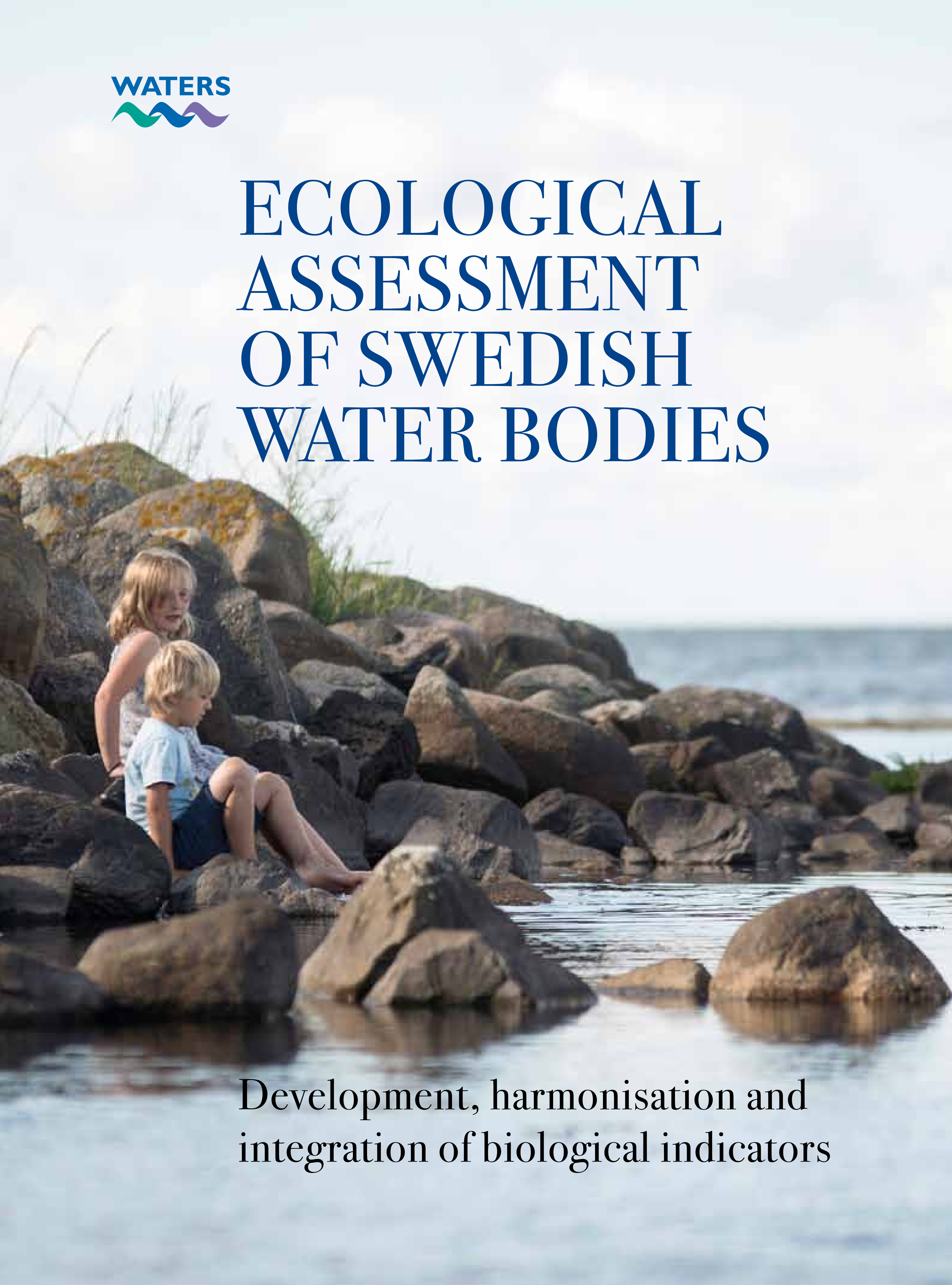
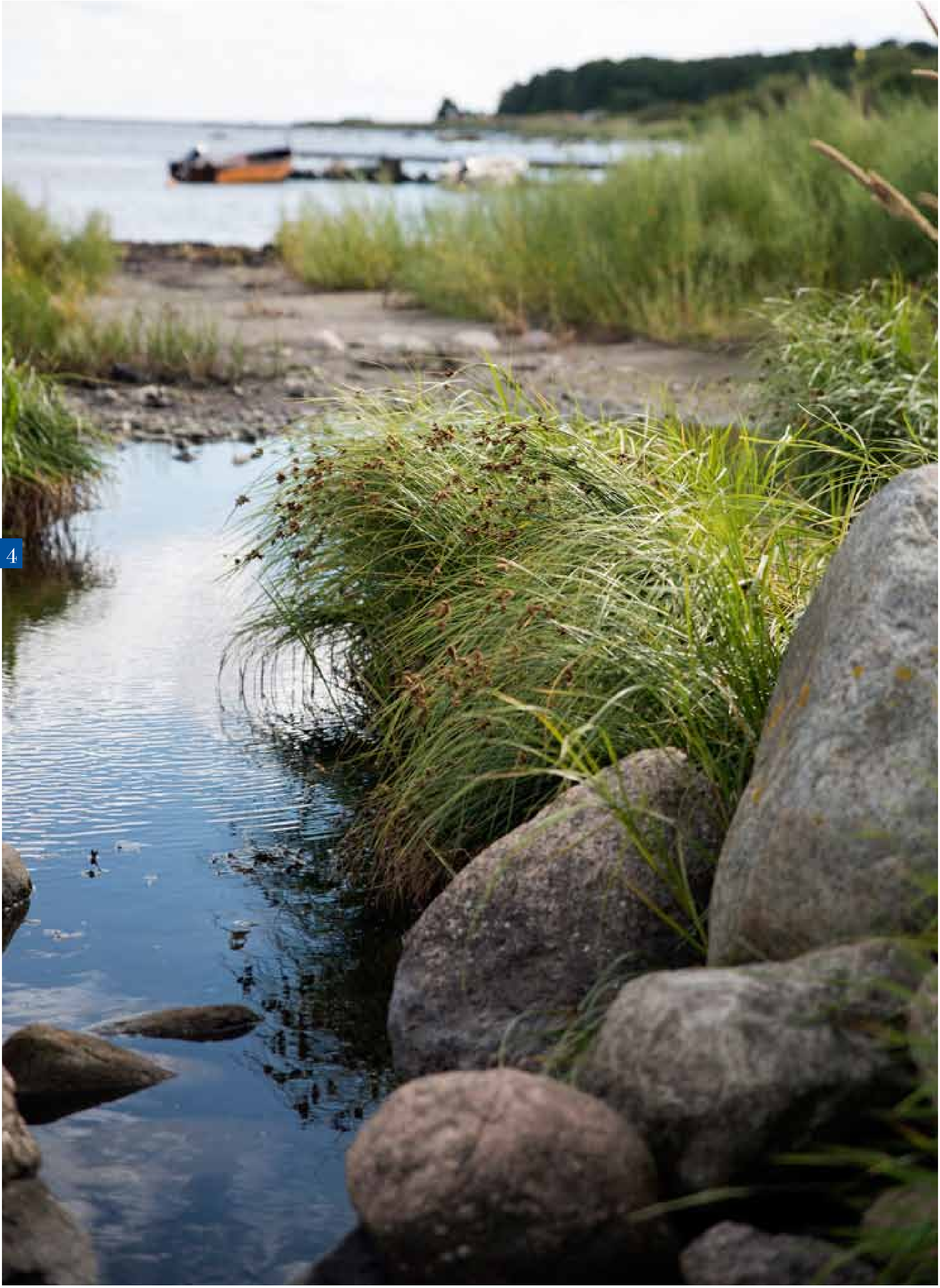
WATERS

ECOLOGICAL ASSESSMENT OF SWEDISH WATER BODIES



Development, harmonisation and integration of biological indicators



Foreword

The Water Framework Directive was incorporated into Swedish legislation in 2004. Its aim is to develop sustainable management of European surface- and groundwater and with the implementation of the directive there was a need for development of monitoring programmes and tools for assessing water quality. In response to these needs the Swedish Environmental Protection Agency opened a call for research on “Biological assessment criteria in aquatic environments”. As a result of this call, a consortium of 11 partners, was granted a total budget of 47 million SEK funded by the Swedish Environmental Protection Agency’s Environmental Research Grant, to pursue the research programme WATERS during 5 years.

WATERS has developed better indicators and methods for classification and references in Swedish lakes, streams and coastal waters. This provides a valuable basis for a new generation of biological assessment criteria to be defined by the authorities. WATERS has also developed general methods for assessing and reducing uncertainty in classification of ecological status, as well as harmonised and transparent methods for integrated assessment. In combination, these results are expected to improve and simplify future status assessments according to the Water Framework Directive.

Several teams of ecologists specialised in limnic and marine waters have contributed, led by Jacob Carstensen, Aarhus university, Richard Johnson, Swedish university of agricultural sciences, Leif Pihl, University of Gothenburg and Sofia A Wikström, Stockholm university. The programme has been coordinated by Mats Lindegarth at the Swedish institute for the marine environment. We wish to express our sincere gratitude to everyone that has contributed to the success of the programme.

WATERS steering group has consisted of the coordinator, team leaders and responsible contacts at the Swedish Agency for Marine and Water Management and the Swedish Environmental Protection Agency. These have been represented by Mats Svensson and Cecilia Lindblad respectively.

The authors alone are responsible for the content of this report.

The Swedish Environmental Protection Agency
and the Swedish Agency for Marine and Water Management, September 2016

Förord

I Sverige infördes EUs ramdirektiv för vatten i svensk lagstiftning år 2004. Syftet var att skapa långsiktigt hållbar förvaltning av våra vattenresurser. I och med direktivets genomförande, behövde övervakningsprogram och befintliga verktyg för bedömning av vattenkvalitet ses över och utvecklas. För att möta de ökade kraven utlyste Naturvårdsverket 2009 ett forskningsprogram med titeln: ”Akvatiska biologiska bedömningsgrunder”.

Forskningsprogrammet WATERS, med elva deltagande partners, beviljades genom denna utlysning medel ur Naturvårdsverkets miljöforskningsanslag. WATERS totala budget var 47 miljoner kronor.

Under sin femåriga programtid har WATERS utvecklat bättre indikatorer och justerat klassificeringar och referensvärden för våra vattenförekomster. Detta är ett värdefullt underlag till de kommande biologiska bedömningsgrunderna. WATERS har också utvecklat metoder för hantering av osäkerhet i bedömningarna, liksom metoder för transparens och harmonisering av sammanvägd ekologisk status. Sammantaget kommer resultaten att förenkla vattendirektivsarbetet.

Flera grupper limniska och marina ekologer har ingått, ledda av Jacob Carstensen, Aarhus universitet, Richard Johnson, Sveriges Lantbruksuniversitet, Leif Pihl, Göteborgs universitet, samt Sofia A Wikström, Stockholms universitet. Koordinator för programmet har varit Mats Lindegarth vid Havsmiljöinstitutet och Göteborgs Universitet. Vi riktar ett stort tack till alla som bidragit till att göra programmet framgångsrikt.

I WATERS styrgrupp har programansvarige samt forskare från respektive fokusområden ingått samt representanter från Havs- och vattenmyndigheten och Naturvårdsverket. Mats Svensson har representerat Havs- och vattenmyndigheten och vid Naturvårdsverket har Cecilia Lindblad ansvarat för programmet.

Författarna ansvarar ensamma för rapportens innehåll.

Naturvårdsverket och Havs- och vattenmyndigheten i september 2016

Deliverable 1.1-4, WATERS report no 2016:10

Title: ECOLOGICAL ASSESSMENT OF SWEDISH WATER BODIES
Development, harmonisation and integration of biological indicators

Publisher: Havsmiljöinstitutet/Swedish Institute for the Marine Environment,
P.O. Box 260, SE-405 30 Göteborg, Sweden

Published: October 2016

ISBN 978-91-639-2098-1 (Havsmiljöinstitutet/Swedish Insitutet for the Marine Environment)

ISBN 978-91-87967-35-1 (Havs- och vattenmyndigheten/Swedish Agency for Marine and Water Management)

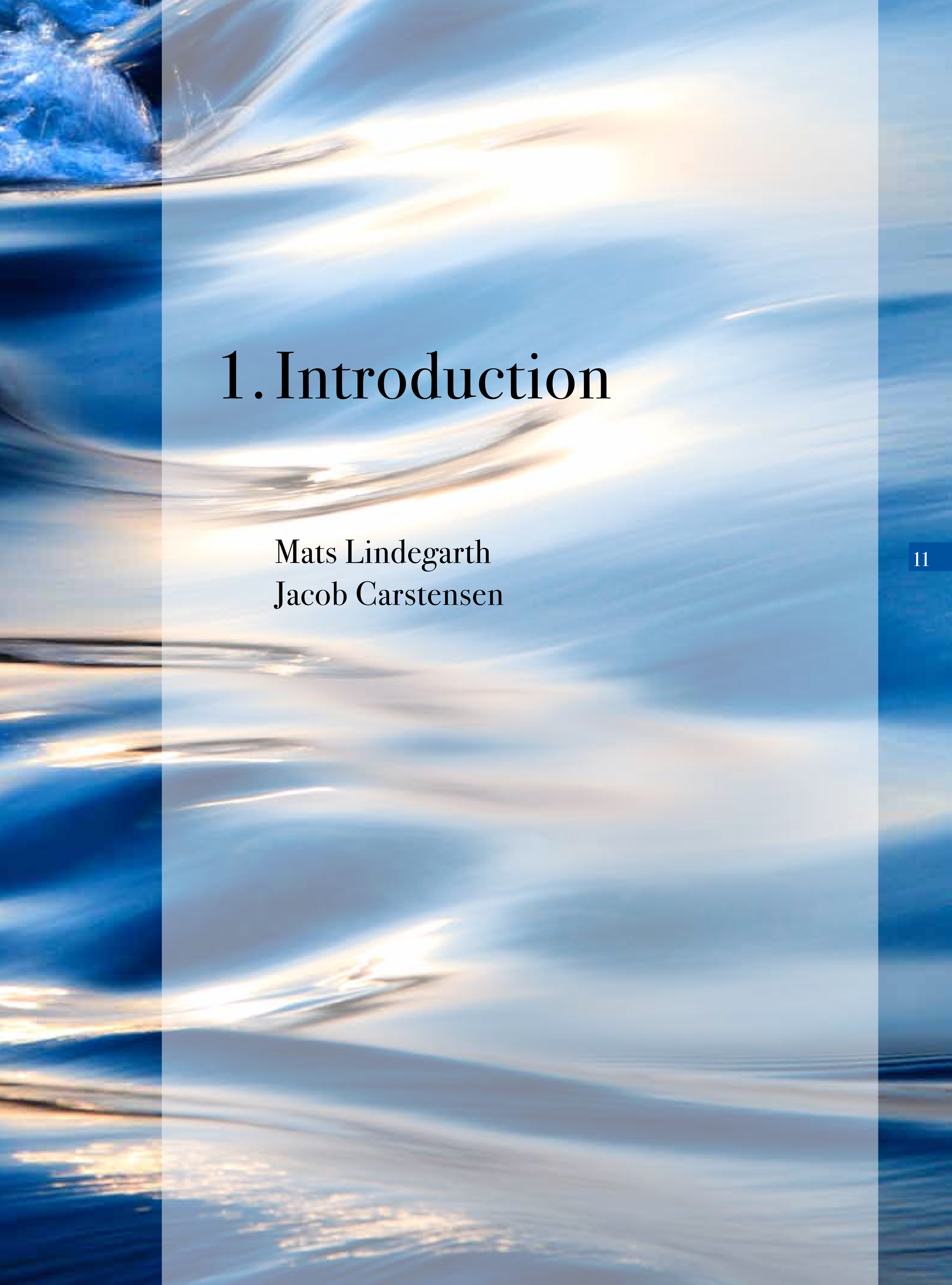
Please cite as:

Lindegarth M, Carstensen J, Drakare S, Johnson RK, Nyström Sandman A, Söderpalm A, Wikström S A (Editors). Ecological Assessment of Swedish Water Bodies; development, harmonisation and integration of biological indicators. Final report of the research programme WATERS. Deliverable 1.1-4, WATERS report no 2016:10. Havsmiljöinstitutet, Sweden.

Photo:

Anders Asp, Thomas Axenrot, Mats Blomqvist, David Börjesson, Magnus Dahlberg, Fraucke Ecke, Dan Evander, Joakim Hansen, Helena Höglander, Martin Karlsson, Eva Kyhlberg, Jason Neuswanger, Lars Ohlson, Birgitta Renöfält, Alfred Sandström, Annika Söderpalm, Amélie Truchy, Susanne Qvarfordt.

Table of contents



1. Introduction

Mats Lindegarth
Jacob Carstensen

Pressures on aquatic ecosystems have increased due to human activities, and as a consequence many systems have now reached a degraded state in which ecosystem services are compromised. Today, the most prominent pressures are eutrophication, hydromorphological modification, pollution, acidification, and overfishing. Increasing evidence of these pressures leading to deterioration of aquatic ecosystems, obtained from numerous scientific studies during the twentieth century, has prompted political actions at the local, regional, and national levels. Managerial frameworks for reducing the ecological effects of human pressures began to be introduced in the 1970s, primarily aiming at setting emission targets (e.g. Total Maximum Daily Loads in the USA) or relative emission reduction targets (e.g. 50% reduction of nitrogen inputs relative to a baseline period). However, the European Water Framework Directive (WFD) introduced a paradigm shift by changing the focus of targets from emissions to recipients, by declaring that all European surface waters should have good ecological status by no later than 2027. This change of objective is meaningful and also poses greater implementation challenges, particularly given the growing evidence of climate change.

The purpose of the WFD is, among other things, to prevent further deterioration, protect and enhance the status of aquatic ecosystems, and to promote sustainable water use. This means that the definitions of environmental objectives and targets are based, not only on scientific ecological arguments, but also involves social and economic aspects. In contrast, however, assessments of ecological status are only based on the best available scientific knowledge and principles. The ecological status is also an important basis for setting environmental objectives and for evaluating the need for measures. When the WFD was enacted in 2000, most of this scientific under-

standing was conceptual, but implementation demands quantitative knowledge in order to obtain operational tools for determining appropriate management measures, acknowledging that implementing these measures may be costly for society. Consequently, over the years, considerable scientific effort has been devoted to improving our quantitative understanding of ecosystem responses to pressures, developing tools for assessing ecological status, and determining the most cost-effective measures to restore good ecological status for those water bodies in a degraded state. Essentially, scientists have been challenged to develop tools that environmental managers in national agencies can apply to fulfil the objectives of the WFD.

