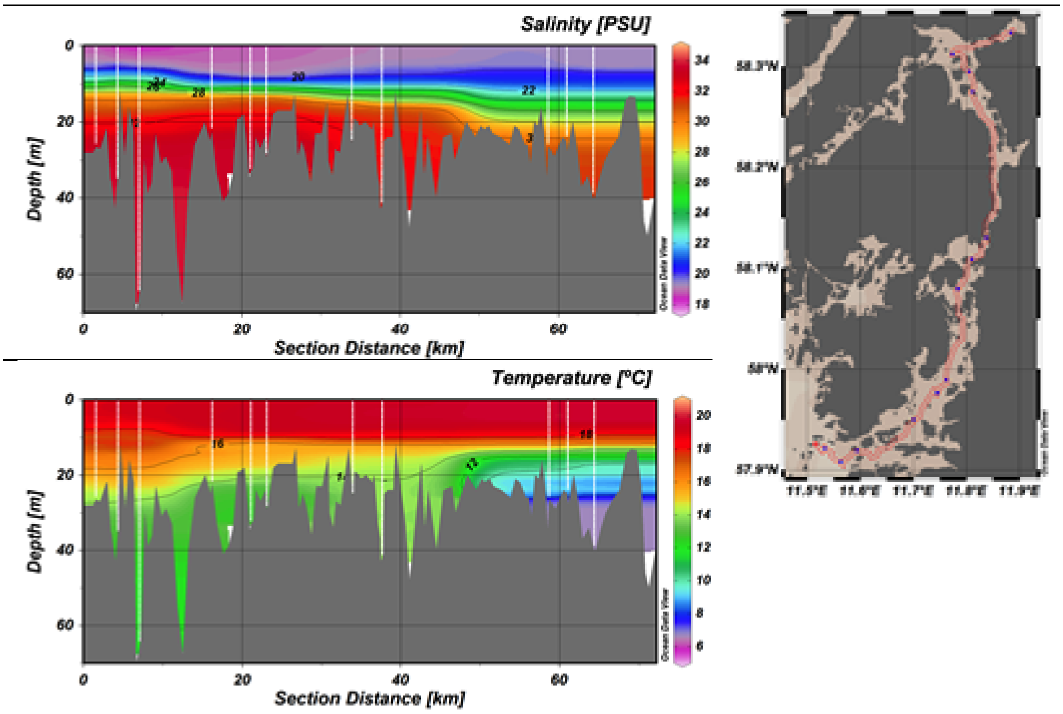


Figure 2: Left oxygen data from sensor on CTD vs. oxygen from Winkler method and right chlorophyll fluorescence from CTD vs. chlorophyll a from water samples.

Figure 3 below show the general hydrographic conditions in the area between the Marstrand fjord and the Havsten fjord. The water was strongly salinity and temperature stratified during the period June-August. Temperature stratification was strengthened at the end of period.



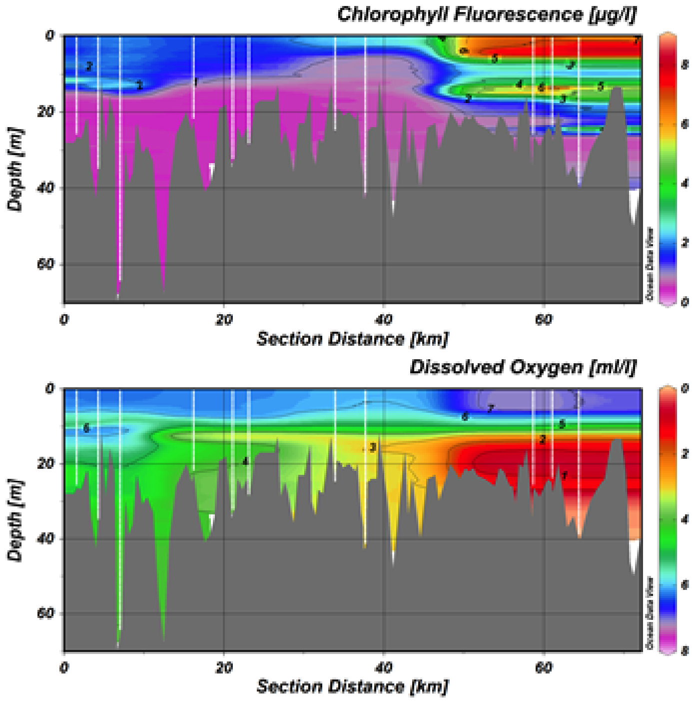


Figure 3: Graphs show the vertical distribution of salinity, temperature, chlorophyll fluorescence and oxygen on 22 August 2012. Marstrand fjord is to the left and Havsten fjord to the right. White dots indicate CTDFO-casts.

Figure 4 and Figure 5 show surface variability. Near surface salinities were lowest in the Hakö fjorden-Askerö fjord area. This was mainly observed in July. A plausible explanation may be influence from river Göta älv. Near surface silicate concentrations were high during the low saline conditions indicating riverine input. Oxygen conditions were good in the near surface water but concentrations were low in water deeper than approximately 15 m from the Askeröfjord and inwards. Chlorophyll fluorescence was used as a proxy for phytoplankton biomass. This showed high biomass 0-10 m and also thin layers of phytoplankton often found at approximately 15 m depth. A general observation is that the variability between the sampling occasions was high. This is likely to reflect short term algal blooms and short term changes in hydrographic conditions resulting from variable weather conditions.

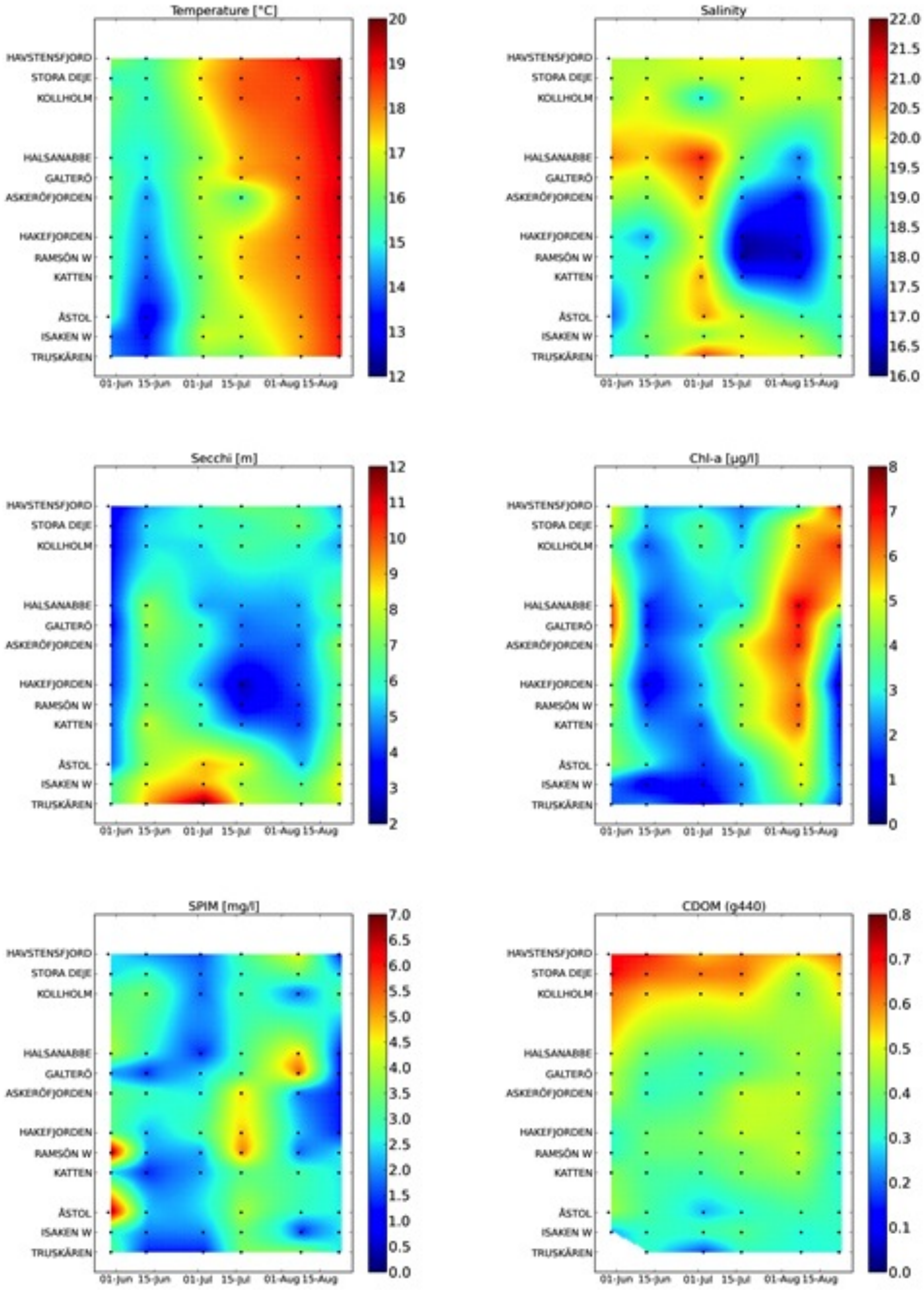


Figure 4: Graphs show surface values of selected parameters. The different parameters are described in Table 1.