The WATERS programme has helped improve the scientific foundation of WFD implementation by developing tools for assessing ecological status in both inland and marine surface waters in Sweden. The current report is a synthesis of the research conducted during the course of the programme (2011–2016). In Chapter 2 we briefly describe the basic concepts and requirements stipulated by the WFD as well as the overall challenges that we faced at the beginning of the WATERS programme. In Chapters 3 and 4, we review current ecological indicators used for Swedish lakes and coastal areas, and describe our development of new indicators to substitute and supplement existing ones. In Chapter 5 we describe the harmonisation and aggregation of indicator information to derive an overall ecological status. Finally, we summarise our general recommendations for status assessment and monitoring for the WFD in Sweden and provide an outlook for the upcoming implementation of these findings in practical water management.

2. The WFD implementation and its challenges

Mats Lindegarth
Jacob Carstensen
Richard K Johnson
Sofia A Wikström

The adoption and implementation of the Water Framework Directive (WFD) has had major impacts on European policies and management of coastal and inland surface waters. At the centre of WFD implementation is a cyclic process for adaptive management, which involves (1) characterisation of pressures and impacts of catchments, (2) definition of environmental objectives, (3) design and implementation of programmes of measures, (4) monitoring and assessment of ecological status, and (5) integration of information in a river basin management plan (Figure 2.1).

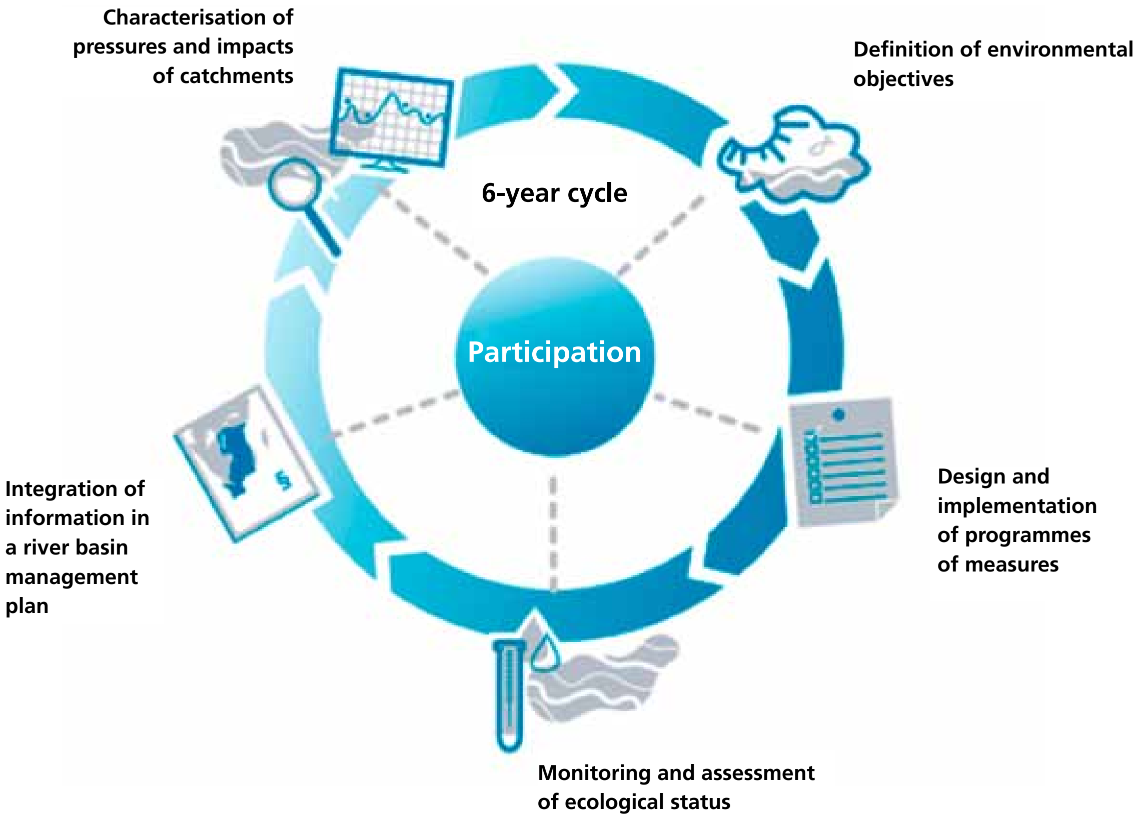


Figure 2.1: The successive steps in the water management cycle defined by the WFD (modified from the Swedish Water District Authorities).

The fundamental unit for status assessment and pressure analysis according to the WFD are spatial units defined as water bodies. The WFD cycles were initiated after a 10-year preparation period envisaging a total of three cycles (Table 2.1), after which good ecological status (or good ecological potential for heavily modified water bodies) should have been achieved for all surface waters in Europe. Hence, good ecological status should be achieved no later than 2027. Furthermore, the WFD dictates that ecological status must not deteriorate between cycles.

Table 2.1. The WFD sets clear deadlines for each of the requirements, which add up to an ambitious overall timetable. Key milestones with reference to specific articles in the Directive are listed below (modified from www.ec.europa.eu/environment/water/water-framework/).

Year	Milestone	Article
2000	Directive entered into force	25
2003	Transposition in national legislation	23
	Identification of River Basin Districts and Authorities	3
2004	Characterisation of river basin: pressures, impacts and economic analysis	5
2006	Establishment of monitoring network	8
	Start public consultation (at the latest)	14
2008	Present draft river basin management plan	13
2009	Finalise river basin management plan including programme of measures	13 & 11
2010	Introduce pricing policies	9
2012	Operationalise programmes of measures	11
2015	Meet environmental objectives	4
	First management cycle ends	
	Second river basin management plan and first flood risk management plan	
2021	Second management cycle ends	4 & 13
2027	Third management cycle ends, final deadline for meeting objectives	4 & 13

Harmonised assessment system with common concepts and requirements

One of the fundamental components of the WFD cycle is that ecological status and changes therein should be assessed using relevant biological and chemical quality information. Such assessments are necessary to evaluate whether environmental objectives are met, to determine what measures need to be taken, and to evaluate whether programmes of measures are effective.

To develop an assessment system that is ecosystem based, transparent, and harmonised among the EU Member States, the Directive defines a set of common concepts and requirements. Some of the more fundamental ones are as follows:

- European surface waters (i.e. lakes, rivers and coastal waters) are classified into a set of water body types based on large-scale ecoregions and a set of physical and chemical factors;
- assessment of ecological status shall refer to temporal units consisting of six-year periods, and to spatial units consisting of individual water bodies, which are subdivisions based on the overall typology;
- assessment of ecological status shall be based on biological quality elements (BQEs), i.e. phytoplankton, aquatic flora, benthic invertebrate fauna, and fish (except in coastal waters), which are supported by a number of hydromorphological and physicochemical elements;
- the status of each BQE is monitored and assessed using one or several indicators, for which the response to relevant pressures is known and for which type-specific reference conditions can be determined;
- the status of each indicator is assessed using an ecological quality ratio (EQR) that relates the observed state to the type-specific reference condition and may take a value of 0–1;
- the EQR range contains intervals characterised as "high", "good", "moderate", "poor", and "bad" status (Figure 2.2); and
- the integrated assessment of ecological status should be based on the precautionary "one-out, all-out" principle in which the BQE having the lowest EQR determines the overall ecological status.

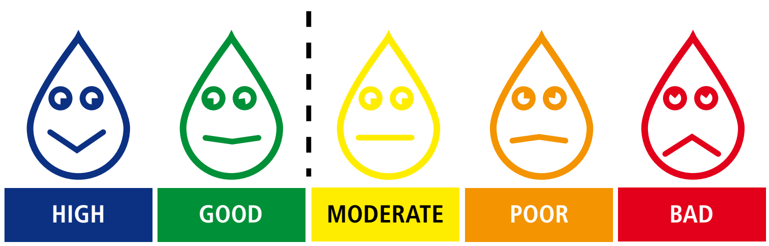


Figure 2.2: Illustration of the different classification intervals within the EQR scale defined by the WFD.

These and a multitude of additional definitions developed in a Common Implementation Strategy (CIS), documented as a set of CIS guidance reports, provide a framework for harmonised assessments among Member States and for different types of surface waters. Nevertheless, implementing and operationalising the WFD in a coherent way in different parts of Europe has been a challenging task for scientists and managers involved in developing and applying tools for ecological assessment. The implementation has largely been driven as a bottom-up process, because biogeographic variations and differences in prevailing pressures may require highly specialised indicators, in addition to constraints imposed by differences in available monitoring data among Member States or even among regions within Member States. Challenges remain before the implementation of the WFD can be considered consistent and complete to a degree that is in accordance with the intentions of the Directive (Hering et al. 2010). Many of these methodological difficulties and shortcomings in practical implementation are also relevant to status assessments of Swedish surface waters (e.g. Rolff 2009).

WATERS has adressed problems in the Swedish WFD implementation

In Sweden, the WFD was formally adopted in Swedish law by Chapter 5 in the Environmental Code, by the Ordinance on Water Quality Management (SFS 2004:660), and by instructions from the Swedish Environmental Protection Agency (SEPA 2008). Advice on the application of assessment criteria were provided in a handbook (SEPA 2007) and additional coordinating documents have been developed by the Water District Authorities, which are responsible for coordinating assessments. Furthermore, the instructions have since been updated by the Swedish Agency for Marine and Water Management (SwAM 2013).

Apart from these formal aspects of implementation, the practical application of the criteria has had large consequences for how water quality is assessed in Swedish waters. Assessments of ecological status resulting from the first management cycle also revealed a number of shortcomings and ambiguities of the assessment criteria and of the monitoring of the biological quality elements, as well as large inconsistencies among them.

The WATERS programme has addressed these inconsistencies with the aim to develop:

- more reliable and sensitive indicators;
- common strategies for defining reference conditions and class boundaries;
- common strategies for assessing uncertainty of estimates and classification using monitoring data; and
- a coherent framework for whole-system assessment

These aims pose several scientific challenges that require empirical and theoretical expertise on individual BQEs in inland and coastal waters, as well as on approaches to integrated assessment across BQEs. The work within WATERS has shown that the specific challenges and methods required differ among BQEs (see chapters 3 and 4). Nevertheless, important steps involve quantifying, testing, and comparing the pressure-response relationships and uncertainties of a number of candidate indicators, and developing and testing general frameworks for the harmonisation of assessment methods.

2.1 Deriving operational indicators for BQEs

Numerous indicators of ecological status have emerged with the advent of the WFD. Birk et al. (2012) reported almost 300 methods for assessing ecological status with an almost even distribution between rivers, lakes, transitional waters, and coastal waters, whereas more assessment methods addressed benthic BQEs relative to phytoplankton and fish. The majority of the assessment methods focused on eutrophication and organic pollution as the main pressure, but hydromorphological alteration was also commonly considered a pressure in rivers and transitional waters. Birk et al. (2012) further found that about one third of the assessment methods were not validated against the pressure, i.e. a putative relationship to the main pressure was not identified. Hence, such assessment methods may not even respond to relevant pressures and may contribute noise rather than information. Moreover, assessment methods with a documented response to a pressure were generally not validated, i.e. testing the relationship on new independent data to confirm the generality of the pressure-response relationship.

Range in both pressure and response needed for statistical significance

Establishing pressure-response relationships requires establishing a suitable range in the pressure relative to other variations in the data. This is mostly achieved by pooling data from multiple water bodies such that there is a broad range in both pressures and responses. However, for an indicator to be operational it also needs to be responsive within a realistic range of the pressure, otherwise the indicator is not a marker of ecological status. For example, Carstensen and Heiskanen (2007) investigated 76 different phytoplankton indicator species and found that only half of them responded significantly to broad changes in nutrient levels and only one species responded sufficiently sensitively to changes in total nitrogen of 20%. In practice, the indicator response to the pressure should be strong enough to render the indicator operational.

Indicators must be quantitative

Operational indicators should also have target values with which to assess the degree of compliance. For the WFD this implies having defined reference conditions and class boundaries (see below). It is also important that these target values be aligned with those of other indicators, such that the indicators provide an unambiguous status assessment. Moreover, it is essential that both indicators and targets be quantitative, implying that it is insufficient to know only whether or not the target has been met; it is just as important to know how far the status is from the target. This information is crucial to the overall integrated assessment.

Standard monitoring should be sufficient to determine whether the target is met

Ability to estimate the precision of an indicator is a fundamental requirement that allows assessment of the confidence with which the indicator addresses the target, and the confidence of the combination of all indicators addressing good ecological status in the integrated assessment. Indicators should be cost-effective, implying that they should be based on standard monitoring data that are not overly costly to obtain and that they should not require excessive amounts of data for estimation with a reasonable degree of precision. New techniques for monitoring have emerged with the potential to deliver more information at lower cost, but the information provided and linkage to existing indicators should be carefully assessed before such techniques are implemented as standard monitoring methods. The cost-efficiency of indicators based on novel techniques producing large amounts of data should consider not only the cost of collecting the data but also the cost of processing and quality assurance.

2.2 Deriving reference conditions and boundaries

The reference condition reflects undisturbed conditions

Effective management of aquatic resources requires knowledge of when a water body's condition differs from its natural condition (i.e. absence of human disturbance) or ecological target, and about what has caused an observed deviation. In recent years, reference conditions have been increasingly used to gauge the effects and magnitude of human intervention. An important consideration when discussing reference conditions is terminology, as discussions often consider reference sites that are undisturbed, minimally disturbed, or least disturbed (see Stoddard et al. 2006 for clarification). According to the WFD, reference conditions are defined as reflecting no or minimal anthropogenic stress, although the use of alternative benchmarks has been discussed (see Pardo et al. 2012). As a pillar of the WFD, Member States are required to identify reference conditions for the purpose of defining reference biological communities. Approaches used to determine reference conditions include: (i) spatial approaches such as surveys, (ii) historical data such as palaeo-reconstruction data, (iii) modelling approaches and hindcasting, and (iv) expert judgement.

Precision can be increased by using site-specific characteristics

In areas where land use has not drastically altered the landscape, the identification of reference conditions is fairly straightforward and spatial methods are frequently used (Johnson et al. 2010). Use of survey data is common as the approach either explicitly (i.e. sites are sampled to include temporal variability) or implicitly (i.e. with space-for-time substitution) includes natural variability. Another reason for the popularity of using survey data is transparency: the definition of what constitutes a reference condition is established *a priori* using reference/pressure filters with inclusion/exclusion criteria being explicitly defined. However, one problem with using spatial, typology-based approaches is that the number of sites needed in order to adequately estimate reference conditions increases with the number of groups used to partition spatial variability. Along these lines, many recent studies have found that spatial approaches do not adequately capture the fine-scale variability of aquatic communities (e.g. Hawkins and Vinson 2000; Davy-Bowker et al. 2006; Aroviita et al. 2009). Hence, although spatial typologies may be useful for partitioning regional variability and providing a useful framework for setting ecological targets, they may often need to be supplemented by site-specific characteristics of the habitat to increase precision. Modelling approaches are often used in areas where humans have extensively altered the landscape over long periods and spatial analogues for estimating reference conditions are few or lacking. If reliable stress-response relationships are known, the reference condition can be predicted by modelling a stress-response relationship for a low, or target, level of stress.

A second modelling approach uses knowledge of relationships between response and predictor variables to predict the expected reference condition (e.g. community assemblage or index). Often an empirical model is calibrated using reference sites; this allows the ecological attributes expected at a site (e.g. probability of taxon occurrence, aka taxonomic completeness) to be predicted from a suite of environmental variables (e.g. Wright et al. 1996; Hallstan et al. 2012). Figure 2.3 illustrates how biological variability within a region may be further partitioned into ecotypes that account for local characteristics. For example, using *a priori* criteria, ecoregion delineations are often used to partition large-scale natural (biogeographic) variability; thereafter more local-scale variables are used to partition the among- and within-site variance of ecotypes. One major distinction between typology- and model-based approaches is that the former relies on categorical variables for partitioning regional variability, whereas the latter often uses continuous variables. Nonetheless, in both approaches, the use of ecological and pressure criteria is considered one of the most cost-effective approaches when screening for potential reference sites/conditions (Wallin et al. 2003).

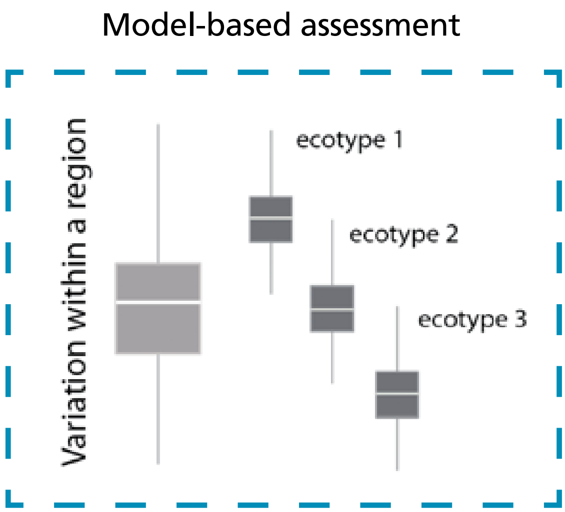


Figure 2.3: Biological variance within a region (e.g. ecoregion) may be further partitioned using variables characterising ecotypes within a region (e.g. habitat features). The dotted line indicates that a site-specific model may be used to predict the biological variance of ecotypes and regions.

Class boundaries define five levels of ecological status

Establishing baseline reference conditions, which can be used to measure the effects of human activities, is only the first step towards determining the ecological status of inland and coastal waters. The WFD also stipulates that the level of human impact on the structure and function of aquatic ecosystems needs to be defined in terms of ecological quality elements (WFD Annex V and Table 2), and that Member States must be able to identify five levels of impact (or ecological status) using these ecological quality elements.

Indices of biological quality are of little value without some knowledge and quantitative estimate of their precision and confidence in assigning individual water bodies to ecological status classes. Quantification of uncertainty is explicit in the WFD: "Estimates of the confidence and precision attained by the monitoring system used shall be stated in the river basin monitoring plan". Variation in pressure-response relationships may be due to poorly characterised cause-effect relationships between species and environment, as well as to (i) sampling variation and sampling method, (ii) sample processing and taxonomic identification errors, and (iii) natural temporal and spatial variation. Standardised field and laboratory protocols are two methods frequently used to reduce variance associated with sample collection and processing. For example, the WFD requires that "Methods used for the monitoring of type parameters shall conform to the international standards listed below … or to such other national or international standards which will ensure the provision of data of an equivalent scientific quality and comparability" (see Section 1.3.6 in Annex 5, European Commission 2000).

Varying approaches to setting class boundaries

In reviewing approaches used by the 28 Member States for setting class boundaries, Birk et al. (2012) found that statistical approaches were most commonly used (45%), followed by ecological relationships (37%) and expert opinion (18%) (Figure 2.4). Methods were classified as statistical if boundaries were established mathematically without using ecological relationships. Methods were classified as expert opinion if boundaries were set by experts without the use of ecological or statistical approaches. For example, in the intercalibration of stream benthic invertebrates, Erba et al. (2009) used a statistical approach when establishing class boundaries. The high/good boundary corresponded to a fixed percentile (i.e. 25th percentile) of the variability of the response metric at reference sites. The good/moderate boundary was set by multiplying the high/good boundary by 0.75 and the moderate/poor boundary by multiplying the high/good boundary by 0.50. Consequently, 25% deviation from the high/good boundary of reference sites is considered a slight deviation and corresponds to good status.

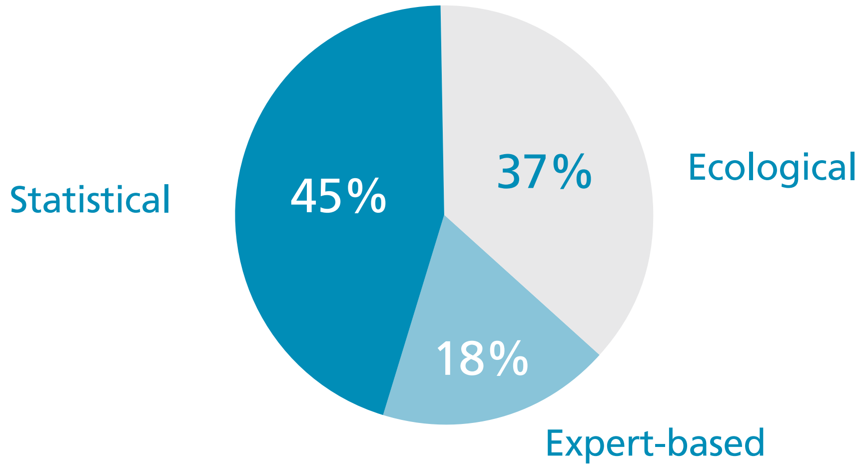


Figure 2.4: Methods for setting ecological boundaries. **Ecological** = Using discontinuities of pressure-response relationships. Use of divergent pressure-response relationships (e.g. paired metrics). **Statistical** = High/good boundary determined using near-natural reference sites and other boundaries set using equidistant division. **Expert-based** = Calibrated against pre-classified sites. Modified from Birk et al. (2012).

Besides addressing reference conditions, two other foci of the WATERS programme have been the harmonisation of approaches used 1) to establish reference conditions and 2) to set class boundaries in the Swedish assessment. To increase the harmonisation of methods for establishing reference conditions and setting class boundaries, Johnson et al. (2014) defined important concepts, identified strengths and weaknesses of current approaches, and suggested future directions. For coastal and marine waters we have focused on the main pressures affecting biodiversity in order to better define pressure criteria for establishing reference conditions or alternative benchmarks. For inland surface waters, the focus was on revising the typology-based approach to decrease the number of water body types. Both marine and freshwater systems used modelling to determine whether precision in ecological assessment increased compared with typology-based approaches.

2.3 Combining indicators

Worst BQE determines overall ecological status

The WFD requires that the overall ecological status of a water body should be assessed based on the status of various BQEs (Table 2.2) using the one-out, all-out principle. Application of this principle essentially means that the overall status is set by the lowest-ranking BQE of those mandatory for a given water body. The BQEs are further assessed by means of a range of sub-elements (called "parameters" in SwAM 2013), differing across water category and BQE. Furthermore, these sub-elements can be assessed by single or multiple indicators. Hence, the WFD entails a hierarchical structure going from overall assessment, to BQEs and sub-elements, and further to indicators (Figure 2.5).

Table 2.2. Indicative parameters to be included in biological assessment methods for the surface water categories and BQEs. The table gives an overview of the parameters mentioned in CIS Guidance No. 7 (optional issues within parentheses). From CIS Guidance No. 14.

Surface water category	Biological Quality Element	Taxonomic composition	Abundance [a]	Disturbance-sensitive taxa	Diversity	Age structure	Frequency and intensity algal blooms	Biomass	Absence of major taxonomic groups	Taxa indicative of pollution
Rivers and lakes	Phytoplankton	X	X				X	X[b]		
	Macrophytes and phytobenthos	X	X							
	Benthic invertebrate fauna	X	X	X	X				X	
	Fish fauna	X	X	X		X				
Transitional waters	Phytoplankton	X	X				X	X		
	Macroalgae	X	X							
	Angiosperms	X	X	X						
	Benthic invertebrate fauna	X	X	X	X					X
	Fish fauna	X	X							(X[d])
Coastal waters	Phytoplankton	X	X		(X)		X	X		
	Macroalgae and angiosperms		X	X	(X[c])					
	Benthic invertebrate fauna	X	X	X	X			(X)		X

a) depth distribution/cover for macroalgae and angiosperms, b) only lakes, c) only macroalgae, d) bioaccumulation bioassays

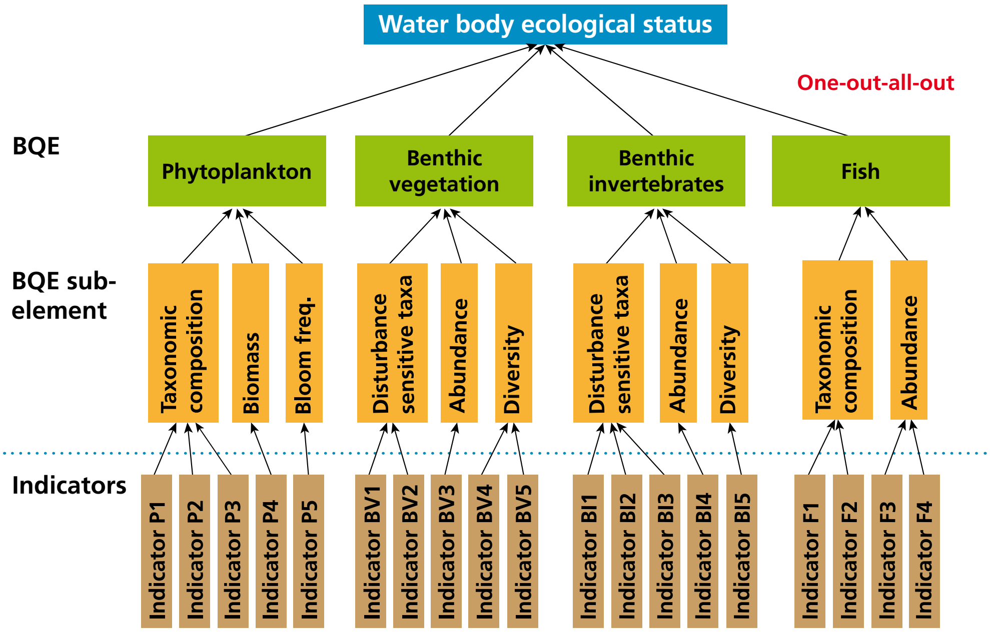


Figure 2.5: The overall principles of information aggregation from indicators to the overall ecological status assessment. The WFD stipulates that the one-out, all-out principle applies to BQE, whereas the principles of aggregation for the other levels of the hierarchy are not given.

Whereas the integration principle for deriving the overall assessment (i.e. one-out, all-out) is given in the WFD, the integration of sub-elements and indicators can be done as deemed appropriate. This lack of harmonisation is due to biogeographic differences regarding the importance of sub-elements, differences in pressures and their effects on sub-elements and indicators, and differences in monitoring yielding quite diverse indicators across Member States and even within countries.

Supporting elements affect status assessments in various ways

In addition to the BQEs used in assessing ecological status, the status of supporting elements should also be included in the overall status assessment. Supporting elements, which include physicochemical and hydromorphological conditions, can be used to modify ecological status determined by the BQEs. According to the principles laid out in CIS Guidance No. 13, high and good status can only be achieved if the supporting elements are also of high and good status, respectively. Furthermore, supporting elements often play a crucial role for various types of expert assessments when data on the BQEs are insufficient or absent. Practical experiences in the Swedish context have shown that this aspect of supporting elements is of great importance, as well as in subsequent steps of the management cycle where they are used to design remedial measures.

A generic approach to combining indicators to overcome lack of harmonisation

The WFD and accompanying guidelines describe how to integrate information from status assessments of BQEs and supporting elements, whereas the information underpinning the BQEs and supporting elements can be integrated in various ways by Member States. Because of this lack of harmonisation for the integration of sub-elements and indicators, diverse solutions have emerged ranging from rule-based assessment systems to composite indicators encapsulating multiple sources of information in a single metric. To overcome this lack of harmonisation, a major objective of the WATERS programme has been to develop a generic approach to combining indicators and sub-elements based on the hierarchical structure of the WFD. CIS Guidance No. 13 suggests the averaging or weighted averaging of indicators responding to the same pressure, whereas the one-out, all-out principle is recommended when indicators can be partitioned into groups responding to different pressures, in case there are multiple dominant pressures.

Assessments transparent at all levels

This framework for integrated assessment should provide transparent information at all levels within the hierarchy, implying that all the details of the integration can be retrieved and assessed. It is also important that the framework be flexible, such that new indicators can easily be added to the assessment. Current assessment systems seldom incorporate such flexibility. Nevertheless, new indicators emerge and it is an important aim of WATERS to develop an integrated assessment system that allows the set of indicators to be expanded in a modular fashion. Finally, the framework should also include an assessment of the confidence of the status classification, which means that the framework should also provide a methodology for aggregating indicator uncertainties to the level of sub-elements and overall status assessment.

Ecological status assessment implicitly assumes that the status of all BQEs is assessed, such that it is based on a broad assessment of relevant organisms. However, most monitoring programmes do not conduct sampling for all BQEs in all water bodies and in many cases a reduced set of BQEs is monitored. In such cases, it is relevant to extrapolate information from other water bodies that are ecologically similar and exposed to the same pressures in order to obtain a coherent ecological status assessment based on all required BQEs. Thus, one objective of the WATERS programme has been to develop a framework for incorporating information on non-monitored BQEs from other water bodies in order to obtain a complete ecological status assessment.

2.4 Estimating confidence in classification

Two types of uncertainty

According to the WFD, status assessments of BQEs should be based on monitoring data. Because such assessments are always associated with some degree of uncertainty, the Directive specifically comments on uncertainty and monitoring programme requirements. The central aspects of uncertainty identified by the WFD are precision and confidence in classification (see also CIS Guidance Nos. 7 and 13). These two concepts are related, but while precision (measured as standard error or confidence interval) refers to the uncertainty of an estimated average of an indicator, confidence in classification refers to the probability that a certain status classification (e.g. "good" and "moderate") is correct (Figure 2.6).

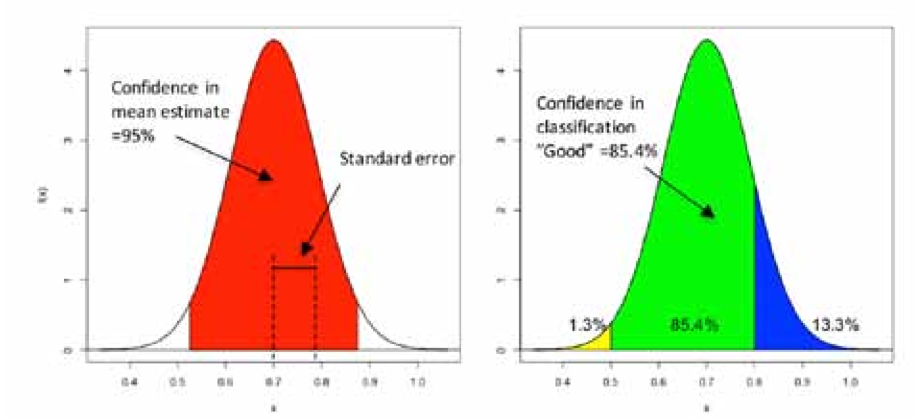


Figure 2.6: Illustration of the terms precision and confidence for an arbitrary indicator denoted x and having a mean value of 0.7. Left: mean ± 95% confidence interval and standard error; right: confidence of classification using a good/moderate boundary of 0.5 and a good/high boundary of 0.8. Percentages represent probabilities for different classes with a 85% probability for the good status.

Briefly stated, the precision of an estimated mean of an indicator in a given assessment period and water body is determined by its variability in time and space. These properties are largely beyond our control, but are determined mainly by the properties and dynamics of the indicator. Precision, however, is also affected by the quality of efforts to monitor the indicator. Measurement errors are minimised by the standardisation of sampling protocols and analytical methods, but the number of samples and the allocation of samples among times and locations that are sampled within a water body also have profound consequences for the precision of the estimated mean.

Because confidence in classification is affected directly by the precision of the mean estimate, it is similarly determined by variability in the data and sampling efforts. In addition, the confidence in classification is also affected by the size of the estimated mean and by how close this is to one or several class boundaries. Given that the precision is held constant, less confidence must be placed in classifications based on an estimated mean that is near a class boundary than to one further away. Therefore, unlike the precision, the confidence in classification is difficult to improve by increasing monitoring efforts because it depends on the actual status of the indicator.

Previous WFD guidance incomplete

Although the Directive states that "frequencies [of sampling] shall be chosen so as to achieve an acceptable level of confidence and precision" (e.g. p. 55), it is important to note that no quantitative definition of "what is acceptable" is defined. Instead, the Directive states that it is necessary that the uncertainty be quantified and that "estimates of the confidence and precision attained by the monitoring system used shall be stated in the river basin management plan". These are important but vague remarks in light of the complexity of the natural spatial variability of aquatic systems and the diverse monitoring designs often used for the various BQEs. Providing robust and reliable methods for dealing with this problem, however, has been a focus within WATERS.

The Swedish assessment criteria and guidance documents (SwAM 2013; SEPA 2007) contains some general guidelines as well as BQE-specific information on how uncertainty should be addressed within the Swedish assessment procedures. Furthermore, most BQEs provide some form of minimum requirements for the spatial and temporal allocation of samples in order for the assessment criteria to be formally valid. Inspection of these documents, however, reveals that they do not cover all aspects of the requirements defined by the Directive and that developing additional guidelines and coherent methods could substantially improve the assessment of uncertainty and, even more importantly, achieve more reliable and relevant estimates of ecological status.

First, it is notable that for none of the BQEs do the assessment criteria provide guidance on how to estimate average status for the most important spatio-temporal resolution, i.e. the whole assessment period for an entire water body. For some BQEs, it is stated that data from at least three years are required, which means that averages across years may be estimated, but for most BQEs, only yearly averages appear to be estimated and requirements for sampling during multiple years are not mentioned. Consequently, no guidance is provided on how to assess uncertainty at the assessment-period scale. Considering the complex structure of monitoring programmes, involving, for example, replicate years, sites, and samples, required to achieve representative estimates, such procedures are far from trivial.

Another potential problem is that different measures of precision are used for different BQEs. All the indicators for lakes and streams employ the standard deviation as a measure of precision, while for BQEs in coastal waters, the criteria make no mention of measures of precision, use a one-sided bootstrap confidence interval, or use the standard deviation as a measure of uncertainty (SwAM 2013). Recommendations about the desired level of precision are given for only one BQE. Furthermore, in lakes and streams, standard deviations indicative of methodological uncertainties are given in tables (except in the case of fish, for which an empirical formula is given). Some of these differences may relate to differences in sampling designs (i.e. differences in spatial and temporal replication). Nevertheless, this diversity in recommendations and routines clearly could obscure the coherent assessment of uncertainty and lead to arbitrariness in the handling of uncertainty in whole-system assessment.

The WATERS uncertainty framework relieves confusion

These and other shortcomings, inconsistencies, and guideline oversights regarding how to assess uncertainty have likely contributed substantially to the relative confusion about how to account for uncertainty in Swedish status assessments according to the WFD. In the practical application of assessment criteria during the two first management cycles, quantitative assessments of precision and confidence in classification have been practically absent despite the requirements defined in the Directive.

To address these issues, WATERS has proposed developing a unified framework for estimating relevant components of variability and for combining several relevant components with information on monitoring designs. This allows the more realistic estimation of precision and confidence than is permitted by current procedures and opens prospects for harmonising the handling of uncertainty. Furthermore, this framework provides a solid foundation for attempts to reduce uncertainty by optimising monitoring designs and incorporating important environmental factors as covariates.

3. Inland WATERS

3.1 General introduction to inland waters

As mirrors of landscape, lakes and watercourses reflect and integrate the effects of activities occurring in their catchments. Human activities have altered many of the local and regional drivers of biodiversity and ecosystem function, with freshwaters considered among the most perturbed habitats on Earth (Giller et al. 2004; Dudgeon et al. 2006). As lakes and streams nested in terrestrial surroundings are highly valued for their scenic and aesthetic values, as well as providing a number of valued ecosystem services, including water purification, nutrient cycling, flood regulation, and recreational activities, alterations or losses of these valued habitats have strong ramifications for our wellbeing.

Globally, the main factors affecting freshwater habitats include pollution, land use, water abstraction, altered hydrology and geomorphology, and invasive species. Eutrophication is considered the most pervasive pressure affecting the biodiversity and functioning of lakes (Schindler and Vallentyne 2008), while altered hydrology and morphology as well as elevated nutrients are considered the most pervasive pressures affecting watercourses. In Sweden, other pressures such as forestry and airborne pollution (e.g. acidifying compounds and Hg deposition) affect the biodiversity and ecosystem services of many surface waters. Although monitoring programmes have largely emphasised single pressures, there is increasing awareness that freshwater ecosystems are affected by multiple pressures from catchment land use, airborne contaminants, and direct pollution (e.g. Young et al. 2005; Johnson et al. 2006; Strayer and Dudgeon 2010). For example, agriculture results in multiple stressors affecting the integrity of aquatic ecosystems: elevated nutrient concentrations result in increased productivity and decreased oxygen concentrations (Hynes 1970; Tonkin and Death 2012); inputs of fine sediment result in loss of in-stream habitat (Piggott et al. 2012); pesticides affect populations of non-target species (Köhler and Triebskorn 2013); and water abstraction alters hydrology and connectivity (Lange et al. 2014). With the emerging view that multiple pressures are prevalent and that interactions among stressors are complex, there is growing awareness that management decisions based on individual stressors alone will likely result in inappropriate programmes of measures, such as remediation of the wrong pressure.

According to the Water Framework Directive (WFD), multiple taxonomic groups are to be used in monitoring programmes for classifying the ecological status of each water body (European Commission 2000). Using biological quality elements (BQEs), the Directive states that water bodies should have or achieve good ecological status. High ecological status is defined as follows: "There are no, or only very minor, anthropogenic alterations to the values of the physico-chemical and hydromorphological quality elements for the surface water body type from those normally associated with that type under undisturbed conditions. The values of the biological quality elements for the surface water body reflect those normally associated with that type under undisturbed conditions, and show no, or only very minor, evidence of distortion". Good ecological status is defined as follows: "The values of the biological quality elements for the surface water body type show low levels of distortion resulting from human activity, but deviate only slightly from those normally associated with the surface water body type under undisturbed conditions". Water bodies that display moderate or more severe deviation than that normally associated with the water body type need to be restored to at least good ecological status.

The use of biological variables in monitoring aquatic integrity has a long history in Europe (e.g. Metcalfe 1989). With the implementation of the WFD, multiple BQEs are now being used in biomonitoring (e.g. Birk et al. 2012) and many countries have developed and implemented assessment tools using the five BQEs (e.g. Poikane et al. 2016), for example: phytoplankton and benthic diatoms have been used in assessing the effects of acidification and eutrophication (e.g. Battarbee et al. 1997; Kelly et al. 2009); macrophytes and fish are considered reliable indicators of alterations to flow and habitat quality (e.g. Gorman and Karr 1978; Bain et al. 1988; Tremp and Kohler 1995); and benthic invertebrates are commonly used in monitoring the effects of organic pollution, acidification, and alterations in hydromorphology (e.g. Armitage et al. 1983; Buffagni et al. 2004; Sandin et al. 2004; Poikane et al. 2016).