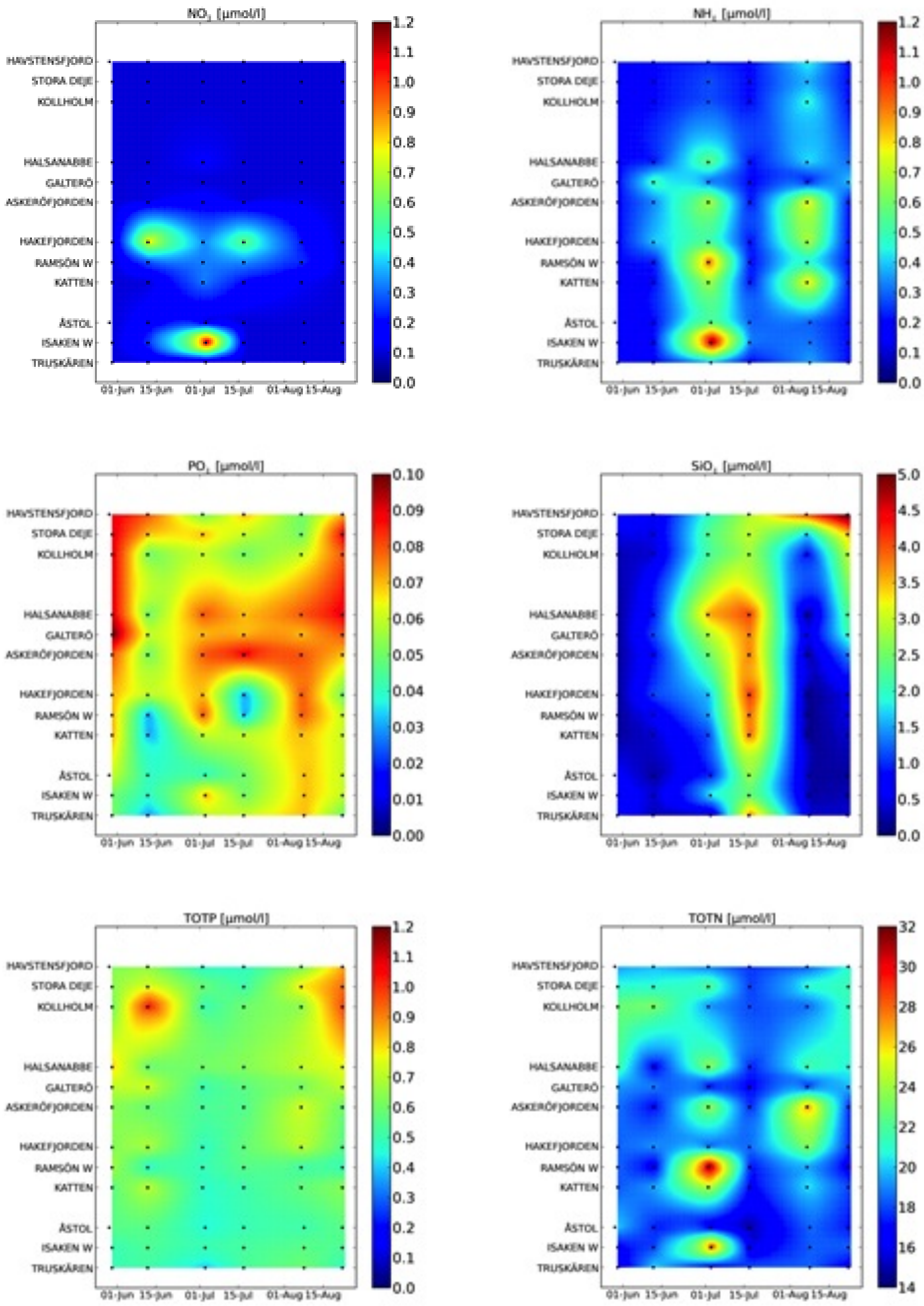


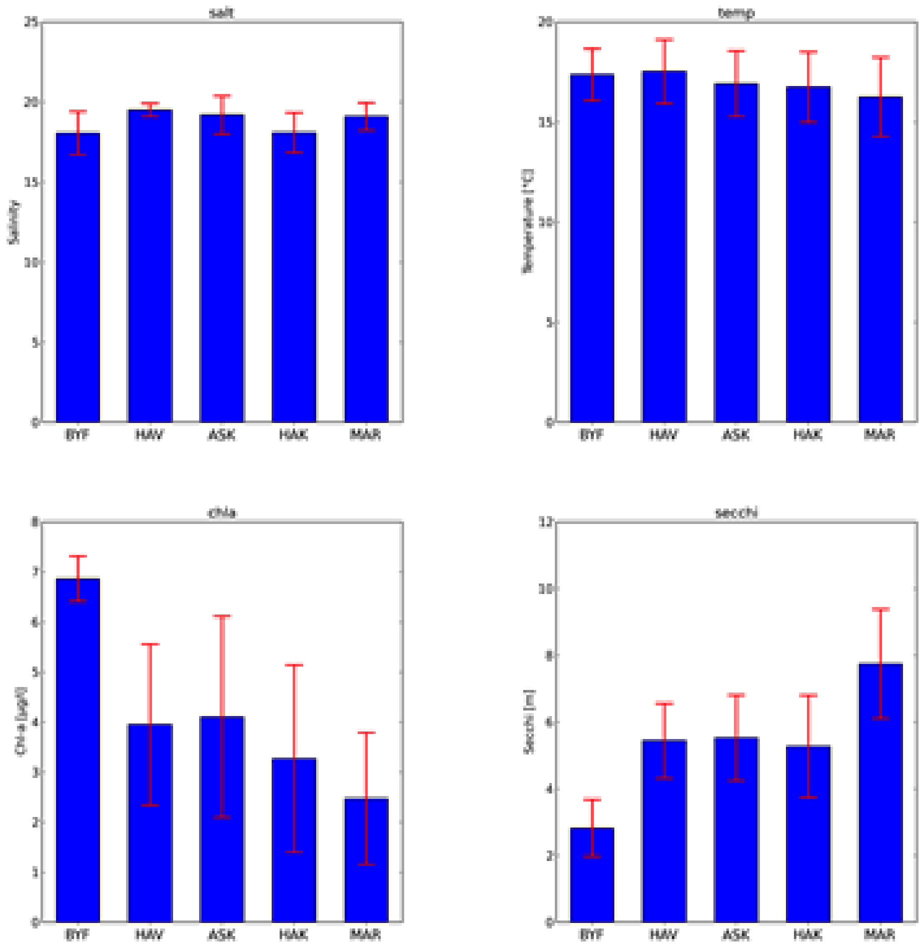
Figure 5: Graphs showing surface values of selected parameters. The different parameters are described in Table 1.

The highest values of Secchi depth were observed in the Marstrand fjord. This area is influenced by off shore water from the Baltic current. Chlorophyll data from water samples indicate that the highest biomass of phytoplankton was found in early and mid-August. Data on Coloured Dissolved Organic Matter (CDOM) show highest values in the

Havsten fjord area. This is likely to be an effect of terrestrial runoff from land with forests.

Concentrations of inorganic nutrients were in general low during the study, at least compared to winter conditions. This is expected in a summer study.

A summary of the mean values for each sea area is presented in Figure 6. Results indicate gradients in Coloured Dissolved Organic Matter, Secchi depth and in Silicate concentrations (comparing error bars, no statistical tests performed) and possibly in some other parameters.



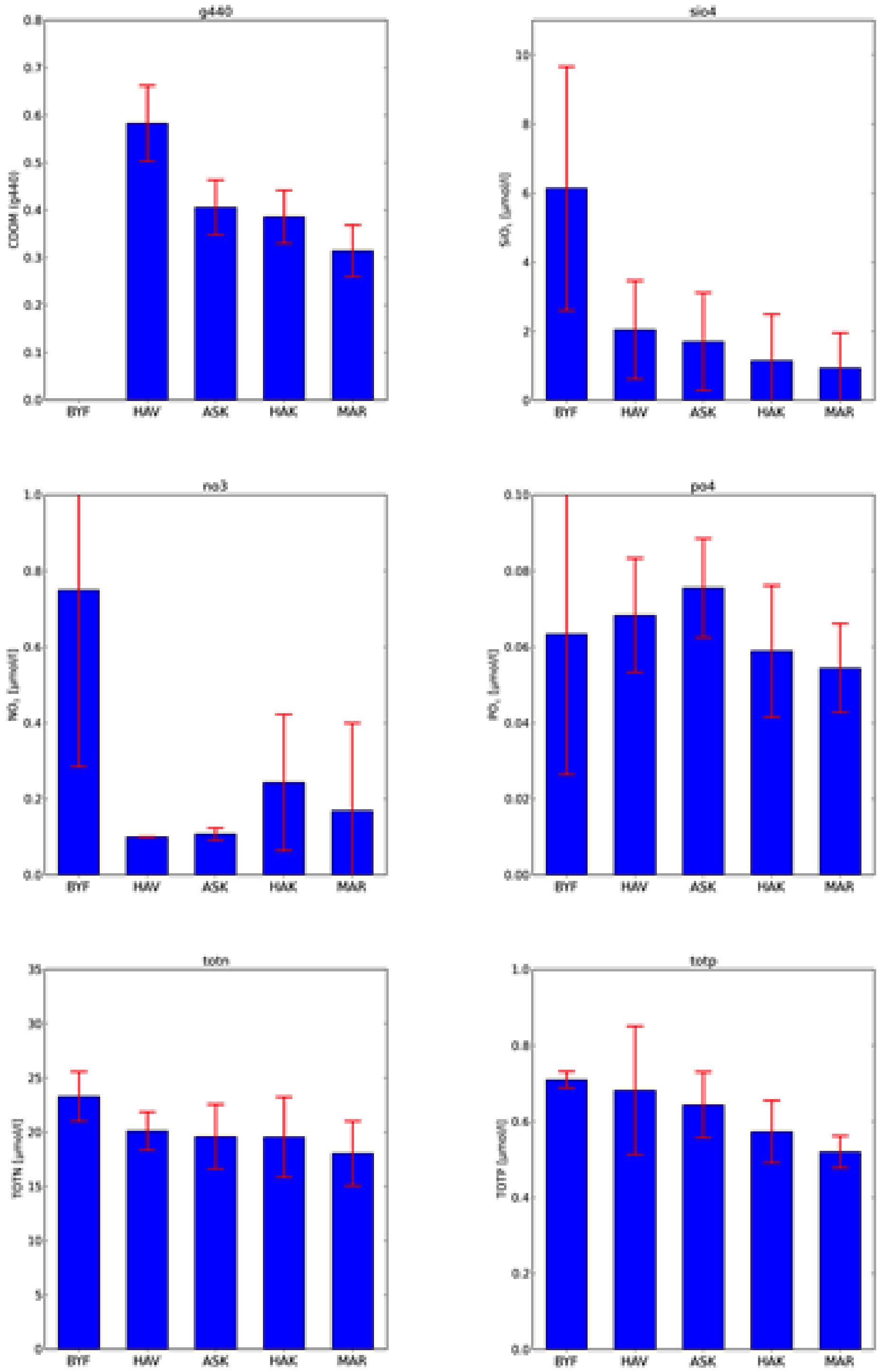


Figure 6: Graphs showing average values of selected parameters. Means represent data from six sampling occasions and three stations for each area. The different parameters are described in Table 1.

Effect of attenuating substances on Secchi depth

This exercise was summarised by Jacob Carstensen from Aarhus University, and Bengt Karlsson, Swedish Meteorological and Hydrological Institute.

During the marine gradient studies on the east and west coast of Sweden Secchi depths and different water quality variables have been measured along the expected nutrient gradient. As described in the section above there were pronounced gradients in Secchi depths, increasing from land towards the sea. However, a key question is: what is causing this gradient in water transparency?

To address this question it is first assumed that the Secchi depth represents ~10% of the surface light (although this assumption is not critical, as will be discussed later), and that light is attenuated with depth according to the Lambert-Beer equation

$I = I_0 \exp(-k_d*z)$, where k_d is the light attenuation coefficient.

Using the assumption for Secchi depth (z_{SD}) it is found that

$z_{SD}=-\ln(0.1)/k_d$, where k_d varies with the concentrations of attenuating substances in the water, i.e. dissolved organic matter (DOM) absorbs light, suspended particulate organic matter (SPOM) absorbs and scatter light, and suspended particulate inorganic matter (SPIM) scatter light. Although the effect of absorption and scattering are different the overall effect on light attenuation can be approximated to a reasonable degree by

$k_d=k_0 + k_{DOM}*DOM + k_{SPOM}*SPOM + k_{SPIM}*SPIM$

where k_0 is the background attenuation by water and other substances not included in the other components, k_{DOM} is the DOM-specific attenuation coefficient, k_{SPOM} is the SPOM-specific attenuation coefficient, and k_{SPIM} is the SPIM-specific attenuation coefficient. Both gradient studies have measured DOM (as absorbance at 440 nm), SPOM and SPIM in discrete water samples simultaneously with Secchi depth. For describing the variation in Secchi depths the average concentrations of DOM, SPOM and SPIM in the top 5 m water column were calculated.

The equation for z_{SD} with the equation for k_d inserted constitute a non-linear regression model that can be solved by non-linear ordinary least squares regression (e.g. PROC MODEL in SAS or nls() in R). Applying this non-linear regression model to the data from the east and west coast gradient studies separately resulted in deviating parameter estimates (Table 4). The parameters from the west coast gradient study were all significant, whereas the parameter for DOM at the east coast gradient study was close to zero and not significant.

TABLE 4: PARAMETER ESTIMATES FROM NON-LINEAR MODEL RELATING SECCHI DEPTH TO CONCENTRATIONS OF ATTENUATING SUBSTANCES IN THE WATER COLUMN. P-VALUES ARE THE PROBABILITIES THAT THE ESTIMATE IS EQUAL ZERO. NUMBER OF OBSERVATIONS WERE N=70 AT THE WEST COAST AND N=17 AT THE EAST COAST.

Model	West coast		East coast	
parameter	Estimate	P-value	Estimate	P-value
k_0	0.064964	0.0308	0.182936	0.0002
k_{DOM}	0.471983	<.0001	-0.02616	0.8649
k_{SPOM}	0.022348	0.0197	0.548543	0.0020
k_{SPIM}	0.025195	0.0065	0.219932	<.0001

In order to further analyse the deviating parameter estimates, scatter plots of the three explanatory variables were examined for the two gradient studies, showing that strong correlations between the attenuating substances were present in the data from the east coast gradient study, whereas correlations in data from the west coast were substantially smaller (Figure 7). Obviously, it was not possible to estimate independent relationships for the attenuating substances on the east coast.

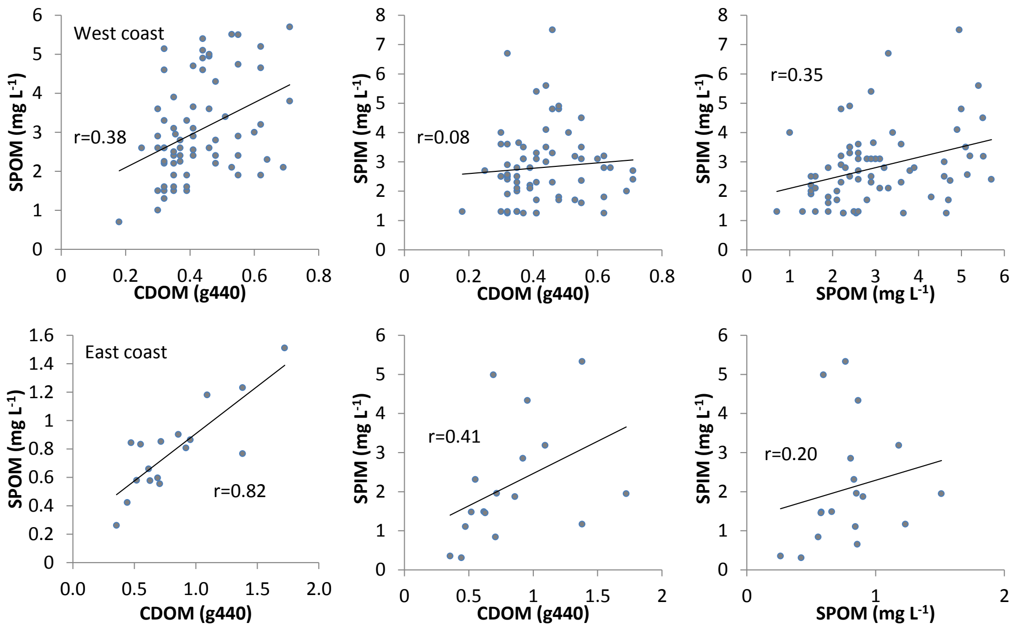


Figure 7: Correlations between CDOM, SPOM and SPIM for the two gradient studies: West coast (top) and east coast (bottom). Correlation coefficients are listed for each plot.