In a recent review of methods for assessing the ecological status of surface waters in the Nordic countries, Anderson et al. (2016) found that many methods are currently used and therefore recommended a more harmonised approach in order to better understand the transitions between water body categories.

More robust tools for assessing the ecological status of Swedish inland waters needed

In 2007, Sweden revised its national classification schemes to better meet the requirements of the WFD (SEPA 2007). Although major advances were made in developing robust tools for assessing the ecological status of lakes and streams using community composition, a number of challenges remained. In brief, these can be summarised as: (i) knowledge of the responses of different taxonomic groups to different pressures, (ii) characterisation and harmonisation of reference conditions, and (iii) quantification of the uncertainties associated with measuring taxonomic composition and response indicators and classifying ecological status.

To design and implement cost-effective management programmes and to scientifically underpin conceptual models of human-induced changes in aquatic ecosystems, more knowledge of the response signatures of taxonomic groups to common pressures is needed. Response signatures of taxonomic groups from different habitats and to different pressures were assessed in WATERS using the precision (coefficient of determination, R^2) and sensitivity (magnitude of change, slope) of stress-response relationships. Large-scale datasets were created by combining biological and environmental data from regional and national monitoring programmes as well as from the dedicated field study of selected pressures (gradient studies).

The effective management of aquatic resources requires knowledge of when a water body differs from its natural condition and what has caused the divergence from the expected unimpaired condition. Reference conditions are increasingly being used to gauge the effects and magnitude of human intervention. A number of problems have emerged, however, in the use of reference conditions. For instance, as discussed below, there is now a growing body of literature indicating that classifications based solely on large-scale, landscape-level predictors are insufficient to capture the fine-scale variability of aquatic communities. In WATERS we have revised the pressure-filter approach used to define reference conditions and quantified the efficacy of typology- and model-based approaches for establishing reference conditions.

The use of biological quality elements and indicators requires knowledge of the precision and confidence of these measures. Here we analysed various sources of uncertainty, focusing on natural and pressure-induced variability. Uncertainty measures of all BQEs were assessed using a harmonised approach for quantifying variance components, culminating in the construction of a library of measures of uncertainty. In addition, in quantifying the uncertainty associated with macrophyte sampling, we tested the use of remote sensing with unmanned aircraft systems. This method provides sub-decimetre resolution of aerial images that can be used for mapping macrophyte stands. For benthic diatoms, we tested the response of indicators at both large and small spatial scales in both lakes and streams, as well as the use of the BenthoTorch for quantifying algal biomass using chlorophyll-*a* and measures of other pigments. For fish, the emphasis was on the uncertainty associated with the use of abundance measures such as biomass and number of fish in individual gill-nets.

Aims and approaches

The main aims of the inland waters part of WATERS have addressed knowledge gaps concerning pressure-response relationships and uncertainties associated with using different BQEs in ecological assessment. Briefly stated, although the specific aims varied among the BQEs, the overall aims can be summarised as to:

- collate existing biological and environmental data from regional and national authorities
- improve methods for field sampling
- agree on a harmonised approach for classifying reference conditions
- harmonize statistical approaches used in data analyses
- validate existing indicators and, where necessary, suggest new indicators
- quantify the response of different BQEs to selected pressures
- quantify uncertainty associated with the use of BQEs

The following sections review the methodology currently used to assess the ecological quality of lakes and streams using phytoplankton, benthic diatoms, macrophytes, benthic invertebrates, and fish. For each BQE, large datasets were created by combining data from regional and national authorities. A common understanding of what constitutes minimally disturbed, reference conditions was agreed upon and the WATERS inland group revised the pressure filter accordingly. In revising the pressure filter, consideration was given to information readily available (from the Water Information System Sweden, VISS) or easily obtained (i.e. water chemistry and land use data) (see Ten pressures in proposed reference filter, Section 5.1). Threshold values for water quality were updated using national classification criteria and new parameters (i.e. forestry, alteration of hydromorphology, and invasive species) were included as exclusion criteria.

3.2 Phytoplankton in lakes

OVERVIEW OF CURRENT INDICATORS

Sensitivity to nutrients and high turnover rates make phytoplankton good indicators

Phytoplankton are well known indicators of eutrophication, as some parts of the phytoplankton community reproduce very quickly at high nutrient concentrations and often accumulate at the surface of the water in algal blooms. Such algal blooms often consist of cyanobacteria that can be toxic to animals, including humans, which drink or are exposed to them. In addition, a high biomass of phytoplankton increases the risk of oxygen deficiency in the water, for example, due to respiration at night and when the biomass eventually decomposes. This especially occurs in the bottom waters of lakes that are temperature stratified and may also happen throughout the water column at night in small shallow lakes, affecting almost all higher trophic levels in the lake (Brönmark and Hansson 2005). The response to nutrients is direct as phytoplankton assimilate nutrients directly from the water for use as building blocks during energy fixation and growth. The high growth and turnover rates of phytoplankton make them good indicators for other environmental changes as well, especially as phytoplankton communities typically have a high diversity with species-specific requirements, though those relationships have been less well studied.

Table 3.1. Phytoplankton indicators used in assessing the ecological quality of lakes in Sweden and changes suggested by WATERS.

System	Indicator	Acronym	Pressure	Suggested revisions and comments
Lake	Total Biomass	TotBio	Nutrient load	No revision recommended; harmonize with new typology
Lake	Trophic Plankton Index	TPI	Nutrient load	Recommended revisions: use new European Plankton Trophic Index (PTI) instead; harmonize with new typology
Lake	Proportion of cyanobacteria	%Cyano	Nutrient load	Recommended revisions: use total biomass of cyanobacteria; develop class boundaries related to health risks
Lake	Number of taxa	Taxa	Acidity	No revision recommended
Lake	Chlorophyll-*a*	Chla	Nutrient load	Recommended revision: harmonize with typology and start to use interchangeably with TotBio, averaging when both indicators are available

Pressures other than eutrophication need to be evaluated

According to the WFD, three aspects of phytoplankton should be used to assess the ecological status of a water body: biomass or abundance, composition, and bloom frequency and intensity (European Commission 2000, Annex V). The Swedish implementation of the WFD currently involves the use of total phytoplankton biomass, Trophic Plankton Index (TPI), and the percentage of cyanobacteria to assess the ecological status of lakes and the ecological effects of nutrient loading to lakes (SwAM 2013).

The main pressure emphasised by the WFD is eutrophication. However, there is also a need to use biological indicators in assessing other types of pressures. For example, in Sweden, the number of phytoplankton taxa is used as an index to classify lakes in different acidity classes (Willén 2007; SEPA 2010). Furthermore, pressures related to human activities now available from maps, such as mining, forestry, farming, and urbanisation, have to be tested regarding their effect on phytoplankton community composition.

To establish reference conditions, we evaluated whether site-specific reference values based on models using similar parameters as used for the WFD typology, i.e. altitude, mean depth, alkalinity, and humic content, can be used instead of using the same reference value for all lakes within a type. Site-specific reference values may reduce the uncertainty in assessments of lakes near typology-based borders and of rare lake types.

OVERVIEW OF INDICATOR DEVELOPMENT

Data on phytoplankton and water chemistry from all over Sweden

In WATERS, we used a large recently compiled phytoplankton dataset comprising data on 806 lakes for which there are data from at least one year between 2000 and 2012. This dataset is based on monitoring data from national and regional monitoring programmes all over Sweden, where the participating laboratories and county boards have uploaded phytoplankton and sometimes water chemistry data to the national data host at http://miljodata.slu.se/mvm/. All lakes were sampled in a similar way using a standard protocol and analysed using certified laboratories. In developing indicators, we focused on the summer months July and August. August is the dominant sampling month for phytoplankton in Swedish lakes, and most lakes are sampled for phytoplankton only once a year. The statistical analyses focused on community composition and indicator responses to TP, but TN and pH responses were studied as well. The effects of pressure exerted by forestry and mining or of nutrient pressure exerted by various land uses, measured, for example, as percentages of agricultural and urban area, were also tested.

New phytoplankton indicators displayed better relationships to environmental pressures

The intercalibration exercise showed that chlorophyll-*a* was a good indicator with a strong relationship to nutrient-enrichment pressures (Phillips et al. 2008; Carvalho et al. 2012). In WATERS we compared the current Swedish indicator total phytoplankton biomass with chlorophyll-*a*, and demonstrated that both indicators displayed a strong positive linear relationship to TP and that the amount of variance explained was higher for chlorophyll-*a* (80%) than for total biomass (65%). Land use data representing typical eutrophication pressures were also analysed. Farmed areas explained more of the variation than did urban areas and, when combined, these pressures explained 31% of the variation in chlorophyll-*a* ($n = 385$) and 28% of the variation in total biomass of phytoplankton ($n = 761$).

The current Swedish Trophic Phytoplankton Index (TPI) had a slightly stronger correlation with TP than did the newly developed European PTI (Plankton Trophic Index, PTI; Phillips et al. 2012; Figure 3.1). However, TPI seems to be able to have almost any value in the middle of the TP gradient and therefore has greater error than does PTI, i.e. a RMSE of 0.96 compared with 0.31. This is because TPI overlooks indicator species with growth optima in this part of the nutrient gradient, meaning that it consists mainly of very sensitive or very tolerant species. The same pattern with large variation in the middle of the pressure gradient for TPI is seen when it is compared with the proportion of farmed and urban areas in the catchment, in which case the European PTI performs better (43% of variance explained and RMSE = 0.38 vs. 42% and 1.17, respectively, for TPI). The advantage of the European PTI is that it includes score values over the whole range of nutrient concentrations, whereas the Swedish TPI is missing score values in the middle of the pressure gradient, which leads to unnecessarily high variation.

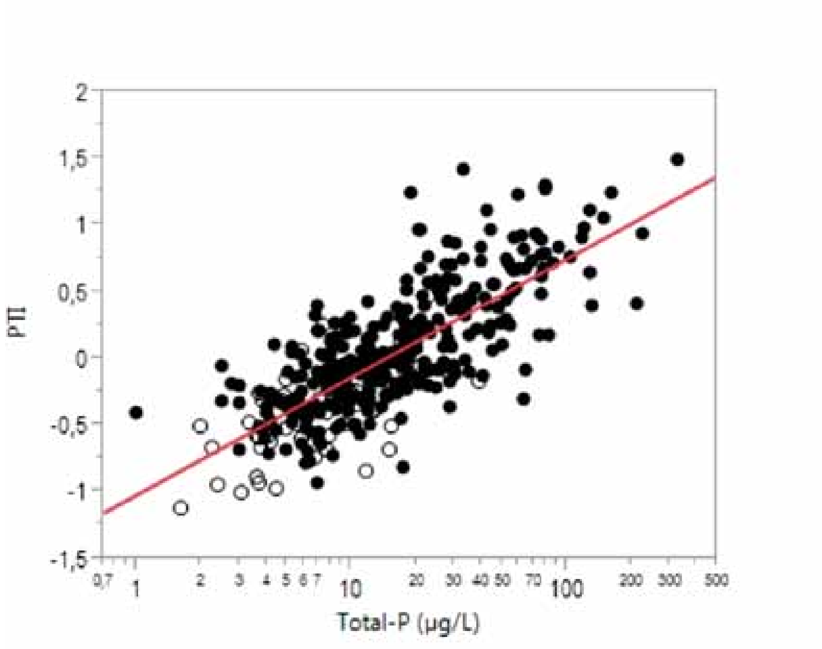


Figure 3.1: Relationship between the European PTI and total phosphorus in 361 Swedish lakes (mean values, July – August 2000 – 2012). White dots indicate lakes that pass the criteria for reference lakes. PTI = -1.02 + 0.385 log(Total-P), $p > 0.0001$, adj $R^2 = 0.588$, RMSE = 0.31.

Blooms of cyanobacteria are related to health risks

In oligotrophic lakes with a low biomass of phytoplankton and a relatively high biomass of non-blooming cyanobacteria, the proportion indicator including cyanobacteria has not worked well. Using the WATERS dataset, we found that cyanobacterial biovolume performed better along nutrient gradients than did cyanobacterial proportion, but not as well as did the indicators for total biomass and composition. At low nutrient concentrations, the biomass of cyanobacteria is always low, but it is clear that at higher nutrient concentrations cyanobacterial biomass can be either high or low, resulting in very low predictability. Instead, cyanobacteria may be used to identify the nutrient concentration or land use where there is a risk of having a certain biomass of cyanobacteria, as suggested by Carvalho et al. (2012), i.e. the risk of having cyanobacterial toxins in the water. The World Health Organization has identified the cyanobacterial concentrations at which there is risk of health-related problems involving cyanobacteria. WHO uses three levels of safe practice (WHO 2003):

1) Relatively low probability of adverse health effects at 20,000 cyanobacterial cells mL^{-1} or 10 µg chlorophyll-*a* L^{-1} with dominance of cyanobacteria, with short-term adverse health outcomes such as skin irritation and gastrointestinal illness.

2) Moderate probability of adverse health effects at 100,000 cyanobacterial cells mL^{-1} or 50 µg chlorophyll-*a* L^{-1} with dominance of cyanobacteria. The added risk is potential for long-term skin irritation and gastrointestinal illness.

3) High probability of adverse health effects when there is cyanobacterial scum formation in areas where whole-body contact and/or risk of ingestion/aspiration occur. The added risk is potential for acute poisoning.

Cell numbers are then converted to cell volumes using the volume of a typical cell such as *Microcystis* with a cell diameter of 4.5 µm (Hillebrand et al. 1999). For example, Norway has adopted this reasoning and has developed class boundaries at 0.16 (high/good), 1.00 (good/moderate), 2.00 (moderate/poor), and 5.00 mg cyanobacteria/L (poor/bad; Norwegian Environment Agency 2013). In the WATERS dataset, we found that at total phosphorus levels higher than 20 µg L^{-1} or total nitrogen concentrations higher than 500 µg L^{-1} there is a risk of health-related problems involving cyanobacteria in Swedish lakes.

Agriculture, mining, and urbanisation have strong impacts on the phytoplankton community

Both anthropogenic and natural variables are important when describing phytoplankton composition. Alkalinity, water colour, and altitude were the most important of the natural variables; of the human pressures available from maps, the extents of intensive agriculture, mining, and urban areas were the most important. Because information on metals was virtually absent from the dataset, it is not known whether the effects of mining are caused by nutrients or metals. Furthermore, models including water chemistry, such as nutrient or pH levels, were always stronger than those without, highlighting the importance of the chemical environment for phytoplankton (Drakare , in manuscript).

Typology-based versus site-specific reference values: a matter of choice

Type-specific reference values have been discussed and agreed upon at the European level (Lyche Solheim et al. 2014), and the typology used for assessing ecological status in lakes in Sweden was modified accordingly after the intercalibration exercise (SwAM 2013). In general, type-specific reference values are low, being well correlated with the ratios calculated between observed and expected reference values as well as with the ecological quality ratio (EQR) and the pressure gradient (Lyche Solheim et al. 2014). Useful models could be built for chlorophyll-*a*, total biomass, and PTI using similar variables as for typology groups, i.e. altitude, latitude, longitude, catchment area, lake area, lake mean depth, absorbance, and alkalinity. The models produced EQR values more weakly related to the pressure gradient than did using type-specific reference values; on the other hand, these models take into account all kind of lakes, including highly humic or alkaline lakes that were not included in the intercalibration.

Descriptive data are missing for most lakes

Only 16% of the 806 lakes could be categorised according to the eight typologies used during the EU intercalibration. Highly humic and mountain lakes were excluded from this comparison and, in addition, mean depth was not available for many lakes. In WATERS we also typed according to a suggested new Swedish lake typology with 48 types in theory (Drakare 2014). We were able to group 267 lakes or 33% into 19 of the suggested types in the proposed new Swedish typology (Supplementary information). Reference lakes were only found in half of the types, making it difficult to estimate reliable type-specific reference values for most of the lake types in the new typology. Using models, it was possible to calculate site-specific reference values for 30 – 40% of the 806 lakes depending on the index. Thus, in both cases there is a need to collect descriptive data for lakes either to group lakes into the correct type for obtaining type-specific reference values or to use in models for calculating site-specific reference values.

SUGGESTED REVISIONS OF ASSESSMENT METHODS

Chlorophyll-*a* is a better indicator of eutrophication than is total biomass

As chlorophyll-*a* had a stronger or similar relationship to nutrient load-related pressure variables as did the total biomass of phytoplankton, this indicator may well be used interchangeably with total phytoplankton biomass in status assessments. As three Member States use chlorophyll-*a* and phytoplankton biovolume interchangeably or average them when both indicators are available (Carvalho et al. 2012), we suggest doing so in Sweden as well. The current status of the chlorophyll-*a* indicator for assessing ecological status in lakes should be raised from being indicative to being a full indicator to be included in the overall status assessment.

PTI is a robust indicator of phytoplankton composition

We suggest starting to use the European PTI in Swedish assessment tools after developing class boundaries and setting reference values for this indicator. The PTI addresses the genus level and may be developed by including certain key taxa at the species level for genera that have both sensitive and tolerant species; however, it is currently unclear whether that is necessary, as PTI seems robust enough as it is.

Algal blooms can be assessed with biovolume of cyanobacteria

We suggest using the biovolume of cyanobacteria as the indicator to assess algal blooms until high-frequency methods to measure bloom frequency and intensity become available for large-scale lake monitoring. Cyanobacteria can be used to identify whether there is a risk of cyanobacterial toxins in the water. If cyanobacterial biovolume is to be used to indicate algal blooms, we suggest developing class boundaries that relate to health problems, i.e. corresponding to the levels suggested by the WHO. This means that neither type- nor site-specific reference values may be needed for this indicator, as cyanobacterial class boundaries will be the same all over Sweden if they are related to health risks. The UK and Norway use the indicator only to lower the status of a lake, not to increase the lake status in the absence of cyanobacteria (WFD-UKTAG 2014; Norwegian Environment Agency 2013). We suggest using cyanobacteria in a similar way in Swedish assessment tools.

New data needed to predict reference conditions for lakes

In addition to chlorophyll-*a* and phytoplankton species composition, more information is necessary to perform assessments, namely, lake mean depth, alkalinity, and water colour. These data enable the prediction of the reference conditions needed to calculate the EQRs and allows the lake to be assigned to a type. Each lake needs to be assigned to a type to be able to report according to the WFD. As these data are crucial for both assessment and reporting, we suggest that they be requirements when designing monitoring programmes and for reporting to the national data host.