Using the k_d-relation above with the concentrations of the different attenuating substances measured at different stations on the west coast, a pronounced pattern of light attenuation is found showing increasing attenuation from the Skagerrak towards Byfjorden (Figure 8). This gradient of increasing light attenuation is mainly caused by increases in the CDOM concentration, increasing its relative proportion to light attenuation from 45% to almost 60%. These results are consistent with the general perception of CDOM contributing most to light attenuation along the Swedish coast.

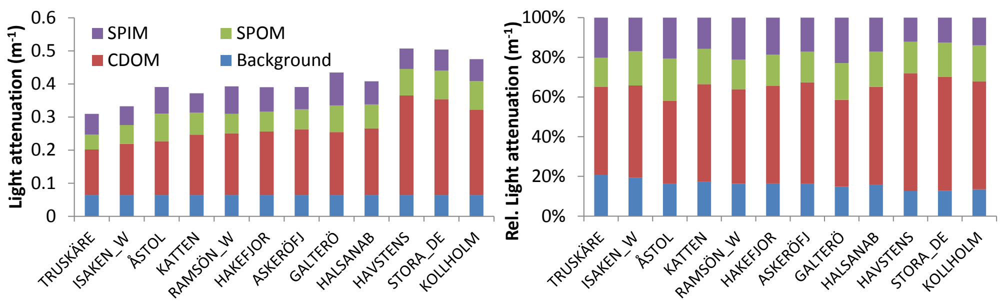


Figure 8: The estimated attenuation of light by different substances at the stations in the west coast gradient study (left) and their estimated proportion to the total light attenuation (right). Stations are ordered along a gradient from the open sea to Byfjorden.

Fish in 45 lakes

This exercise was summarised by Thorsten Balsby from Aarhus University, and Kerstin Holmgren, Swedish University of Agricultural Sciences.

For assessment of ecological status, according to the Water Framework Directive, the lake fish community should be sampled one or more times during each six-year water management cycle. Fish communities of small to intermediate sized Swedish lakes are sampled using benthic, multi-mesh Nordic gillnets, according to a European standard method (EN 14757). Sampling period is fixed to the late summer period (in Sweden from mid-July to August), when deeper lakes are thermally stratified. The recommended or default sampling effort (number of nets) increases with area and maximum depth of the lake. Nets are 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, 35-50 m, 50-75 m and > 75 m. Originally, the lake-specific recommended sampling effort was intended to give an acceptable level of precision, e.g. for detecting differences in abundance or biomass of dominating fish species, between different years in the same lake. The following exercise estimated different sources of sampling variance, for exploration of alternative sampling designs within six-year water management cycles.

The analysis aims at estimating the variance contributions of year-to-year variation, depth zone variation, variation in number of nets used for estimating various indicators of fish communities in freshwater lakes in throughout Sweden. The indicators that we used in these analyses were total fish biomass (g per benthic gillnet) and abundance (number of fish per benthic gillnet). Ultimately the analysis could device better ways to optimize monitoring efforts used for evaluating the status of the fish stock in lakes.

Description of data and analysis

The dataset contained fish catches in each benthic gillnet, for 45 lakes between 2007 and 2012. Catches were aggregated over all fish species, and given as either total fish biomass (g) or total number of fish caught in each net. In 17 lakes samples were collected in multiple years, 15 lakes with annual samples and two lakes sampled twice. In most lakes that were sampled in multiple years the nets were set at semi-permanent sites, i.e. more or less replicated between years. Net positions were, however, numbered in the order nets were set each year and within-lakes sites were not necessarily sampled in the same order each year. Therefore, the current dataset does not permit estimation of within-site variance between years. All samples were taken in July and August.

We analysed data for each separate lake and for all the 45 lakes combined.

Mixed model was used to estimate the variance contributions of each random variable (random factors in *CAPITAL* letters and fixed in *lowercase*). We assumed that residuals followed a normal distribution. The full model used

Response parameter = μ + lake + YEAR + DEPTH + REPLICATES (eq. 1)

Each net acted as a replicate for a lake and is estimated as the residual variance. For the site specific models several of the lakes were only sampled once during the six–year period and for those lakes the variance contribution could not be estimated for year. Likewise some shallow lakes only had 1 depth zone, which also required a modification of the model.

The variance estimates for the combined model could be used to estimate the total variance under different allocations of monitoring effort. In the estimation of the total variance within a six-year period we have to account for the possibility that variance have been sampled in all six years: $V[\bar{y}] = \frac{s_Y^2 * (1-\frac{a}{6})}{a} + \frac{s_{Depth}^2}{b} + \frac{s_e^2}{abn}$

(eq.2)

where *a*, *b* and *n* are the number of sampled years, depths and replicates respectively. Additional fixed variables might further reduce the variation between lakes, e.g. altitude, average air temperature (1961-1990), and freshwater eco region. However, altitude would usually be unique for each lake. As none of these variables varied between years within lake in the site specific model, these variables were not included in the model.

Results

In this analysis the variance contribution of the model parameters was assessed based with regard to biomass and number of fish caught per net. The overall model for biomass showed huge variance contributions for all model parameters (Table 5), indicating that getting estimates of biomass was associated with much uncertainty. There were huge variations in biomass between nets in different depth strata and between nets in general whereas year to year variation contributed with a smaller proportion of the variance.

The overall model for number of fish caught resulted in smaller estimates of variance for each of the model parameters than for the biomass (Table 5).

Table 5: Variance estimates for the overall models for biomass (g) and number of fish caught per net.

Parameter	Variance for Biomass	Variance for number
Year	1081	14.4
Depth stratum	299691	156.3
Residual	731503	1222.4

In the following we use the variance estimates for number of fish caught to estimate the effect of monitoring schemes. Variance is calculated using the estimates from Table 5 and equation 1. The variance is calculated for all combinations of: 8, 16 or 24 nets, for 1, 2, 3 depth zones and for 1, 3 or 6 years. As most lakes have several depth strata the effect of year and number of nets in a lake is illustrated for 3 depth strata (Figure 9).

The figure suggests that the effect of increasing the number of years of monitoring reduces variance more than using more nets. If monitoring was only done in one year the overall variance varied between 81 and 115, whereas if monitoring was done for 3 years within the 6 year period the variance varied between 60 and 71. The differences in variance between monitoring for 3 compared to 6 years only reduce the variance with 6 to 15 for a given number of nets.

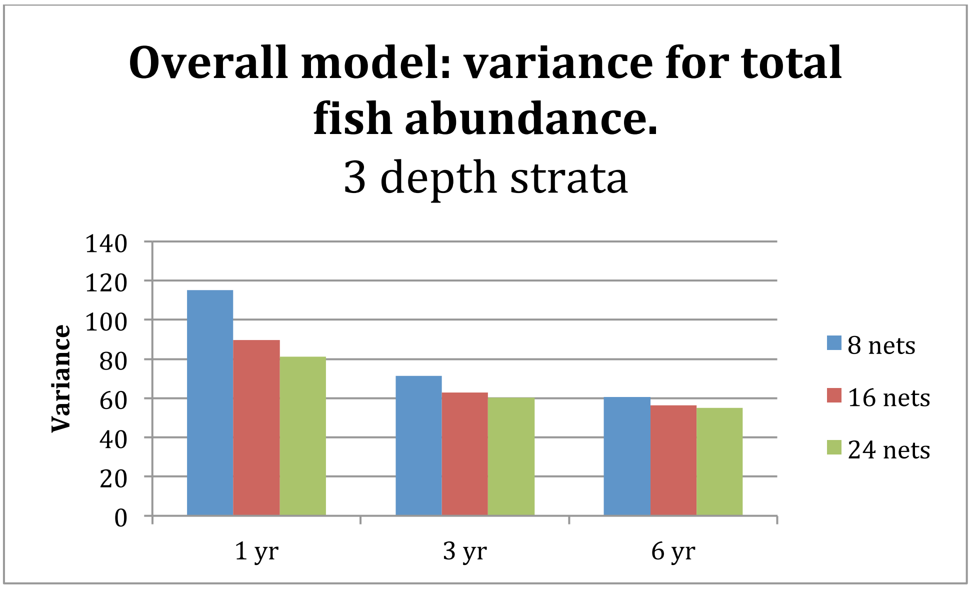


Figure 9: Estimates of total variance in fish abundance (numbers per benthic gillnet), for the overall model for combinations of sample size (number of nets per sampling occasion) and sampling frequency (years per six-year management cycle).

Conclusions from fish analysis

This exercise illustrated that finding an optimal sampling design representing a six-year period is a quite different task than optimal sampling for detecting differences between two specific years (e.g. before and after some restoration treatment) or for monitoring of long-term trends. By using data from lakes sampled during a six-year period, including a subset of lakes sampled for long-term trends, we could estimate different sources of variance. Estimated variances from an overall model, showed that fish abundance during a six-year period, rather than in specific years, might be estimated with higher precision by allocating the same total effort in three different years compared to setting all nets in a single year.

Diatoms (microphytobenthos) in lakes and streams

This exercise was summarised by Maria Kahlert, Swedish University of Agricultural Sciences.

The objective of this study was to 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 pristine streams and lakes, and deviations from those communities, a question that is required by the WFD to be answered. Therefore, the present exercise was done to do a first analysis to find Swedish reference communities, and to study how clearly they would be separated from impacted ones (matter of uncertainty of assessing any deviation).

Furthermore, we wanted to investigate if today's diatom indices, developed mainly for streams, can be used in lakes as well, i.e. do their responses to environmental variables differ between streams and lakes? There is no diatom index for lakes, and one simple way until new methods are developed would be to use the existent stream indices in lakes, if they respond in a similar way to the stressors in question. Therefore, the present exercise was done to test if there were significant differences in the response.

Data and analysis

The complete collection of "all" Swedish stream diatom and environmental background data was used for this analysis, from national and regional monitoring programs and from research projects; additionally data from lakes collected in a PhD study by Steffi Gottschalk were used. The stream data included 1142 streams with 51 environmental variables, and 100 lakes.

Analytical approach. With NPMANOVA and ANOSIM was tested if there were significant different diatom flora in the seven Swedish ecoregions. After that analysis some sites had to be removed as they turned out to be outliers (lake sites in stream dataset, single types without replicates), the analysis was repeated. A SIMPER analysis and an IndVAL analysis were done to see which taxa were typical for the different ecoregions. It was also tested with K-mean clustering to let the diatom community composition structure the outcoming groups (7 groups chosen; Figure 10). The software PAST was used for calculations.

IPS was correlated versus Tot-P, and ACID versus pH for streams and lakes, and the correlations were compared using GLM with STATISTICA.

Outcome of the statistical analysis. The *Achnanthidium minutissimum* group needed to be taken out as it was everywhere and very abundant. All ecoregions of Sweden have significant different diatom communities. K-mean clusters added up mainly on a CA's first axis. It was not possible to explain the results for the ecoregions with background variables in the short time available.

ACID was not significantly different between streams and lakes when sites > pH 8.4 were taken out of the analysis (bias by lakes). IPS was different, but the part of variation added by this difference was very much smaller than the variation explained by the environmental variable Tot-P. It must also be born in mind that IPS is not only explained by P, but also by organic pollution, where we did not have data to test.

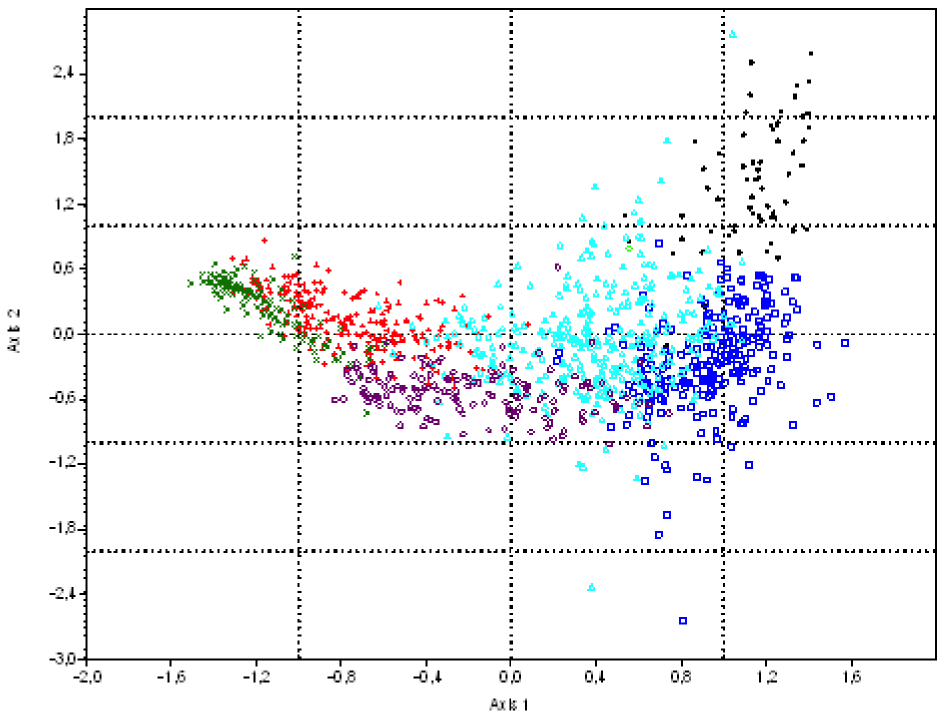


Figure 10: K-mean clusters extracted from Swedish diatom communities in streams (7 clusters), plotted in CA.

Macroalgae along the Swedish coast

This exercise was summarised by Sofia Wikström from AquaBiota and Dorte Krause-Jensen and Jacob Carstensen from Aarhus University.

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.

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).

Data on N and P concentrations and salinity are taken from the Coast Model, SMHI, which has values modelled for each coastal waterbody. A total of 284 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.

Analysis of total cumulative cover

The analysis was done in two steps. First, we established a model for the decrease of total cumulative cover with depth for each waterbody. We wanted to exclude data from the uppermost part of the zonation, where the cover is likely set by physical disturbance rather than light availability, and only model the decline in cover from the depth of maximum cumulative cover. In order to do that, we checked plots of total cumulative cover against depth for each waterbody. The peak depth was typically observed between 0-3 m across water bodies. We assumed that these differences could be explained by differences in wave exposure, but we did not investigate this in further detail since the focus was on relating macroalgae cover to eutrophication. Consequently, in order to reduce the specific influence of wave exposure on the data we excluded all data from <3 m depth. This data restriction can be further refined by identifying the peak depth specific to each waterbody, and relate this to data on physical exposure.

With this data set, we ran the model

log (cum cover macroalgae) = area + area*depth

where area is the waterbody-specific intercept and area*depth is the waterbody-specific slope. These parameters (waterbody-specific intercepts and slopes) were extracted from the model and combined with nutrient levels from the Coast Model.

In the next step, the slope from this model was tested against summer total N. We hypothesised that the slope would be steeper with decreasing Secchi depth, but we do not have Secchi depth recordings for all water bodies. We know that Secchi depth depends on chlorophyll concentrations, which are connected to nutrient concentrations, but also to other factors such as POM and CDOM. There was a weak but significant relationship (R^2=0.0719; p=0.0143) between the slope and total N (log-transformed) (Figure 11A).

We further tested the intercept against salinity. We hypothesised that since the total cumulative cover is calculated as the sum of individual cover recordings, the intercept should be positively correlated with species diversity and thus with salinity. As predicted, there was a positive correlation between intercept and salinity (R^2=0.0814; p=0.0061) (Figure 11B), but similar to the regression for the slopes this correlation was also relatively weak and there was a lot of scatter. Another issue is that the salinity data cluster in two groups, one from the east coast (salinities of 2.5-8) and one from the west coast (salinities of 25-28), and consequently the regressions assumes linearity between these two clusters.

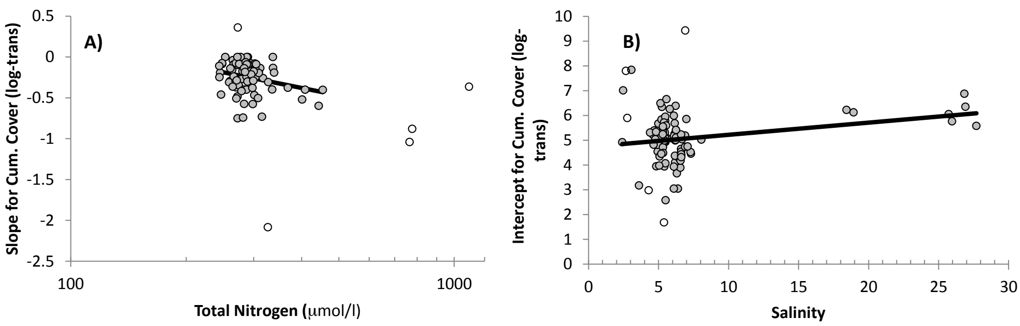


Figure 11: A) Slope for cumulative cover (log-trans) decrease with depth versus TN and B) intercept for cumulative cover versus salinity. Five observations were not included in the analysis as they were considered outliers (2 observations) or highly influential on the slope-regression versus TN.