REMAINING CHALLENGES

More background data needed to estimate reference values

The main challenge when it comes to phytoplankton is to estimate reference values. To estimate the reference values, we need background data (i.e. lake mean depth, alkalinity, and water colour) for each lake to be able to use typology or use the data in models. Mean depth is time consuming and expensive to measure but has to be done only once for each lake. The missing water chemistry data in some cases have probably been measured but not uploaded to the data host; in other cases, an extra sample needs to be added for alkalinity and water colour analysis when taking the phytoplankton sample. The data availability is not currently good enough to be able to type most lakes, which means that many lakes cannot be assessed at all. This is also true of most reference lakes, 78% of which could not be typed according to the suggested new typology, meaning that we could use phytoplankton data from only 31 of 139 reference lakes. There are probably enough reference lakes to be able to develop both reliable type- and site-specific reference values as soon as we have the needed information. For rare lake types, models will likely be needed in order to set reference values. Discussion with authorities is needed in order to develop a strategy for obtaining the needed lake information.

Selecting natural conditions from modified ones is a challenge

Phytoplankton indicators are well correlated with both nutrient and acidity gradients, so the challenge is to select which nutrients or acids are the result of human activities. Soils in agricultural areas are usually naturally rich in nutrients, and this must be taken into account when setting reference values. In the WFD, high alkalinity has been used as a proxy for natural nutrient-rich conditions, but high production of phytoplankton due to human activities can also increase the alkalinity. It would be much better to use detailed soil maps and combine the data derived from them with a method to estimate the natural phosphorus retention when estimating the reference phosphorus concentration for each lake. Chemistry assessment tools will give some of the answers and should be taken into account when setting reference conditions for phytoplankton. For example, the current phosphorus assessment tools (SwAM 2013) use similar variables that worked well in the WATERS programme to describe natural conditions for phytoplankton – i.e. water colour, lake altitude, and lake mean depth. However, the phosphorus assessment tools for lakes need to be updated with phosphorus reference conditions for lakes in agricultural areas as has been done for running waters.

The results from comparing phytoplankton with land use may also be developed further, as interesting patterns were found for land use categories paid little attention before, such as mining and forestry.

For the acidity gradient, the current phytoplankton indicator gives an answer only about the acidity not the acidification. We have not found a phytoplankton indicator that works better than this or works for acidification, not acidity. In Sweden, the Model of Acidification of Groundwater in Catchments, MAGIC (Moldan et al. 2004), has been adopted as the official chemical assessment tool for classifying waters affected by acidification. MAGIC predicts the preindustrial pH, and lakes where current pH values differ from preindustrial levels by more than 0.4 are considered acidified (Fölster et al. 2007).

In conclusion, the suggested new indicators for phytoplankton are promising and will work well when combined into a multimetric index. This can be done as soon as we have the required additional data of lake mean depth, alkalinity and water colour, at least for lakes that can be classified as references, and chemical assessment tools to be able to develop reliable reference conditions and class boundaries.

3.3 Benthic diatoms in streams and lakes

OVERVIEW OF CURRENT INDICATORS

Diatoms are strongly affected by their chemical environment

Diatom indicators are widely used in environmental assessment, for example, in many European countries (Kelly et al. 2009) and elsewhere (Stevenson et al. 2007), due to the tight coupling between diatom assemblages and the chemical environment. The Swedish assessment criterion for benthic freshwater diatoms is based on two main indicators: IPS and ACID (Kahlert et al. 2007, 2008; Kahlert and Jarlman 2009). IPS stands for Indice de Polluo-sensibilité Specifique and is an indicator of the impact of nutrients, primarily phosphorus, and organic pollution (Cemagref 1982). This index is combined with two supporting parameters, i.e. per cent of pollutant-tolerant taxa (%PT) and the Trophic Diatom Index (TDI) (Kelly 1995, 1998), where %PT indicates organic pollution and TDI indicates the impact of nutrients. The IPS value is used to classify the ecological status and the supporting parameters verify the assessment (Kahlert et al. 2007, 2008; Kahlert and Jarlman 2009).

ACID stands for the ACidity Index for Diatoms and is an index created to separate different acidity classes (Andrén and Jarlman 2008). ACID can also be used to find waters at risk of being acidified by anthropogenic impacts. However, to assess whether the acidity is natural or caused by humans, it is necessary to continue the assessment using chemical models such as MAGIC (Kahlert and Jarlman 2009). The Swedish standard for diatoms provides a measure of the uncertainty of IPS and ACID and briefly describes its use in ecological status classifications (Kahlert et al. 2007; SEPA 2007).

Table 3.2. Benthic diatom indicators used in assessing the ecological quality of streams in Sweden and changes suggested by WATERS.

System	Indicator	Acronym	Pressure	Suggested revisions and comments
Streams	Indice de Polluo-sensibilité Specifique	IPS	Nutrient load and organic pollution	Use updated taxon list, as well as the updated indicator values, for assessment of both streams and lakes
Streams	Percent of pollutant-tolerant taxa	%PT	Organic pollution	Use updated taxon list, as well as the updated indicator values, for assessment of both streams and lakes
Streams	Trophic Diatom Index	TDI	Nutrient load	Use updated taxon list, as well as the updated indicator values, for assessment of both streams and lakes
Streams	Number of taxa	ACID	Acidity	Use updated taxon list, as well as the updated indicator values, for assessment of both streams and lakes.
			Heavy metals and herbicides	Add this indicator as a new method of using diatoms to detect pollutants in routine environmental assessment

The method for monitoring benthic diatoms currently used in Sweden is based on European standards (CEN 2003, 2004), where the sampling and handling of samples in the laboratory are standardised for stream sampling. These methods are used to assess the status of streams according to the Swedish assessment criteria. Recently, however, European monitoring standards have been updated to include sampling in lakes as well (CEN 2014a, 2014b). These updates have been implemented in the Swedish monitoring guidelines (Jarlman et al. 2016), but benthic diatoms are currently not part of the Swedish assessment criteria for lakes.

Current methods for diatom assessment are demanding

In 2010, The Swedish Environmental Protection Agency (SEPA) judged the Swedish diatom method to be a cost-effective and reliable indicator of nutrient impact, organic pollution, and acidity (SEPA 2010). SEPA criticised the method for its dependency on particular software, its need for very experienced analysts, and the potential difficulty of finding suitable substrates for the collection of diatom samples. Water managers requested further development of a method for lakes, for detecting pesticide and metal impacts, for Nordic stream types, and for sulphide clay stream types. Concerns were also raised about the fact that the class boundary between good and moderate status occurred at a different phosphorus concentration from the boundary defined by the chemical assessment criteria.

Several important diatom assessment challenges have been identified as important for Sweden and other countries (Rimet 2012; Kelly 2013). First, diatom assessment has largely focused on the impact of nutrients. The response of existing indicators to other stressors is less well understood, or no relevant indicators have been developed at all. For example, although the impact of organic pollution can be detected by IPS, it is still not well understood why, and under which conditions, some species thrive in organic pollution while others do not. There are indicators such as ACID for assessing the impact of acidity, but the combined effect of acidity and nutrients is difficult to assess. For heavy metals or pesticide impacts, no indicators have yet been developed. Overall, indicator development has traditionally been based on statistical modelling, with relatively little conceptual support.

Finally, the traditional way of using only diatoms in environmental assessment does not really fit the need to assess changes in ecosystem functions and ecosystem services, which would at least require the analysis of the entire group of phytobenthos, not only diatoms.

OVERVIEW OF INDICATOR DEVELOPMENT

The current diatom method thoroughly tested in WATERS

The WATERS programme has addressed these points of criticism and these challenges. We have verified the existing method and assessed its uncertainty, tested it for lakes, and have also partly addressed new issues such as the impact of new stressors and the assessment of the entire benthic algal community. In detail, we tested whether the current diatom method responds reliably to the impact factors for which it was developed, i.e. nutrients and pH, to verify its use for environmental assessment. We updated the Swedish diatom taxon list, and checked whether all taxa have indicator values corresponding to the ecological backgrounds where the taxa are found. We also tested how the current methods perform in comparison with methods using other organism groups, to formulate guidelines as to which biological indicators should be used in environmental assessment. Furthermore, we tested whether the current stream diatom method can be used in lakes as well, as this was one of the identified major gaps regarding the use of diatoms for environmental assessment. We then started to develop new diatom indicators to assess other stressors not currently included in the assessment, such as hydromorphology, forestry, and pollution. Finally, we assessed the entire benthic algal assemblage, not only diatoms, to be able to correlate the response of the diatom method with the response of the entire algal community, which, from an algal perspective, is the relevant ecological component.

European approaches to diatom assessment largely in agreement

In a review of the use of benthic diatom assemblages and sampling methodology in environmental assessment, we found broad-scale agreement between assessment methods used around Europe (Kelly et al. 2009; Kelly et al. 2014 and references therein). The same standards for diatom sampling and assessment are used around Europe, and although indices may differ slightly, the impact of high nutrient concentrations is assessed similarly.

For diatom work within the WATERS programme, data compiled from national and regional monitoring programmes were used. These data and updates are now available at the national data host at SLU (http://miljodata.slu.se/mvm/). We also updated the Swedish diatom method for environmental assessment with an updated list of Swedish diatom taxa and their indicator values to be used for the calculation of IPS, TDI, %PT, and ACID (Jarlman et al. 2016; http://miljodata.slu.se/mvm/).

In agreement with expectations from international studies, we found that benthic diatom assemblages in Swedish streams responded mainly to pH and nutrient gradients (Kahlert and Trigal, submitted). Furthermore, as predicted, we found good correlations between the indicators ACID and pH and between IPS and total phosphorus (TP) (Figures 3.2 and 3.3). As for response to nutrients, the largest changes in diatom assemblage structure following a nutrient gradient were found at about a TP of 20 µg L^{-1}, which is the nutrient threshold earlier demonstrated to distinguish the good from high ecological status classes. At about TP 60 µg L^{-1} and higher, good ecological status is distinguished from moderate status (Kahlert et al. 2007).

We found that the assessment of nutrient status and acidity using the current Swedish diatom method entails relatively low error (Figures 3.2 and 3.3). Regarding ACID, 70% of the index variation is explained by

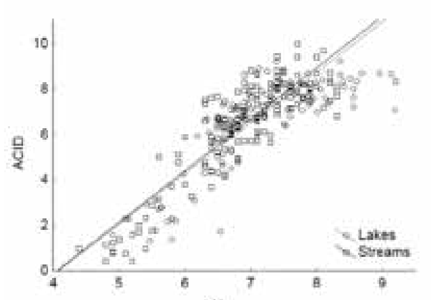


Figure 3.2: ACID as a function of pH at each sampling site. Fitting was restricted to pH < 8.4. Lake ACID: $p < 0.0001$, $r^2 = 0.67$; stream ACID: $p < 0.0001$, $r^2 = 0.74$

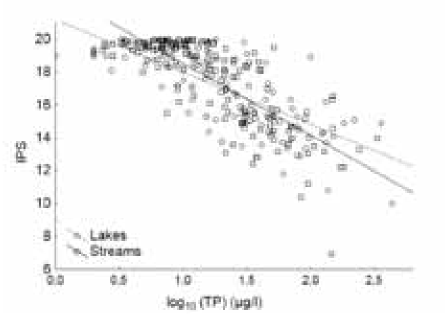


Figure 3.3: IPS as a function of total P. Lake IPS: $p < 0.0001$, $r^2 = 0.55$; Stream IPS: $p < 0.0001$, $r^2 = 0.63$

pH (r^2 = 0.74 for both streams and lakes). Regarding IPS, which was developed to reflect not only nutrients but also organic pollution, about 63% of the IPS variation is still explained by TP for both streams and lakes (Kahlert and Gottschalk 2014; Kahlert 2011). However, it must be kept in mind that IPS is only expected to relate closely to TP in the ecological status classes moderate to high, as IPS is not only an indicator of nutrients.

Diatom indicators respond similarly in lakes and streams

By comparing lakes and streams along environmental gradients, we could demonstrate that the Swedish diatom assemblages as well as the calculated indices respond significantly and similarly in both habitats (Kahlert and Gottschalk 2014). Thus, despite differences in taxonomic composition between streams and lakes, the resulting diatom assemblages are similar, enabling the use of the same diatom indicator in both habitats. The Swedish diatom method for environmental assessment was updated on the basis of this finding (Jarlman et al. 2016).

We found that pollution had a negative impact on the structure of diatom communities and on diatom valve formation. Indicators of a high load of heavy metals and herbicides were over 1% of diatom valves being deformed, a small number of diatom taxa in a sample (<20), and low species diversity (Shannon's index < 2). We accordingly developed an indicator for the use of benthic diatoms in screening for pollutants (Kahlert 2012). This indicator was implemented in the Swedish method (Jarlman et al. 2016) and is now used by water managers. However, the variation is large, and we need to know more in order to understand the response of benthic diatoms to pollutants.

The BenthoTorch has the potential for quantification

In seeking a simple way to measure benthic algal biomass and to quantify different algae groups, we tested a new instrument, called the BenthoTorch. We found that the instrument has potential for quantifying total algal biomass expressed as µg chlorophyll-*a* cm^{-2}, but that its output for the biomass of various algal groups must be used with great caution (Kahlert and McKie 2014). Moreover, because the necessary equipment is very expensive, there are also doubts about the cost-effectiveness of the method.

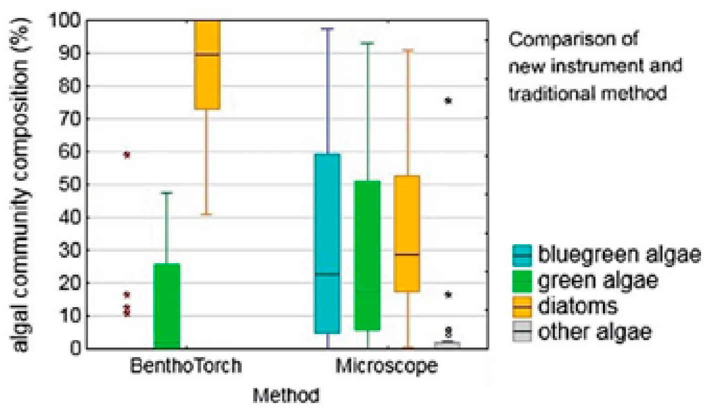


Figure 3.4: Algal group composition is assassed differently with the BenthoTorch and the microscope. From: Kahlert and McKie (2014).

SUGGESTED REVISIONS OF ASSESSMENT METHODS

Adjust indicator values of diatom taxa and apply the method in lakes

We suggest using the updated Swedish taxon list, as well as the updated indicator values, for the assessment of both streams and lakes. We base this recommendation on the finding that the current indicator responds as expected with low error levels for nutrients and pH. Furthermore, as many taxon names have been recently changed, we recommend the use of the updated list of Swedish taxa. Additionally, many TDI index values in the list were adjusted following a detailed evaluation of this index in Sweden TDI was developed for the United Kingdom, and is only now being evaluated in detail for use in Sweden.

Continue developing use of diatoms for detecting impact of pollutants and other stressors

We suggest including the new method of using diatoms to detect pollutants in routine environmental assessment to test how the new index should be incorporated into ecological status classification. We found that diatoms can indicate the impact of pollutants, especially of heavy metals, but also of herbicides. Regarding hydromorphology and forestry, more data are needed to assess the impacts of these pressures on diatoms.

Use microscopy to monitor benthic algae

The BenthoTorch results did not meet our expectations when assessing the biomass of all algal groups. Therefore, we recommend continued use of microscopy for the assessment of benthic algal community structure. The BenthoTorch was, however, found to provide a fast and cost-effective analysis of benthic chlorophyll-*a* in oligotrophic streams.

REMAINING CHALLENGES

Diatom flora typical of different stream types need to be defined

As early as 2007 we could demonstrate (Kahlert et al. 2007) that the reference values for unimpacted streams do not differ significantly between ecoregions, and therefore decided on a single value for all of Sweden. Similar results have been obtained for all of Europe (Kelly et al. 2009). Because certain stream types were underrepresented, in 2007 we were unable to test whether the limit between good and moderate status should differ between ecoregions, that is, whether the eutrophication of a boreal stream is perhaps indicated by higher IPS values than in the boreonemoral and nemoral regions (Illies region 14).

Unfortunately, in WATERS we were unable to study the differences between stream types using the new collected data, because we could not separate reference streams from impacted ones due to the fact that the compiled data do not contain information on the classification of impacts according to Swedish assessment criteria (SwAM 2013), such as assessments of eutrophication, acidification, or other pressures such as hydromorphological changes, liming, heavy metal impacts, and sulphide clay impacts. The dataset also contains too few data on certain stream types, and no data on lakes. We need to link diatom data to stressor data from the national data host (http://miljodata.slu.se/mvm/) and VISS databases (http://www.viss.lansstyrelsen.se). Furthermore, we would like to test the new suggested stream typology with benthic diatom assessments.

More detailed analysis of different stressors needed

More detailed analyses of the impact of single stressors are also needed. This would allow for better assessment of whether the biological classifications are in line with models classifying chemical impacts and with selection of the most sensitive organisms for certain stressors.

Finally, we need to harmonize our traditional knowledge of benthic algal ecology with new, soon-to-be-available methods for species identification based on DNA barcoding. These will replace or complement current microscopic methods and possibly also currently used species and indicators.

3.4 Macrophytes in streams and lakes

OVERVIEW OF CURRENT INDICATORS

Vascular plants the most studied taxonomic group

The application of macrophytes (i.e. aquatic vegetation including macroalgae, bryophytes, charophytes, and vascular plants) in describing lake properties has a long tradition, especially in Sweden (Svenonius 1925; Arwidsson 1926; Lohammar 1938, 1949, 1965; Wallsten 1981). In general, vascular plants are the most studied taxonomic group of the macrophytes and eutrophication is the most studied environmental pressure (Penning et al. 2008a, 2008b; Kolada et al. 2014), even though macrophytes are sensitive to other and multiple pressures (Rørslett 1991; Toivonen and Huttunen 1995; Ecke 2009; Mjelde et al. 2013).

Currently used indicator systems rely mainly on the sensitivity and tolerance, respectively, of individual species along the studied environmental gradients. This approach has resulted in indicators based on weighted averages of the indicator values of individual species estimated mainly at the presence/absence scale (Ecke 2007; Kolada et al. 2014) or indicators considering the number or abundance of sensitive and tolerant species (Schaumburg et al. 2004b, Penning et al. 2008b). The Irish assessment system (using the Free index) is truly multi-metric as it includes the maximum depth of colonisation as well as the relative frequency of *Chara* spp. and elodeids (Free et al. 2006).

Table 3.3. Indicators used in assessing the ecological quality of lakes and streams in Sweden and suggested changes.

System	Indicator	Acronym	Habitat, pressure	Suggested revisions and comments
Lake	Trophic Macrophyte Index	TMI	Whole lake, eutrophication	Recommended revisions: set ecologically meaningful class boundaries; use quantitative macrophyte data; and further develop remote sensing-based assessment
Stream				Indicator needs to be developed

Several limitations of current lake indicators

The limitations of current indicators and assessment systems can be summarised in the following four categories: neglect of bryophytes and macroalgae, linear response of indicators and metrics along the pressure gradient, field campaigns needed for data supply are resource intensive, and hydromorphological pressures are seldom implemented in national assessments.