Analysis of censored Secchi depths

Despite comparatively large uncertainties associated with Secchi disk readings (e.g. interpretation bias); this simple measurement is still frequently used to assess water quality in marine ecosystems and lakes. The Secchi disk reading yields a quantitative estimate of a single observable optical property, a combined measure of the beam attenuation coefficient and the diffuse attenuation coefficient of the medium. Water transparency measured by Secchi depth is such a fundamental monitoring variable, encapsulating several aspect of eutrophication, that it has been provocatively proposed as the only measure needed to assess ecological status for lakes (Peeters et al. 2009 in Carstensen 2010). However, Secchi depths should be considered a proxy for eutrophication where the cause-effect is still yet left unknown.

Problems that can be addressed to Secchi disk readings

Various surveys in shallow coastal ecosystems that are likely to have a coupling or a response to water transparency (e.g. macrophytes and fish) are inevitably going to face the problem with the Secchi disk being visible at the bottom, meaning that it is underestimating the actual Secchi depth. Secchi depth (SD) equal to the bottom depth (BD) provides partial information on the actual Secchi depth ($SD \geq BD$), i.e., the real Secchi depth would have been larger if not limited by the bottom depth. In statistics, this is termed censoring. However, although these statistical methods have existed for a long time, apparently they haven't yet penetrated the aquatic ecology science, where such a data often are discriminated and only finite or "true" values are considered.

Secchi depth measurements along a typical Swedish west coast fjord from 0 to 10 metre depths generates around 8 % "true" values in the stratum 0-6 m and over 90 % "true" values in the 6-10 m stratum.

Table 6: Example of data from WATERS gradient study

AREA	STATION	DEPTH	STRATA	SECCHI	CENSORED
Byfjorden	1	0,7	0-6	*	1
Byfjorden	2	2,7	0-6	*	1
Byfjorden	3	4	0-6	3	0
Byfjorden	4	1	0-6	*	1
Byfjorden	5	4	0-6	*	1
Byfjorden	6	2	0-6	*	1
Byfjorden	7	1,7	0-6	*	1
Byfjorden	8	0,7	0-6	*	1
Byfjorden	9	1,8	0-6	*	1
Byfjorden	10	0,9	0-6	*	1
Byfjorden	11	1,7	0-6	*	1
Byfjorden	12	3,9	0-6	3	0
Byfjorden	13	1	0-6	*	1
Byfjorden	14	5,8	0-6	5,5	0
Byfjorden	15	1,7	0-6	*	1
Byfjorden	16	4	0-6	3	0
Byfjorden	17	2,5	0-6	*	1
Byfjorden	18	6	0-6	4,5	0

How to handle censored Secchi depths?

Statistical analysis of censored data can be performed using methods developed for so-called survival analysis. In the case of Secchi depths some of the depth records are measured values, while others are greater-than-or-equal values. The latter occurs, when the disc is visible at the bottom of the sampling site.

If the Secchi depth records shall be interpreted as measures of light attention it is necessary to take into that the depth records are censored. This can be achieved by organizing the input to the statistical analysis into two columns. The first column contains a depth record that is either a measured Secchi depth or the maximum depth at the sampling site. The second column merely indicates whether or not the depth is a measured depth or a censored value. An optional third column can be used to indicate different strata of sampling sites.

A standard survival analysis of such a data set produces an output dataset of depth records that are either measured Secchi depths or estimates of what the Secchi depth would have been if the maximum depth at the sampling site would have been sufficiently large. Like any other statistical methods the estimation is based on some assumptions. Unless otherwise stated that all observations are statistically independent and that, within each stratum, the true (non-censored) values would be normally distributed.

In the software package SAS, a standard survival analysis can be performed using `proc lifereg`, and a sample code that is applicable to right-censored values can be written as follows:

```
Proc lifereg data=tjarno.secchi;
   Class area;
   Model secchi_censored*censored(1)=area /distribution=normal;
   Output out=tjarno.secchi_estimates P=pred std_err=standard_error;
Run;
```

Here, the variable area is used to define different strata. The variable secchi_censored contains all depth records and the variable censored was set to 1 for all censored values and 0 for non-censored values. The output dataset contains estimated (predicted) Secchi depths and standard errors of the predictions.

R has a survival package that can do the same analysis as `proc lifereg` in SAS. Further information can be found on the following link: http://www.ddiez.com/teac/surv/R_survival.pdf.

Results

When area-specific mean values (including censored data) of Secchi depths have been calculated for each strata these are weighted with respect to total number of stations. The outcome of the analysis is shown in Figure 12.

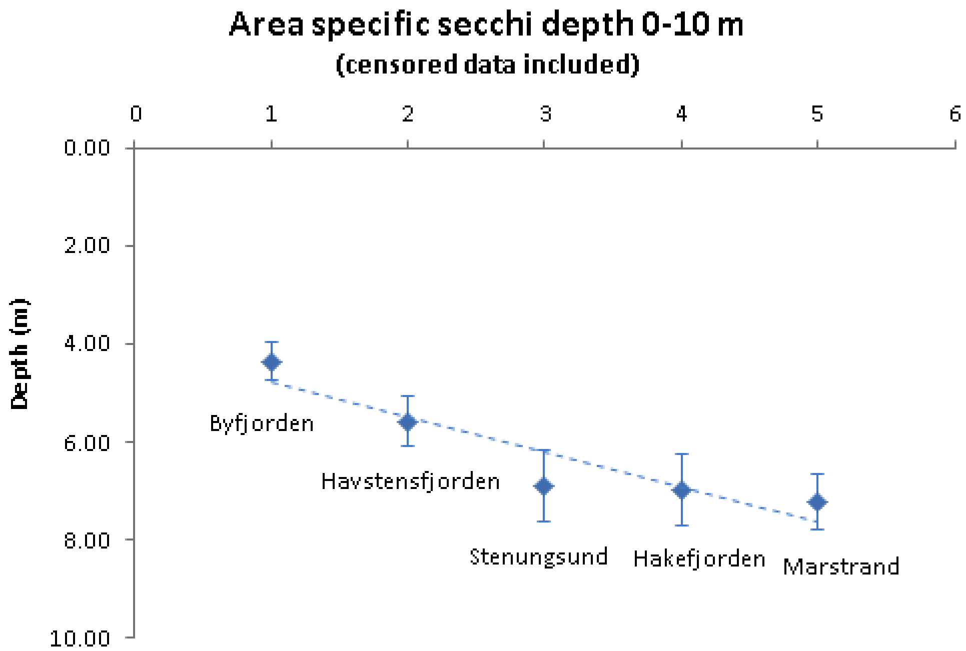


Figure 12: Area-specific means of Secchi depths estimated by means of censored data regression or survival analysis.

Uncertainty components in stream diatom monitoring

This exercise was summarised by Mats Lindegarth from Gothenburg University, and Ragnar Lagergren, Länstyrelsen in Västra Götalands Län.

The River Basin District Authorities (RBDA) and County Administration Boards (CAB) have identified needs for a system assessing the confidence in classifications. In response to this the RBDA's have developed a tool for confidence assessment (River Basin District Authorities 2013). This tool is based on availability of biological data, pressure data and concepts related to (but not identical to) the definitions of uncertainty from the WFD (i.e. precision and confidence). WATERS, on the other hand has recently developed and published a comprehensive framework for quantification and assessment of uncertainties associated with monitoring of biological quality elements (Lindegarth et al. 2013). This framework is a fundament for future work on estimation and reduction of uncertainty in current and future monitoring within WATERS and it could also provide answers to some of the issues raised in the RBDA tool for confidence assessment. Therefore the aim of the work within this group was to explore the relevance of WATERS uncertainty framework for RBDA confidence assessment.

Discussions touched upon a number of issues related to uncertainties (e.g. estimation of precision and confidence in classification, acceptable levels of confidence or precision and relationships between sampling designs and confidence). Nevertheless, the main aim was to use the uncertainty framework to develop estimates of confidence in classifications based on monitoring data. As an example, data on the benthic diatom index (IPS) from sixteen streams in Västra Götaland were used (Figure 13).