In addition, the Swedish assessment system was developed for a small dataset that included only 49 reference sites as well as mainly historic (back to 1926) and only qualitative macrophyte data (i.e. presence or absence).

Mosses and macroalgae are central to the community structure of macroinvertebrates and lake ecosystem functioning (Brusven et al. 1990; Parker et al. 2007). Despite the importance of these plants, most studies of macrophyte-based assessment systems focus on vascular plants (Schneider and Melzer 2003; Penning et al. 2008b; Poikane 2009), include few bryophyte or macroalgae species, or pool species at lower taxonomic levels (i.e. genus or family) (Kolada et al. 2014). Few assessment systems explicitly include bryophytes and/or macroalgae (e.g. Pall and Moser 2009).

The applicability of indicators in water quality assessment is limited if the indicator response along a pressure gradient is linear. An important step in the assessment process is to identify threshold values of a pressure variable based on the responses of certain species, communities, indices, and/or metrics.

Commonly used field methods for sampling lake and river vegetation, such as transect methods (Baattrup-Pedersen et al. 2001; CEN 2004), are not only labour intensive and restricted to small spatial scales, but may yield inconsistent results (Dudley et al. 2012) due to spatial variability (Spears et al. 2009), the number of sampled transects (Leka and Kanninen 2003), and multiple observers (Staniszewski et al. 2006). Given technical advances in platforms and sensors from mono- to multi- and hyperspectral (e.g. Edwards and Brown 1960; Howland 1980; Valta-Hulkkonen et al. 2003), remote sensing has the potential to overcome many of the limitations of field methods.

OVERVIEW OF INDICATOR DEVELOPMENT

Responses to total phosphorus in focus

In WATERS, we tried to overcome the above four main obstacles supported by a large, recent, and quantitative dataset on macrophytes. The neglect of bryophytes and macroalgae as well as the responses of metrics along the pressure gradients were studied on a macrophyte dataset covering 233 Swedish lakes sampled in 2003 – 2013. The lakes were also sampled for water chemistry and distributed throughout the country, even though most lakes are located in south-central Sweden. All lakes were sampled using a transect method allowing for the calculation of frequencies per species and lake (CEN 2004) and since 2007, all sampling has followed the same protocol (n = 225) (SEPA 2015). The statistical analyses focused on species and indicator responses to total phosphorus (TP) concentrations in lakes, because TP is the most studied environmental pressure facing macrophytes in lakes (Penning et al. 2008a, 2008b; Poikane 2009; Kolada et al. 2014), although responses to hydromorphological pressures were studied as well. Other pressures (including pH) were initially evaluated, but when testing multiple pressures, it was difficult to identify the contribution of individual pressures to explaining macrophyte responses. In addition, the response of macrophytes appeared to be strongest along the TP gradient.

Improved class boundaries with ecological relevance

The correlation of the TP concentration in lakes with the existing Trophic Macrophyte Index (TMI, Ecke 2007) (r = – 0.63, P < 0.001) was weaker, though still highly significant, than with the Norwegian TIc index (Mjelde 2013) (r = – 0.76, P < 0.001) or the Intercalibration Metric (Kolada et al. 2014) (r = 0.79, P < 0.001). Previously, sudden drops in the abundance of large isoetids, including *Isoëtes spp., L. uniflora,* and *L. dortmanna,* have been used to identify the high/good boundary for the assessment of ecological status according to the WFD (Ecke 2007; Penning et al. 2008b). In contrast, the other boundaries, i.e. good/moderate, moderate/poor, and poor/bad , have in most European macrophyte-based assessment systems been identified using purely statistical approaches, for example, using percentiles or equal-distance methods (Poikane 2009; Birk et al. 2012), and suffer from a lack of ecological relevance. In WATERS, we were able to identify thresholds based on sudden drops in the frequency of several macrophyte species, which represent boundaries between different ecological status classes (illustrated in Figure 3.5).

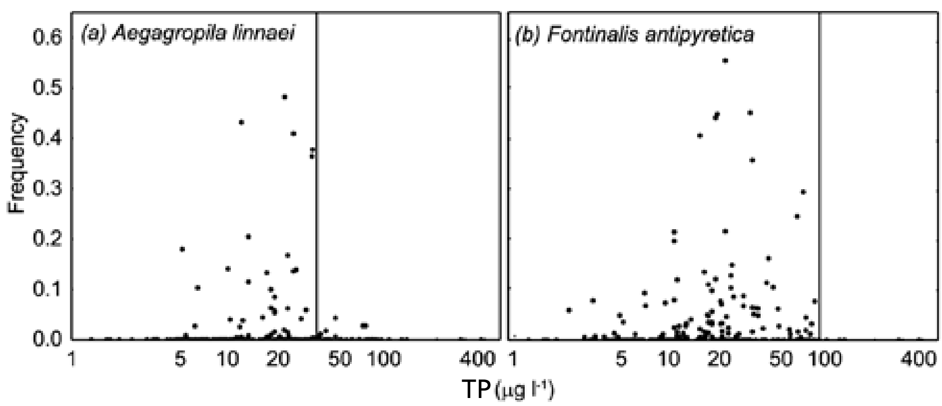
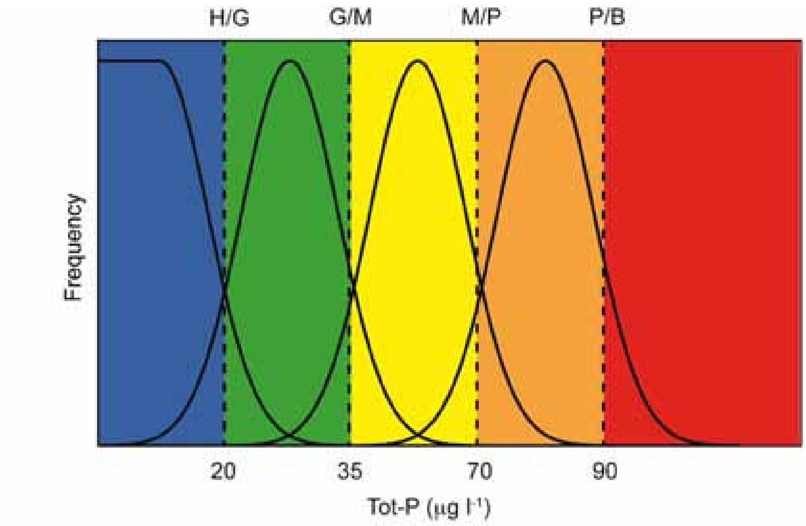


Figure 3.5: Frequency of the macroalgae *Aegagropila linnaei* (a) and the bryophyte *Fontinalis antipyretica* (b) along the gradient of total phosphorus (TP). Vertical lines indicate visually fitted sudden drops in frequency at 35 (a) and 90 µg TP l^{-1} (b), respectively.

Bryophytes in particular were found to be valuable for the identification of thresholds, as 13 of the total of 39 species displaying sudden drops (Figure 3.6) were bryophytes. These results suggest that there is potential to use sudden drops in the frequency of several macrophyte species in identifying all class boundaries. According to the suggested system, the high/good boundary is at 20 µg TP l^{-1}, the good/moderate boundary at 35 µg TP l^{-1}, the moderate/poor boundary at 70 µg TP l^{-1}, and the poor/bad boundary at 90 µg TP l^{-1} (Figure 3.6). The suggested high/good boundary, here supported by several macroalgae and bryophyte species, agrees well with the previously suggested boundary based on large isoetids (Penning et al. 2008b).



Species representative of boundary

High/Good 20 µg l-1	Good/Moderate 35 µg l-1	Moderate/Poor 70 µg l-1	Poor/Bad 90 µg l-1
Batrachospermum spp.	***Aegagropila linnaei***	***Fontinalis hypnoides***	*Alisma plantago-aquatica*
Calliergon spp.	***Calliergonella cuspidata***	*Potamogeton pusillus*	*Butomus umbellatus*
Drepanocladus longifolius	***Drepanocladus sordidus***	*Ranunculus circinatus*	***Fontinalis antipyretica***
Fontinalis dalecarlica	***Fissidens fontanus***	***Riccicarpos natans***	*Hottonia palustris*
Isoëtes echinospora	*Hippuris vulgaris*	*Sium latifolium*	*Potamogeton berchtoldii*
Juncus bulbosus	*Isoëtes lacustris*	*Sparganium erectum*	*Sagittaria sagittifolia*
Lobelia dortmanna	*Plantago uniflora*		*Stratiotes aloides*
Nostoc zetterstedtii	*Ranunculus peltatus*		*Stuckenia pectinata*
Ranunculus reptans	***Sarmentypnum exannulatum***		*Zannichellia palustris*
Scorpidium scorpioides	*Stuckenia filiformis*		
Sparganium angustifolium			
Sparganium gramineum			
Subularia aquatica			
Utricularia intermedia			

Figure 3.6: High/good (H/G), good/moderate (G/M), moderate/poor (M/P), and poor/bad (P/B) ecological status class boundaries along the total phosphorus gradient (TP) as suggested by sudden drops in the frequency of the studied macrophyte species. The class boundaries are illustrated by representative species displaying sudden drops in frequency at the respective boundaries. Bryophytes and macroalgae are indicated in bold.

Macroalgae and bryophytes especially contribute to high and good ecological status boundaries

The correlation between the trophic index (Tic) and TP concentrations was only marginally higher when calculated with, versus without, macroalgae and bryophytes (R^2 = 0.23 and R^2 = 0.22, respectively). However, the differences between the trophic index calculated with and without these species was higher at TP concentrations <35 µg TP l^{-1} (variance in δ-Tic: 236.60) than at TP concentrations ≥35 µg TP l^{-1} (variance in δ-Tic: 1.99). This implies that macroalgae and bryophytes contribute important information for identifying especially high and good ecological status. They should therefore be routinely integrated into the monitoring and ecological assessment of lakes. Macrophyte field sampling demands high taxonomic competence on the part of the observers; the existing focus on sampling primarily vascular and charophyte macrophytes in lakes probably reflects this need for high competence. The results presented here strongly suggest that observers of macrophytes in lakes should also be trained in macroalgae and bryophyte identification. At least among the bryophytes, several species can, after sufficient training, be identified in the field, while post-field identification of lake macroalgae is strongly recommended.

Great potential in remote sensing of ecological assessment

In a first step, we tested whether high-resolution (i.e. resolution approximately 5 cm) aerial images acquired with an unmanned aircraft system (UAS) could be used to produce vegetation maps at the species level and to estimate species abundance and biomass.

We identified the species composition of vegetation stands at lake and river sites with an overall accuracy of 94.6% and 80.4%, respectively, and it was feasible to produce a digital vegetation map (Husson et al. 2014a). The UAS method also offers great potential to accurately assess nutrient and trace-element cycling in the riparian zone, an important step in indicator development (Husson et al. 2014b).

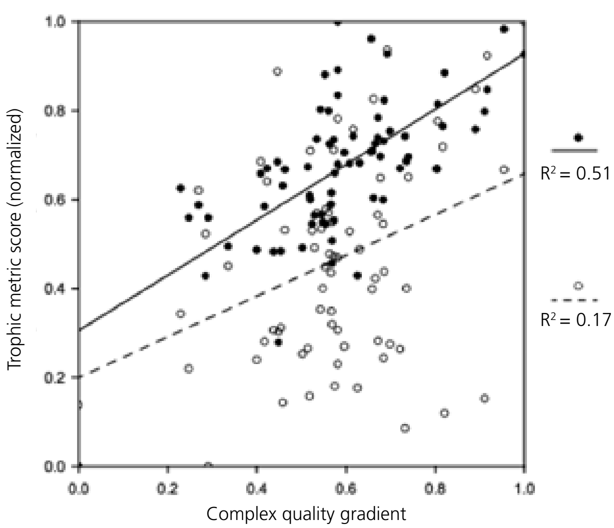


Figure 3.7: Relationship between a normalised trophic metric score and a complex water quality gradient. The metric was calculated for macrophyte taxa detectable with high-resolution remote sensing (filled circles) and for taxa sampled in the field (open circles). Adapted from Birk and Ecke (2014).

In a second step, we evaluated the potential of using only emerging, floating, and floating-leaved taxa detectable by high-resolution remote sensing (i.e. RS-taxa) to assess the ecological status of lakes (Birk and Ecke 2014). Average indicator scores were higher (indicating higher trophic status) for the index calculated with RS-taxa than for the index calculated with all taxa. Correlations of the trophic metric score and total nitrogen concentrations were equally strong for the dataset based on RS-taxa and on all taxa (R^2 = 0.28 and 0.26, respectively). For TP concentrations, the correlation was stronger for the dataset based on all taxa (R^2 = 0.35 and 0.14, respectively), but for a complex water quality gradient (including sulphate, N-species, chlorophyll, and per cent cover of wetlands in the riparian buffer), the correlation was higher for the RS-taxa dataset (R^2 = 0.45 and 0.12, respectively) (Figure 3.7) (Birk and Ecke 2014). Our results suggest that conventional field surveys could be replaced by high-resolution remote sensing at the sub-decimetre scale after successful calibration and validation.

WATERS developed an indicator for hydromorphological pressures

Based on a common Nordic dataset on regulated lakes, we developed a water-level drawdown indicator (WIc) using the ratio between sensitive and tolerant macrophyte species (Mjelde et al. 2013). The indicator correlates well with winter drawdown in Finnish (R^2 = 0.77), Norwegian (R^2 = 0.67), and Swedish (R^2 = 0.73) regulated lakes. The correlations were strongest with winter drawdown in storage lakes (i.e. lakes regulated for hydroelectric power and with a considerable winter drawdown). The WIc index is applicable to low-alkalinity, oligotrophic, and ice-covered lakes, and is suggested to be a useful tool to identify and designate heavily modified water bodies in Nordic lakes according to the WFD.

SUGGESTED REVISIONS OF ASSESSMENT METHODS

We suggest ecologically relevant class boundaries

In the present assessment system for the quality element macrophytes, only the H/G boundary was based on the response of macrophytes along the pressure gradient (Ecke 2007). All other boundaries were based on a statistical approach or were, in the case of P/B, even missing due to insufficient data. Our results enable the setting of ecologically meaningful boundaries between all ecological status classes, as indicated by the identified sudden drops of macrophytes along the pressure gradient. We suggest new boundaries set as follows: H/G boundary at 20 µg TP l^{-1}, G/M boundary at 35 µg TP l^{-1}, M/P boundary at 70 µg TP l^{-1}, and P/B boundary at 90 µg TP l^{-1}.

Quantitative macrophyte data should be used

We were able to demonstrate the major contribution of quantitative macrophyte data to the assessment of ecological status. The current assessment method should therefore be revised towards including and taking advantage of the frequency data sampled according to the Swedish monitoring standard (SEPA 2015). For example, the Norwegian Trophic index (Tic) was identified as a promising indicator for revising the assessment method for the quality element macrophytes. This indicator was developed for qualitative data (Mjelde 2013) but can also be adapted to accommodate quantitative data (Schaumburg et al. 2004a).

Macroalgae and bryophytes are demanding but have high additive value

Sampling and identifying macroalgae and bryophytes are time consuming and require high taxonomic competence on the part of the observers performing the field sampling. In WATERS, we demonstrated that these species groups have a significant additive value in ecological assessment. The extra resources needed in the field appear to pay off in terms of increased reliability and decreased uncertainty in the ecological assessment.

Remote-sensing assessment systems to be implemented at full scale

In WATERS, we demonstrated that high-resolution remote sensing has high potential in the ecological status assessment of humic lakes (Birk and Ecke 2014). The technical development of both sensors and platforms is progressing continuously. This increases the potential to detect and identify not only emerging, floating, and floating-leaved but also submerged vegetation. The technology and remote-sensing assessment prototypes are at a stage that encourages the implementation at full scale.

REMAINING CHALLENGES

Potential indicators identified

In WATERS, we compiled a macrophyte dataset comprising recent quantitative macrophyte data from all over Sweden, including from 91 lake reference sites. However, the dataset and potential indicators still have to be tested against typology groups. We identified several potential indicators that could be used in revising the Swedish assessment criteria, including TIc and ICM (Mjelde 2013; Kolada et al. 2014).

Species responses in rivers need to be evaluated

Macrophyte data from rivers are still rare. In the WATERS programme, we compiled a dataset of macrophyte data from a total of 89 rivers. These data need to be evaluated in terms of species responses along environmental pressure gradients and new indicators need to be developed. Also, a new monitoring standard needs to be developed, as the formerly used standard (SEPA 2003) increases the risk of missing important indicator species.

Responses to multiple pressures merit further evaluation

Most of the indicators evaluated in WATERS were tested against a phosphorus gradient. However, our remote-sensing study indicated that the tested macrophyte indicators responded more strongly to a multiple-pressure gradient than to the single phosphorus gradient (Birk and Ecke 2014). This indicates that macrophytes indeed display responses to multiple environmental pressures, a matter that needs to be evaluated in greater detail in a study designed to reveal single-pressure-responses.

Remote sensing-based assessment systems need to be validated

The identification of individual species, including emerging, floating, floating-leaved, and submerged growth forms, remains a challenge given the sensors currently suitable for UAS deployment. Recent developments in multi- and hyperspectral sensors suitable for UAS that were tested in WATERS are promising and could take the application of remote sensing-based assessment systems further. In remote sensing, the manual segmentation and classification of vegetation stands is time consuming (Husson et al. 2014a, 2014b). However, recent developments in software offer high potential to overcome the drawbacks of manual approaches. In a recent approach, we tested automatic segmentation and classification (Figure 3.8). The results are promising, but require further development and validation.

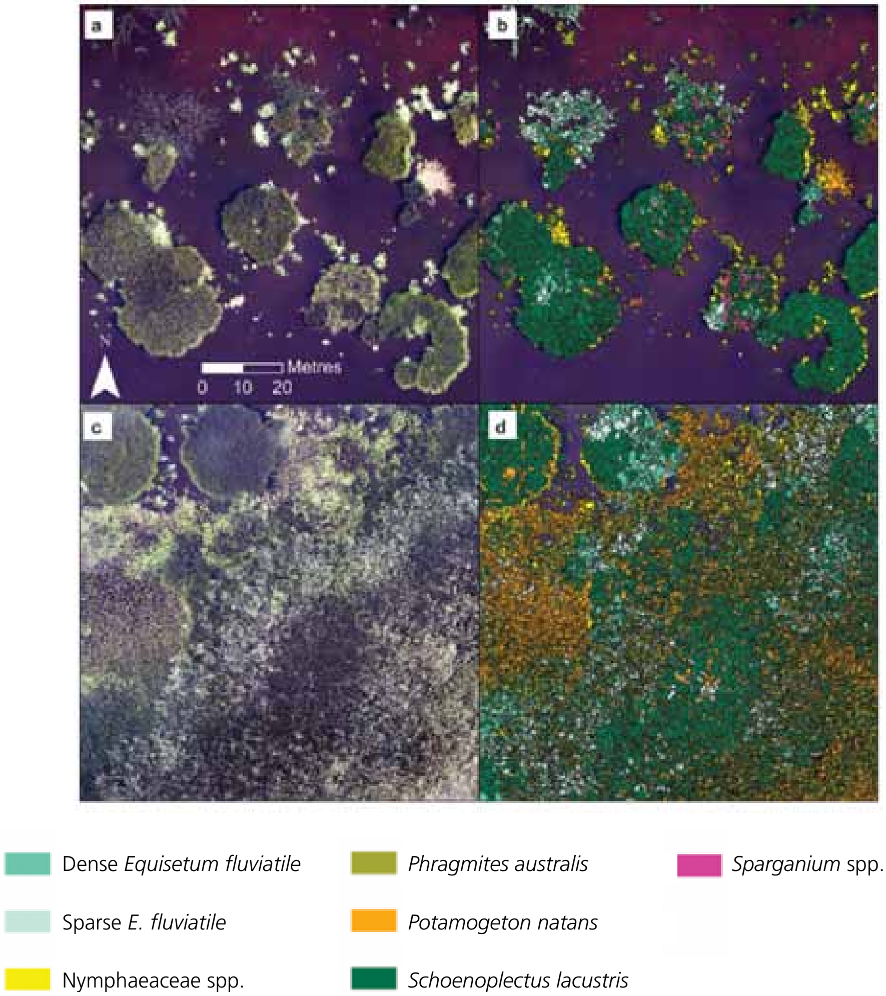


Figure 3.8: Illustration of the automated segmentation and classification (b, d) of vegetation in lake Ostträsket near Skellefteå, county of Västerbotten, based on high-resolution remote-sensing images (a, c): a and b show simple and mainly single-species vegetation stands, while c and d show complex vegetation stands.