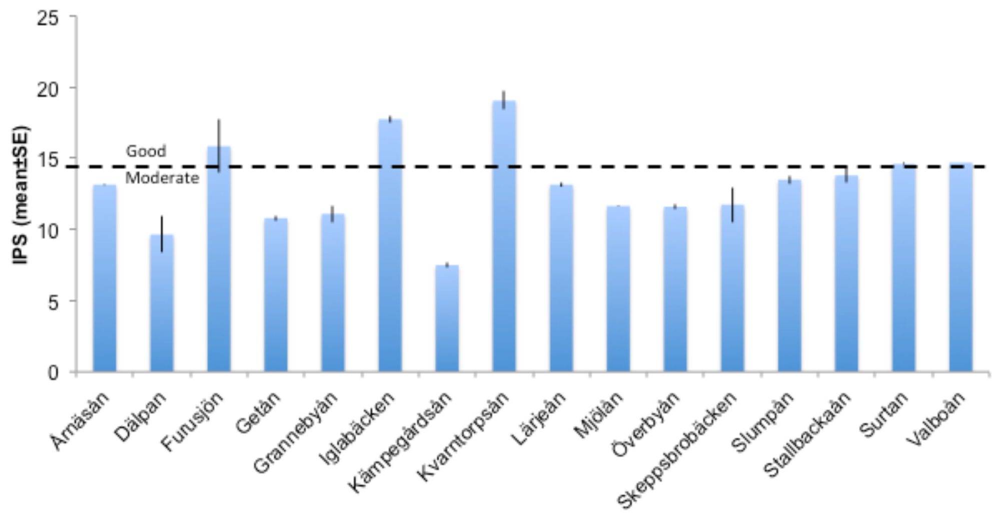


Figure 13: Mean of IPS in sixteen streams in Västra Götaland. Samples are collected between 2008 and 2011 with one pooled sample per stream at 2-3 years per stream.

Example data

Data on IPS were collected according to the Swedish monitoring standard (Naturvårdsverket 2009), which recommends one pooled sample per stream based on five stones collected within 10 m. Thus for each stream (=water body), there was one value of IPS from each of two years (only Surtan was sampled at three years). Samples were collected during 2008, 2010 and 2011 and although the design was not fully crossed, complete model can be described as (see Lindegarth et al. 2013 for formulation of models):

$$y = \mu + YEAR + SITE + YEAR * SITE,$$

where the years and sites are random factors.

Uncertainty of mean estimates

One important issue here is the lack of spatial replication, most notably the lack of replicated sites to represent the whole water body, but also the lack of small scale replication within sites due to pooling of stones. A more complete design would have incorporated replicate sites per waterbody and possibly also replicates within sites to allow assessment of small-scale patchiness. Such a design would have corresponded to the following model:

$$y = \mu + YEAR + WATERBODY + YEAR * WATERBODY + SITE(WATERBODY) + YEAR * SITE(WATERBODY) + PATCHINESS$$

and a variance around the total mean of:

$$V[\bar{y}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_{WB}^2}{b} + \frac{s_{Y*WB}^2}{ab} + \frac{s_{S(WB)}^2}{cb} + \frac{s_{Y*S(WB)}^2}{abc} + \frac{s_e^2}{abcn}.$$

Here the $s_Y^2, s_{WB}^2, s_{Y*WB}^2, s_{S(WB)}^2, s_{Y*S(WB)}^2, s_e^2$ represent variance components for different sources of variation and *a*, *b*, *c* and *n* are the number of sampled years, water bodies, sites and replicate samples respectively. *Y* is the number of years in the WFD cycle, i.e. 6. Finally, the corresponding variance for a mean calculated within a particular waterbody (which is the spatial unit to be classified) can be expressed as:

$$V[\bar{y}_{WB}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_{S(WB)}^2}{c} + \frac{s_{Y*S(WB)}^2}{ac} + \frac{s_e^2}{acn}.$$

If there were reliable estimates of variance components, and if we inserted the actual number of samples in this particular design, *a*=2, *Y*=6, *c*=1 and *n*=1, we could estimate the uncertainty of means for individual water bodies as:

$$V[\bar{y}_{WB}] = \frac{s_Y^2 * (1 - \frac{2}{6})}{2} + \frac{s_{S(WB)}^2}{1} + \frac{s_{Y*S(WB)}^2}{2 * 1} + \frac{s_e^2}{2 * 1 * 1}$$

Now, because only one site was sampled per water body in this study, there is no way that we can estimate the variability among sites within water bodies, $s_{S(WB)}^2$ from these data. This component may or may not be important, but the fact that only one site is sampled per water body may indicate that it is of little importance (sampling of only one site per water body appears to be a deliberate monitoring strategy). Nevertheless, it is clear that any existing variability among sites may potentially have large consequences for the uncertainty of estimates within water bodies and that quantitative estimates are needed. While there are studies addressing variability among sites (M. Kahlert, pers. comm.) these were not available here and for the purposes of these analyses, variability among sites was assumed to be negligible, i.e. $s^2_{S(WB)} = 0$. Furthermore, because replicate samples were not available from each site at individual years, variance components due to patchiness and spatio-temporal interactions cannot be separated, but will be estimated as a sum, i.e. $s^2_{Y*S(WB)} + s^2_e$.

Using the data on IPS, variance components were estimated using the restricted maximum likelihood (REML) method with the program R (specifically the library "lme4" and the function "lmer()"; R Development Core Team 2008; Table 7). Because there is an interest of estimating means for particular water bodies, sites were considered fixed factors while Years were considered random. Note also that the component "Residual" estimates the combined variance due to $s^2_{Y*S(WB)}$ and i.e. s^2_e.

Table 7: Output from R showing the linear model and variance estimates.

```
> ########### REML
> vars<-lmer(IPS ~ Site + (1|Year), data=dataset)
Linear mixed model fit by REML
Formula: IPS ~ Site + (1 | Year)
   Data: dataset
AIC      BIC      logLik    deviance   REMLdev
90.88    117.8    -27.44    61.46      54.88

Random effects:
Groups     Name          Variance   Std.Dev.
Year       (Intercept)   0.35077    0.59226
Residual                 0.61178    0.78216
Number of obs: 33, groups: Year, 3
```

The REML analyses show that the variability among years was roughly half the size (0.35) of that of the residual (0.61). The variability among years most likely reflects year-to-year fluctuations in water chemistry that might occur during a WFD cycle (M. Kahlert, pers. comm.). As explained earlier residual variability reflect more complex small-scale and interactive processes. Using these estimates (assuming i.e. $s^2_{S(WB)} = 0$) the total variance of estimated means in a water body can be estimated as:

$$V[\bar{y}_{WB}] = \frac{0.351*(1-\frac{2}{6})}{2} + 0 + \frac{0.612}{2} = 0.423.$$

While this number contains several sources of variability and is strongly affected by the number of years sampled, it is interesting to note that it is not far off from the "margin of error" referred to in the Swedish assessment criteria (i.e. 1 for IPS<13 and 0.5 for IPS>13, SEPA 2010).

Uncertainty of classifications

With this estimate of variability it is finally possible to address the main question posed by the RBDA: how certain (=confident) can we be that a certain classification is correct? If a water body is classified as "better than moderate" (i.e. above the Good-Moderate [G-M] boundary): is this reliable or is it ambiguous? If a water body is classified as "below the G-M boundary": is this reliable or is it ambiguous? How large must the deviation be from the G-M boundary to be considered reliable at a certain level of confidence (i.e. 95%)?

Such questions can be addressed using the procedures described in Annex A of Lindegarth et al. (2013). In short we can use the normal distribution and the variance estimate to estimate the confidence in classification above or below the G-M boundary (or for individual status classes). These analyses show that if the estimated mean is above 15.57, we can be 95% confident that the true mean of the IPS is at least above the G-M boundary, i.e. 14.5 (Fig X). On the other hand, if the estimated mean is 13.43 or smaller, the true mean is with 95% confidence below the G-M boundary (Figure 14). Note, however, that these values do not account for any existing variability among sites and that it is entirely dependent on the monitoring design. These numbers can only serve as illustrations of a procedure in this particular example and cannot be taken as universal guidelines.