3.5 Benthic invertebrates in streams and lakes

OVERVIEW OF CURRENT INDICATORS

Benthic invertebrates have a long history of use in environmental assessment

Knowledge that aquatic organisms reflect the quality of their habitats was documented as early as ancient Egyptian times (e.g. Time Magazine 1993; see Johnson 1993); the Greek philosopher Aristotle even taught his students that worms were born of river slime. The use of aquatic organisms to gauge anthropogenic impacts on rivers has a long tradition, with most studies advocating the use of macroinvertebrates in biomonitoring (Johnson et al. 1993; Hering et al. 2004; Bonada et al. 2006). By comparison, fewer studies have assessed the efficacy of using macroinvertebrate assemblages in biomonitoring lake ecosystems (Brauns et al. 2007; Johnson et al. 2004, 2007). Indeed, only a decade ago, the lack of WFD-compliant macroinvertebrate assessment tools was acknowledged as one of the major gaps impeding full assessment of the ecological quality of lakes (Solimini et al. 2006). Largely driven by WFD implementation, several biological indicators have recently been developed with which to assess the ecological quality of lakes (Brucet et al. 2013). Poikane et al. (2015) reviewed the indicators used by 10 European countries to detect ecological change. Their study found significant pressure-response relationships with acidification (three indicators), eutrophication (five indicators), morphological alterations (five indicators), and the combined pressure of eutrophication and altered morphology (two indicators).

Arguments for using benthic invertebrates in biomonitoring include sound theoretical concepts and predictive models of environment-species relationships, pollution-specific discrimination of human-generated impact, large-scale applicability, and the relatively low costs associated with sampling and taxonomy. In reviewing methods used across Europe to assess the ecological status of freshwater ecosystems, Birk et al. (2012) found that most countries continue to prioritize the use of benthic invertebrates in biomonitoring, using comparable field sampling and laboratory procedures (Friberg et al. 2006). Measures of invertebrate diversity, sensitivity, and ecological traits are commonly used to determine the ecological status of lakes and rivers, whereas measures of abundance are used less frequently (e.g. Johnson et al. 1993).

Benthic invertebrates are sensitive to pressures and therefore used in biomonitoring

Due to their high sensitivity and stress-specific responses to environmental pressures, benthic invertebrate assemblages are commonly used in biomonitoring to assess the ecological status of aquatic ecosystems (e.g. Johnson et al. 1993; Bonada et al. 2006). Indeed, the use of benthic invertebrates in assessing lotic systems has a long tradition, constituting the foundation of many biomonitoring programmes (Hellawell 1986; Rosenberg and Resh 1993). Most stress-response studies and indicator development efforts have focused on lotic systems, with substantially fewer studies addressing the efficacy of invertebrate assemblages for detecting anthropogenic effects on lake ecosystems. One exception is the use of profundal assemblages, in particular chironomid midges, which were used early on in biomonitoring to assess nutrient effects on lakes (Thienemann 1918; Wiederholm 1980). Increasingly, however, littoral invertebrate assemblages, principally comprising sedentary species and an important component of lake food webs, are being used as indicators of nutrient status (Donohue et al. 2009; McFarland et al. 2010), acidification (Stendera and Johnson 2008; Schartau et al. 2008; Johnson and Angeler 2010), and hydromorphological alterations (Brauns et al. 2007; Miler et al. 2015) in classifying the ecological status of lakes.

Five invertebrate indicators are currently used in streams

In this study, we assessed the response of lake and stream invertebrate assemblages to selected pressures. Currently, five indicators are used to assess the ecological status of lakes and streams using invertebrate assemblages (Table 3.4 and SEPA 2007). The Benthic Quality Index (BQI, Wiederholm 1980), comprising 12 profundal chironomid taxa, is used to assess the effects of eutrophication on lakes, while the DJ index (Dahl and Johnson 2004) is used for streams. To assess acidity, two multimetric indices are currently used: the Multimetric Index for Lake Acidity (MILA) for lakes and the Multimetric Index for Stream Acidity (MISA) for streams. Finally, to determine the effects of general degradation, the Average Score Per Taxon (ASPT)

Table 3.4. Indicators used in assessing the ecological quality of lakes and streams in Sweden and suggested changes.

System	Indicator	Acronym	Habitat, pressure	Suggested revisions and comments
Lake	Average Score Per Taxon[1]	ASPT	Littoral, general degradation	No revision recommended. Adjustments may be needed with revised estimates of reference condition.
Lake	Multimetric Index for Lake Acidity[2]	MILA	Littoral, acidity	Recommend revision. Exclusion of Leptophlebiidae and adjustment of threshold values for subindices. Further adjustments may be needed with revised estimates of reference condition.
Lake	Benthic Quality Index[3]	BQI	Profundal, eutrophication	No revision recommended. Adjustments may be needed with revised estimates of reference condition.
Stream	Average Score Per Taxon[1]	ASPT	Riffle, general degradation	No revision recommended. Adjustments may be needed with revised estimates of reference condition.
Stream	Multimetric Index for Stream Acidity[2]	MISA	Riffle, acidity	Recommend revision. Exclusion of Leptophlebiidae and adjustment of threshold values for subindices. Further adjustments may be needed with revised estimates of reference condition.
Stream	DJ index[4]	DJ index	Riffle, eutrophication	No revision recommended. Adjustments may be needed with revised estimates of reference condition.

1) Armitage et al. (1983), 2) Johnson and Goedkoop (2007), 3) Wiederholm (1980), 4) Dahl and Johnson (2004) index (Armitage et al. 1983) is used for both lakes (littoral habitats) and streams (riffle habitats).

During referral of the current classification system to regional authorities in 2008, much of the criticism focused on (i) the performance of the two acidity indicators MILA and MISA, (ii) the identification of reference conditions, and (iii) invertebrate responses to other environmental pressures such as hydromorphology, forestry, and urbanisation.

European intercalibration studies of benthic invertebrate methods used in lakes and wadeable streams have been completed in the last few years (van de Bund 2009; Poikane et al. 2016). Sweden was found to have compliant national methods for lake eutrophication (ASPT and BQI for littoral and profundal habitats, respectively), river eutrophication (DJ index), and lake and river acidification (MILA for lakes and MISA for rivers). Intercalibration work resulted in minor adjustment of some of the class boundaries (SwAM 2013). Intercalibration work on large rivers using ASPT, DJ index, and MISA is ongoing and will be finalised in June 2016.

OVERVIEW OF INDICATOR DEVELOPMENT

Five replicates per habitat identified to lowest taxonomic unit possible

For the calibration and validation of invertebrate indicators, we used quality-assured data collated by SLU. Field sampling was done using national methods. Lake habitats were sampled using standardised kick-sampling with a hand net (stony bottom, littoral habitats) or an Ekman sampler (profundal habitats) (prEN 16150:2010). Usually, five replicate samples were taken from each habitat using a sieve/hand net with a 0.5-mm mesh size. Stream samples were usually collected from riffle habitats using standardised kick-sampling with a hand net (prEN 16150:2010). The hand nets usually had a mesh size of 0.5 mm; however, samples were occasionally taken using a hand net with a mesh size of approximately 1.5 mm (M42 method). Most samples were collected in autumn and some in spring. Samples were preserved in the field using 70% ethanol and species were identified in the laboratory to the lowest taxonomic unit possible, usually species. Species lists were harmonised using a list of 517 operational taxonomic units (OTUs) (SwAM 2013). The lists of OTUs for lake (littoral) and stream (riffle) taxa are currently undergoing revision.

Pressure-response relationships were studied by isolating the pressure gradients of interest using the pressure filter described in Section 5.3. Systems affected by liming, unless otherwise noted, were not included in any of the analyses. In isolating pressure gradients we considered five types of pressures: eutrophication, acidification, forestry (clear-cut logging), urbanisation, and invasive species. For example, when studying relationships to elevated nutrients, we excluded sites affected by the other four pressures.

ASPT index responds to multiple pressures

The ASPT index is considered to provide an overall assessment of general degradation. Regression analysis partially supported this conjecture, with 27% (lakes) and 36% (streams) of the variation in ASPT explained by a gradient of non-natural land use (Supplementary information). Stronger relationships found between ASPT and total phosphorus concentration

(31% for lakes and 45% for streams) suggest that elevated nutrient levels and variables that covary with land use account for much of this variation. For example, Johnson et al. (in manuscript b), in a study of multiple-stressor effects on benthic invertebrate assemblages in lowland streams, found that disturbance of riparian zones explained much of the variability in macroinvertebrate taxonomic composition and species traits. Alterations of riparian vegetation can result in loss of biodiversity and ecosystem function due to changes in water temperature, incident light, and food webs, as well as increased loadings of fine sediment.

Revised indicators of acidity displayed improved performance

International agreements to curb emissions of acidifying compounds have been largely successful, with clear signs of chemical recovery in many Scandinavian surface waters (Stoddard et al. 1999). By contrast, biological recovery from acidification has lagged chemical recovery, and many ecosystems are still displaying signs of acidification (Johnson et al. 2007; Stendera and Johnson 2008; Ormerod and Durance 2009; Johnson and Angeler 2010). Community response to acidification is complex, reflecting the direct physiological effect of pH as well as the effects of associated toxic metals and indirect effects mediated through bottom-up processes (e.g. food availability; Johnson et al. 2007). For example, acidification often results in predictable changes in the balance between predator and prey organisms, with marked increases in certain large predatory taxa and decreases in acid-sensitive taxa (Økland and Økland 1986). Both stream riffle and lake littoral macroinvertebrate communities have been shown to be good indicators of acidification stress (e.g. Fjellheim and Raddum 1990; Johnson et al. 1993; Johnson et al. 2007).

Building on earlier studies using benthic invertebrates to assess acidification in Sweden (e.g. Henrikson and Medin 1986), Johnson and Goedkoop (2007) calibrated the MILA and MISA indices against pH. In WATERS, we focused on three factors that might improve the performance of these two indices: (i) exclusion of leptophlebid mayflies that are relatively tolerant of low pH, (ii) revised threshold values of the six subindices, and (iii) revised measures of reference conditions.

Figure 3.9: A leptophlebid mayfly by Jason Neuswanger, www.troutnut.com.

In both lakes and streams, Gradient Forest analyses of invertebrate response to pH indicated strong shifts in taxonomic composition when pH decreased below 6.0 (Supplementary information). It is well established that the community composition of stream and lake (littoral) invertebrate assemblages often changes at pH levels just below 6.0. For example, Johnson et al. (2007) found shifts in community composition (CA scores) between 5.5 and 6.0, with much lower among-site variability at pH < 5.5. Økland and Økland (1986) argued that gastropods were sensitive to changes in pH, noting a loss of species richness between pH 6.2 and 5.9 and no gastropods recorded at pH ≤ 5.2.

Revising MILA by excluding leptophlebid mayflies and adjusting the threshold values of the six subindices increased the precision. The revised MILA index performed better against pH for all three ecoregions. For example, in the boreonemoral and nemoral regions (Illies region 14), precision increased from an R^2 of 0.531 (RMSE 10.96) to 0.642 (15.08) after excluding leptophlebid mayflies and to 0.709 (11.75) after adjusting the threshold values of the subindices (Figure 3.10 and Supplementary information).

As in the case of MILA, excluding leptophlebid mayflies and adjusting the threshold values of the six subindices increased the precision of MISA (Supplementary information). Currently, the observed-to-expected ratios are typology based using a single value (reference MISA = 47.5) for all of Sweden. Model-based estimates of reference MISA values displayed increased precision compared with typology-based estimates, i.e. 0.23 for typology-based (Figure 3.11) versus 0.41 for model-based estimates. Random Forest modelling was done using variables characterising catchment land cover (e.g. % forest), stream size (e.g. depth and width), in-stream substratum, and water chemistry (e.g. water colour) for 168 acidified and reference streams sampled as part of the national stream survey in 2000. Robust predictor variables were % water in the catchment, water colour and conductivity, and catchment size (a proxy for stream size).

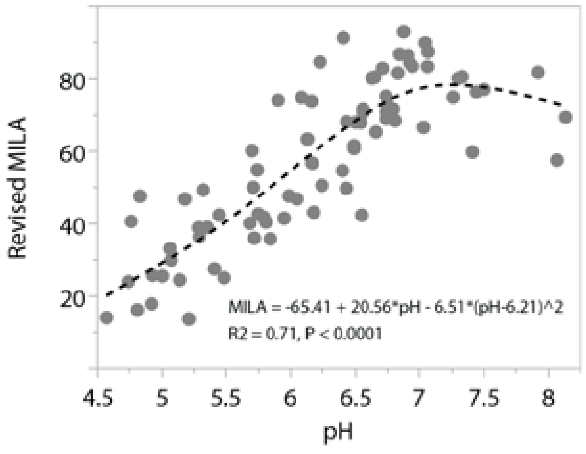


Figure 3.10: Regression of revised MILA against pH in 80 acidified and reference lakes in the boreonemoral and nemoral regions (Illies region 14). The regression line is a smooth fit with lambda set to 1.0.

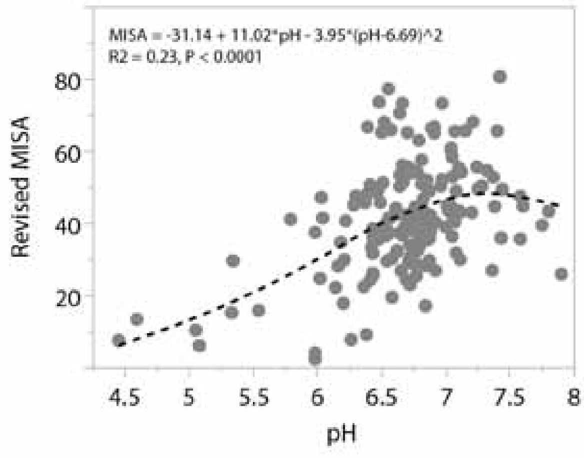


Figure 3.11: Regression of revised MISA index against pH in 168 acidified and reference streams sampled in the boreonemoral and southern boreal regions as part of the national stream survey in 2000. The regression line is a smooth fit with lambda set to 1.0.

Eutrophication of lakes and streams continues to cause biological impairment

Elevated nutrients from agricultural land use often result in the increased population growth of phytoplankton and the depletion of profundal oxygen concentration in many lakes. Profundal invertebrate assemblages are widely used to assess the ecological status of lakes. Building on early classification schemes developed by Thieneman (1922) using chironomid midges, works by Brinkhurst (1974) and Wiederholm (1980) paved the way for more contemporary approaches (Gestmeier 1989; Johnson and Wiederholm 1989; Kansanen et al. 1990; Jyväsjärvi et al. 2010, 2012). The Biological Quality Index (BQI), based on the relative abundances of 12 chironomid taxa (Wiederholm 1980), is still one of the most commonly used methods in Europe.

Analysis of the ratio of profundal chironomid assemblages to total phosphorus concentration confirmed the response of BQI indicator taxa to nutrient enrichment (Supplementary information). The most marked changes in taxonomic composition occurred <10 µg TP L^{-1}, which is near the upper limit of oligotrophic conditions. We also noted shifts in midge assemblages between 40 and 60 µg TP L^{-1}, likely reflecting the transition from mesotrophic to eutrophic conditions. Taxa associated with nutrient-poor conditions, i.e. *Paratanytarsus, Protanypus,* and *Heterotrissocladius subpilosus,* decreased in prevalence at TP concentrations between 20 and 40 µg TP L^{-1} and were replaced by taxa indicative of eutrophic conditions (e.g. *Chironomus plumosus, Endochironomus,* and *Cryptochironomus*).

Several nutrient-poor lakes had relatively low BQI values, indicating the importance of other environmental drivers. Jyväsjärvi et al. (2012) found poor discrimination of impact in shallow boreal lakes using profundal invertebrate assemblages, arguing that even pristine shallow, humic lakes are often dominated by species indicative of eutrophic conditions (e.g. *Chironomus plumosus and Chaoborus flavicans*). These authors also found that models using mean depth and the mean/maximum depth ratio partitioned natural variability better than did a lake-typology approach (Jyväsjärvi et al. 2012). Moreover, in intercalibration work on Finnish lakes, relationships were stronger when shallow lakes (mean depth < 6 m) were not included in the analyses (R^2 of 0.26–0.32 for BQI and TP relationships) (Poikane et al. 2015). In our analyses, modelling natural variability using linear regression (minimum Bayesian information criterion) and Random Forest analyses did not improve the performance of BQI (Supplementary information).

As in the case of lakes, agricultural land use results in increased growth of primary producers such as benthic algae and macrophytes in streams. Moreover, removing riparian vegetation often results in the destabilisation of stream banks and increased erosion. Changes in basal resources, loss of habitat heterogeneity, and altered hydrology are three of the most pervasive pressures affecting the biodiversity and functioning of streams (e.g. Lange et al. 2014). The DJ index, comprising five subindices representing species richness (i.e. number of Ephemeroptera, Plecoptera, and Trichoptera taxa – EPT), taxonomic composition (i.e. relative abundance of Crustacea and EPT taxa), and species tolerance (ASPT index, Armitage et al. 1983, and Saprobien index, Zelinka and Marvan 1961), is used to quantify the effects of agricultural land use on streams in Sweden (Dahl and Johnson 2004). Although calibrated against an organic pollution gradient, the indicator has been demonstrated to be significantly correlated with many of the pressures commonly associated with agricultural land use, such as altered habitat (silt substratum) and catchment-level predictors (e.g. arable and urban land cover) (Dahl and Johnson 2004).