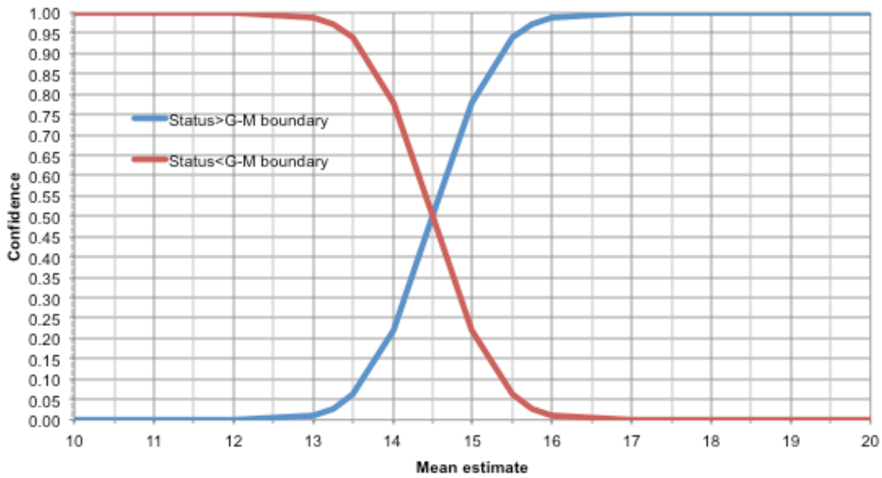


Figure 14: Expected confidence in classifications above or below the Good-Moderate boundary 14.5 using the normal distribution and $V[\bar{y}_{WB}]$=0.423.

For the measured streams in Västra Götaland this means that three streams are classified as better than "Moderate" with high confidence (>95%), while nine streams are classified as below "Good" (Table 8). Two streams are classified as "Moderate" but the confidence for these is smaller than 95% (i.e. 86 and 94%). Similarly, two streams are classified as “Good” but the confidence for a class better than moderate was only 56-62%.

Table 8: Mean estimates, status classes and confidence for IPS in sixteen streams in Västra Götaland. Confidence estimates >95% are shaded green (above G-M boundary) and red (below G-M).

Water body	Mean	Class	p (Status>Moderate)	p (Status<Good)
Årnäsån	13.15	Moderate	0.02	0.98
Dälpan	9.65	Poor	0.00	1.00
Furusjön	15.85	Good	0.98	0.02
Getån	10.8	Poor	0.00	1.00
Grannebyån	11.1	Moderate	0.00	1.00
Iglabäcken	17.75	High	1.00	0.00
Kämpegårdsån	7.5	Bad	0.00	1.00
Kvarntorpsån	19.05	High	1.00	0.00
Lärjeån	13.15	Moderate	0.02	0.98
Mjölån	11.65	Moderate	0.00	1.00
Överbyån	11.6	Moderate	0.00	1.00
Skeppsbrobäcken	11.75	Moderate	0.00	1.00
Slumpån	13.5	Moderate	0.06	0.94
Stallbackaån	13.8	Moderate	0.14	0.86
Surtan	14.6	Good	0.56	0.44
Valboån	14.7	Good	0.62	0.38

Conclusions on uncertainty in diatom stream monitoring

These analyses, using benthic diatoms as an example, show how the “uncertainty framework” developed by Lindegarth et al. (2013) can be used to provide answers to specific needs within the water management. The example show how components of spatial and temporal variability can be combined, monitoring designs can be accounted for and how these can be related to class boundaries to provide accessible assessments of confidence and reliability to status classifications.

It needs to be stressed here, that one potentially serious problem with this particular example is the lack of information about spatial variability among sites within water bodies. The assumption that $s^2_{S(WB)} = 0$ was necessary in order to allow any sort of calculations but currently we have no evidence to support that such an assumption is reasonable. Note however, that quantitative estimates of $s^2_{S(WB)}$, possibly in a “library” of uncertainties as suggested by Lindegarth et al. (2013) could substantially improve the situation even if monitoring programme in a particular water body of interest does not allow such estimation. For the moment, however, because this is currently missing the confidence assessments for these particular streams need to be used with great caution.

Finally it is clear that the discussions and exercises in this group emphasised the need for specific and user-friendly tools that can be used by the CAB and RBDA to deal with issues of uncertainty and confidence. The uncertainty framework can provide the theoretical foundation for tools and as shown in a few examples above, adaptation to specific user-defined problems can readily be made during collaborative sessions like these.

References

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

Lindegarth, M., Carstensen J., Johnson, R.K. (2013). Uncertainty of biological indicators for the WFD in Swedish water bodies: current procedures and a proposed framework for the future. Deliverable 2.2-1, WATERS Report no. 2013:1. Havsmiljöinstitutet, Sweden.

Naturvårdsverket (2009). Påväxt i rinnande vatten – kiselalgsanalys. Version 3:1: 2009-03-13. Available online at: https://www.havochvatten.se/download/18.64f5b3211343cffddb280004863/1325250749399/Påväxt+i+rinnande+vatten-+kiselalgsanalys.pdf

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

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

River Basin District Authorities (2013). Kokbok för kartläggning och analys 2013-2014. Bilaga 1aA. In Swedish.

SEPA (2010). Status, potential and quality requirements for lakes, water courses, coastal and transitional waters: A handbook on how quality requirements in bodies of surface water can be determined and monitored. Swedish Environmental Protection Agency. 2007:4, 421 pp.

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.

Agenda for workshop

Wednesday 30th January 2013		
morning	Travel to Tjärnö	-
12:00	Lunch	
13:00	Welcome and practical details	Mats
13:15	Lecture about indicator development, basic concepts, different approaches.	Ulf
13:45	Lecture on multiple regression	Anders G.
14:00	General linear models (GLM)	Thorsten B.
14:15	Discussion of statistical lectures	
15:00	Presentation of data and problem – solicit cases among participants	
	What type of data do you have? Which are your questions or hypotheses?	
15:30	Coffee break	
16:00	Break out groups to work on data	
	Data sets are expected to be properly organized prior to the workshop	
19:00	Dinner	
20:00	Social get-together and continue work	

Thursday 31st January 2013		
09:00	Feed-back on status from groups (presentations, if any)	
09:10	Uncertainty framework (concepts, eelgrass example)	Jacob
09:30	Uncertainty framework (BQI example)	Mats
09:50	Break out groups to work on data	
10:30	Coffee break	
11:00	Break out groups to work on data	
12:00	Lunch	
13:00	Break out groups to work on data	
15:30	Coffee break	

16:00	Break out groups to work on data
	Expected outcome: Ideas and collaboration for further indicator development
19:00	Dinner
20:00	Social get-together and continue work

Friday 1st February 2013	
09:00	Feed-back on status from groups
09:30	Break out groups to work on data
10:30	Coffee break
11:00	Break out groups to work on data
	Expected outcome: draft indicator and outline of report/paper
12:00	Lunch
13:00	Departure

List of participants

Deltagare	e-post	Organisation
Anders Grimvall	anders.grimvall@liu.se	Havsmiljöinstitutet
Kerstin Holmgren	kerstin.holmgren@slu.se	Akvatiska resurser, SLU
Leif Pihl	leif.pihl@marecol.gu.se	Göteborgs Universitet
Martin Karlsson	martin.karlsson@slu.se	Akvatiska resurser, SLU
Bengt Karlson	Bengt.Karlson@smhi.se	SMHI
Marina Magnusson	marina@marine-monitoring.se	Marine Monitoring AB
Sofia Wikström	sofia.wikstrom@aquabiota.se	AquaBiota Water Research
Thorsten Balsby	thba@dmu.dk	Bioscience, Aarhus Universitet
Maria Kahlert	Maria.Kahlert@slu.se	Vatten & Miljö, SLU
Mats Lindegarth	mats.lindegarth@marecol.gu.se	Göteborgs Universitet
Jacob Carstensen	jac@dmu.dk	Bioscience, Aarhus Universitet
Ulf Grandin	ulf.grandin@slu.se	Vatten & Miljö, SLU
Dorte Krause-Jensen	dkj@dmu.dk	Bioscience, Aarhus Universitet
Ragnar Lagergren	Ragnar.lagergren@lanstyrelsen.se	Länstyrelsen, Västra Götaland

Statistical workshop on gradient studies

The second statistical workshop in WATERS was held at Tjärnö from 30th January to 1st of February 2013 with the aim of indicator development and uncertainty assessment of indicators. A total of 14 persons attended the workshop that included four statistical lectures and group discussions. Statistical analyses of both long-term monitoring data and data from gradient studies were initiated and will be continued in the future. The outcomes of these initial analyses are reported here. The workshop has laid a sound foundation for coming collaboration between biologists and statisticians within WATERS.

WATERS is coordinated by:

Havsmiljöinstitutet
Swedish Institute for the Marine Environment

WATERS is financed by:

NATUR VÅRDS VERKET
SWEDISH ENVIRONMENTAL PROTECTION AGENCY