The total phosphorus concentration accounted for 43% of the variability in the DJ index (Supplementary information). Random Forest modelling revealed that altitude and conductivity were important predictors of the DJ index. Models were better than typology in establishing reference communities. The model-based observed-to-expected ratios regressed against total phosphorus concentration accounted for 22% of the variability, compared with 12% for typology-based observed-to-expected ratios (Supplementary information). In European intercalibration work, the DJ index was found to be strongly (linearly) related to the intercalibration metric (ICM) developed for the Nordic geographic information group (NGIG) (van de Bund 2009). Classification boundaries for Swedish water bodies using the DJ index were generally more stringent than those of other Member States.

SUGGESTED REVISIONS OF ASSESSMENT METHODS

We suggest changing some indicators and adjusting class boundaries

Insights gained through WATERS have resulted in a number of recommendations that should improve the performance of benthic invertebrate indicators in detecting anthropogenic change and classifying ecological status. We recommend revising the two acidity indicators MILA and MISA and revising the list of operational taxonomic units to be used in biomonitoring and in the calculation of indicators. No changes are recommended for ASPT (both lakes and streams), BQI (lakes), or the DJ index (streams), although class boundaries may need to be adjusted with the use of revised typology- or model-based estimates of reference conditions.

Improved acidity indicators by excluding mayflies

The performance of the acidity indicators MILA and MISA improved when leptophlebid mayflies were excluded and after adjusting the threshold values of the subindices. The precision of MILA when regressed against pH increased from 0.531 to 0.709, while the precision of MISA increased from 0.233 to 0.324. Analyses of the percentile distributions of the MILA and MISA subindices using the newly collated datasets indicated marked changes in percentile distributions (Table 3.5). For lakes, the 10th and 90th percentiles of % Diptera and % predators changed markedly, while for streams the number of families and % shredders differed substantially.

Table 3.5. Revised threshold values for standardising subindices to values between 0 and 10 (current values in parentheses).

A. MILA	Percentiles	
	90th	**10th**
% Ephemeroptera of abundance*	32 (27)	0 (0.05)
% Diptera of abundance	53 (86)	0 (26)
Number of Gastropoda taxa	3 (8)	0 (0)
Number of Ephemeroptera taxa*	5 (6)	0 (1)
$AWIC_{family}$ index*	5.56 (5.4)	4.0 (4.8)
% Predators of abundance	61 (19)	5.9 (8.7)

B. MISA Index	Percentiles	
	90th	**10th**
Number of Families*	29.5 (43)	11 (21)
Number of Gastropoda taxa	8.4 (3)	0 (0)
Number of Ephemeroptera taxa*	7 (16)	0 (3)
Ephemeroptera/Plecoptera [% abundance]*	11.7 (10)	0.16 (0)
$AWIC_{family}$ index*	4.8 (5.4)	3.8 (4.8)
% Shredders	2.4 (14)	25 (1.4)

** Revised subindices including Ephemeroptera are calculated without Leptophlebiidae.*

Models versus typology

A rigorous study of typology- versus modelled-based approaches to estimating reference conditions was not possible for all indicators because data were insufficient to characterise habitat features. However, preliminary analyses indicate that model-based estimates of reference conditions using continuous predictor variables result in higher precision than do typology-based approaches. For lakes, models predicting the probability of taxon occurrence (RIVPACS-type models, River Invertebrate Prediction and Classification System) and index values (Random Forest models) displayed higher precision than did typologies (based on ecoregion, altitude, depth, alkalinity, and colour) or a null model. For example, correlations between observed and predicted MILA index values ranged from 0.89 for Random Forest and 0.78 for linear regression to 0.29 for a recently revised typology (Drakare 2014) and 0.07 for the current (ecoregion-based) typology. Likewise, for streams, Random Forest models resulted in higher precision. These results are encouraging, so we recommend further effort to assess the efficacy of model-based approaches to establishing reference communities.

List of operative taxonomic units in need of revision

An expert working group has recommended revising the list of operative taxonomic units (OTUs) first established in 1995 as part of the national lake and stream survey, and currently used for classifying lakes and streams. In revising the list of OTUs, the working group considered (i) the importance of taxonomic identification for biomonitoring and index calculation, (ii) biogeographic distribution (OTUs should work across the whole country) and (iii) nomenclature (updated according to the Swedish Taxonomic Database, DYNTAXA) (Lars Eriksson personal communication, 14 March 2016).

The use of DNA barcoding in identifying organisms is increasing, in particular for groups that are difficult to identify using morphology (e.g. chironomid midges). Therefore, Eriksson et al. (in prep) recommended that samples stored for extended periods of time (e.g. in reference collections) should be preserved in 95% ethanol to allow for future analyses using molecular methods.

REMAINING CHALLENGES

Approach to establishing reference conditions needs to be determined

A number of challenges remain in decreasing the uncertainty associated with using benthic invertebrate assemblages in quantifying the ecological status of lakes and streams. A pivotal component of most ecological assessments is establishing the human-induced deviation from the target/reference condition given known levels of uncertainty. Uncertainty can be decomposed into error associated with quantifying the normal (natural) variability of the water body type, the accuracy of the indicator used to gauge deviation, and error associated with collecting and processing samples. Although progress has been made in evaluating type- and site-specific approaches to establishing reference conditions, more comprehensive study of typology- and model-based approaches is needed. Specifically, many of the variables used in typology (e.g. lake depth and stream slope) were not available for many of the sites when validating current assessment methods. Once the approaches to establishing reference conditions are determined, Pressure-response relationships need to be evaluated, including measures of uncertainty, to determine whether classification boundaries need to be adjusted

Number of taxa for lake BQI could be increased, as in the Finnish revision

At this time we do not recommend changing the lake (profundal) BQI. In Finland, however, the BQI originally based solely on chironomid midges has been replaced by a BQI employing the same algorithm but using benthic invertebrates commonly found in profundal habitats (Jyväsjärvi et al. 2014). Two statistical methods (detrended correspondence analysis and weighted averaging) were used to estimate revised scores for profundal assemblages. This approach resulted in an increase in the number of indicator taxa from seven (current Finnish BQI) to 70 (revised Finnish BQI). The reasoning behind increasing the number of indicator taxa was that many of the BQI scoring midges were rare or absent, resulting in many zero values and making it impossible to calculate the BQI for many water bodies. For Swedish waters, not finding indicator taxa does not seem to be a major issue limiting its use. However, a BQI should be calibrated following the approach outlined by Jyväsjärvi et al. (2014) and the result compared with the current BQI. Along these lines, development of a multimetric index (MMI) for profundal habitats might be attempted.

We recommend that OTUs be used in classifying the profundal habitats of lakes

In proposing a standardised species list for profundal assemblages, the minimum would be the 12 chironomid species that comprise the BQI. In addition, other taxa commonly found in profundal habitats (e.g. oligochaetes and selected crustaceans) should be included to estimate changes in taxonomic richness. Use of molecular methods to identify chironomids and other difficult-to-identify taxonomic groups is increasing, and these methods will likely be an important tool for identifying taxonomically challenging groups. Given the importance of invertebrates in subarctic and arctic food webs, the use of DNA methods should be considered in future studies.

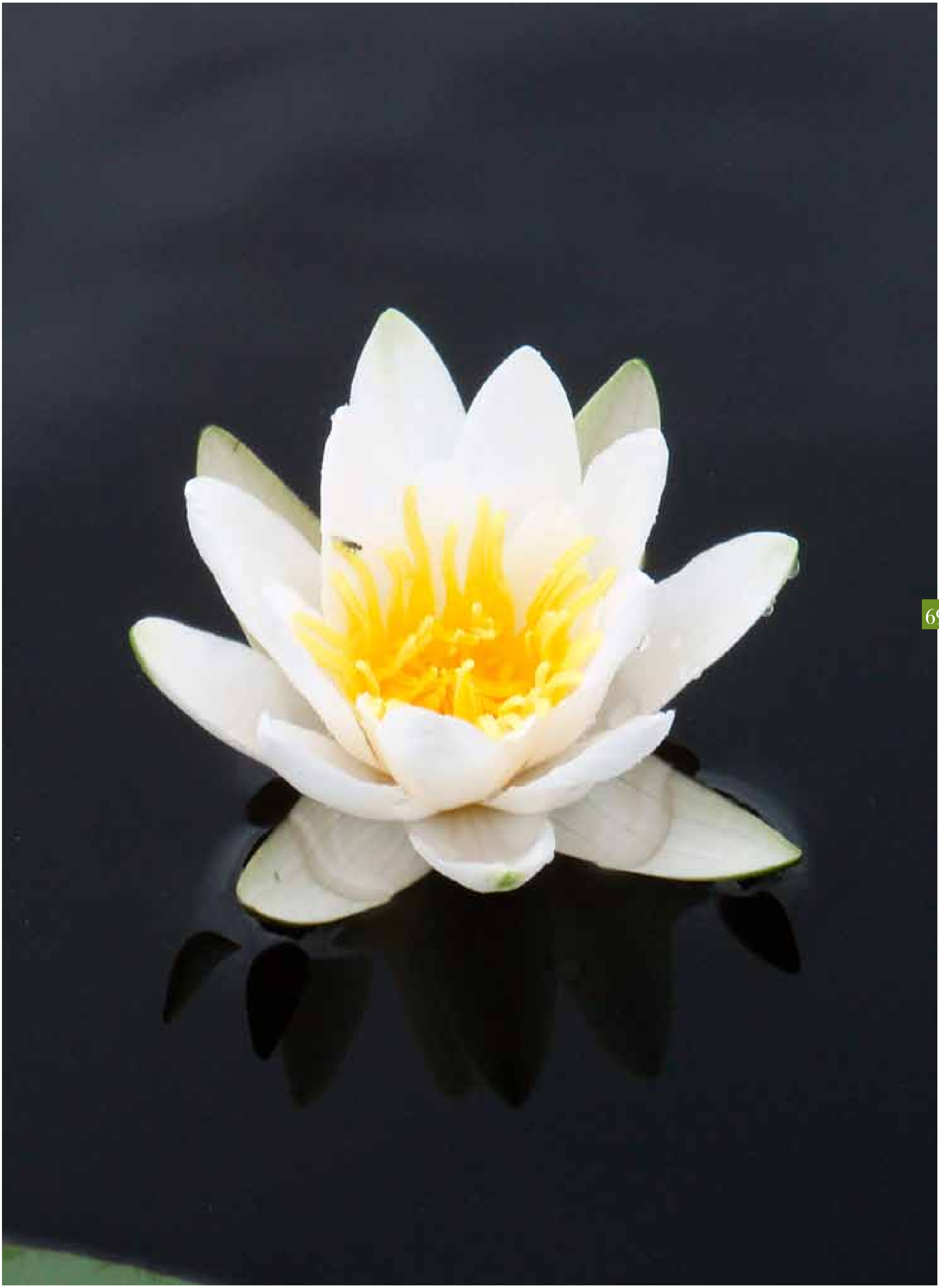
Benthic invertebrate samples are preferably collected from riffle habitats. Sampling efforts should be stratified to the hard-bottom substratum to reduce habitat-specific variability and increase statistical power. As diversity in general and the presence of many sensitive taxa specifically (e.g. the Ephemeroptera, Plecoptera, Trichoptera, and EPT taxa) are often higher in riffle habitats than in soft-bottom sediments, statistical power (i.e. the power to detect change) is also higher, resulting in low false-negative errors. However, restricting sampling to hard-bottom habitats makes finding appropriate sampling sites difficult in low-energy lowland streams. A Swedish study (part of the EU-STAR project) found that riffle habitat sampling resulted in higher measures of diversity than did multi-habitat sampling (STAR-AQEM method) (Sandin et al. 2006). However, as stratified sampling of riffle habitats is considered problematic in lowland streams, we recommend comparing the STAR-AQEM (e.g. prEN 16150:2010) and riffle methods (SS-EN 27 828) currently used in lowland streams.

Effects of forestry, altered hydrogeomorphology, and urbanisation need to be studied

Insights from WATERS have indicated that benthic invertebrate assemblages are affected by pressures not considered in current ecological classifications, pressures such as forestry, alterations of hydrogeomorphology, and urbanisation (Johnson et al., in manuscript a), as well as multiple pressures (Johnson et al., in manuscript b). Future work should focus on developing tools for quantifying the extent of single- and multiple-pressure effects on lake and stream invertebrate assemblages. For example, targeted sampling of streams affected by alterations of hydrogeomorpology and forestry, combined with data collected as part of the WATERS gradient studies, could be used in calibrating indicators to assess these two pressures.

Assessments need to be harmonised at regional and national levels

Finally, to better harmonize assessments among regional and national authorities, user-friendly software should be developed for calculating indicators and reference values and for classifying sites, taking into account the uncertainty associated with the method. For example, after reporting selected physicochemical variables, software can be used to predict the site-specific reference community using either typology or models. In the first step, a measure of taxonomic completeness is used to determine whether the observed (i.e. measured) community deviates from what is expected under unperturbed (i.e. reference) conditions. Biological indices are then calculated as diagnostic tools for determining what pressures might be causing the deviation. Indices are calculated using modelled data of the reference community and measured data from the water body. Ideally, knowledge of potential pressures is used when comparing indicator responses. Classification of taxonomic completeness (i.e. number of observed/expected taxa and O/E values) is done including measures of uncertainty, resulting in classification probabilities across the five ecological status bands.



3.6 Fish in streams and lakes

OVERVIEW OF CURRENT INDICATORS

Fish communities currently described by multimetric indices

Current Swedish fish assessment methods consist of multimetric indices (Table 3.6), each including several indicators describing the structure and function of fish communities. The indicators constituting the main indices EQR8 and VIX cover one or more aspects of abundance, species composition, and age structure as required by the WFD. All indicators included in each index were previously known to be affected by acidification, eutrophication, and/or other human pressures.

Table 3.6. Fish indices used or suggested for use in assessing the ecological quality of lakes and streams in Sweden.

System	Indicator	Acronym	Habitat, pressure	Suggested revisions and comments
Lake	Multimetric fish index for lakes (eight subindices)[1]	EQR8	General degradation, acidity, and nutrients	Consider revision depending on the results of the project on common fish indices for Swedish and Norwegian lakes.
Stream	Multimetric fish index for streams (six subindices)[2]	VIX	General degradation and nutrients/ organic load	No revision recommended.
Stream	Hydrology side index to VIX (five subindices)[2]	VIXh	Hydrology impact	No revision recommended.
Stream	Acidity or morphological side index to VIX (four subindices)[2]	VIXsm	Acidity or morphological impact	No revision recommended.
Stream	Morphological side index to VIX (seven subindices)[3]	VIXmorf	Morphological alteration	New index suggested to complement the current one.

1) Holmgren et al. (2007), 2) Beier et al. (2007), 3) Spjut and Degerman (2015b).

Currently, EQR8 and VIX are applied to small lakes and streams. Monitoring data from fairly small lakes and streams were used in developing the current multimetric fish indices EQR8 for lakes (Holmgren et al. 2007) and VIX for streams (Beier et al. 2007). The fish indicators include aspects of abundance, species composition, and age structure, but the specific indicators and measurement units differ between EQR8 and VIX (Figure 3.12). Site-specific reference values were modelled using multiple regressions to account for continuous variability in water body size and other typology variables at the least-disturbed sites. The standard sampling of whole lakes using Nordic multi-mesh gillnets (CEN 2015; Figure 3.13) is not generally feasible and/or is too expensive in very large and deep lakes. The monitoring data available in the National Register of Survey Test Fishing (NORS) and the Swedish Electrofishing Register (SERS) were specifically used to develop EQR8 and VIX, respectively. In Sweden, the electrofishing standard (CEN 2003) has usually been implemented by wading (Figure 3.13), covering the whole wetted width of a defined stream reach. Such sampling is restricted to shallow stream reaches with predominantly hard substrates. Very large lakes and large and/or slowly running rivers are therefore absent from the calibration datasets. This absence reflects the strong association between the assessment methods used and the habitats sampled in the main current monitoring programmes.

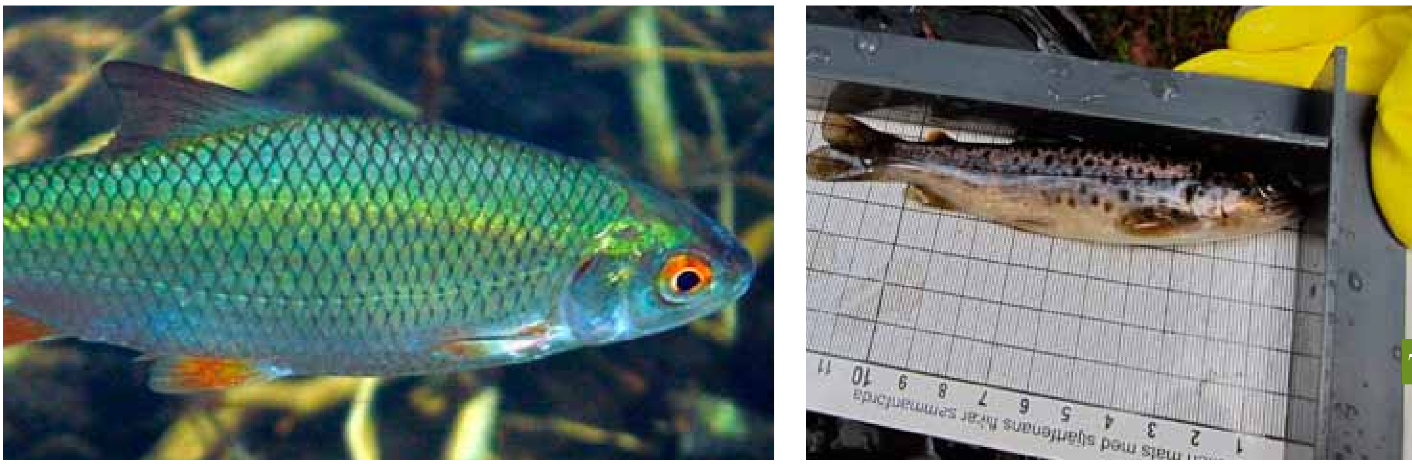


Figure 3.12: Roach (Rutilus rutilus) and brown trout (Salmo trutta) occur in lakes and streams. Roach and other cyprinids are key species in EQR8. Brown trout and other salmonids are central to VIX. *Photos: left, Anders Asp; right, Lars Ohlson).*

Figure 3.13: Two photos of standard fish sampling in a small lake (*left: Magnus Dahlberg)* and in a wadeable stream *(right: Lars Ohlson)*.

OVERVIEW OF INDICATOR DEVELOPMENT

VIX is inter-calibrated but EQR8 is not

When WATERS started in 2011, both VIX and EQR8 were part of ongoing European intercalibration projects (EC 2011). VIX passed the intercalibration (EC 2013), along with eleven other national river fish assessment methods from other Member States. In contrast to VIX, EQR8 did not pass the intercalibration in the Nordic lake fish group (Olin et al. 2014). In this case, four national fish indices were applied to Nordic gillnet catches in lakes from all countries. Index responses were tested against a gradient of total phosphorus, i.e. representing eutrophication pressure. EQR8 decreased slowly with increasing total phosphorus in Swedish lakes. There was, however, no significant response to this pressure when EQR8 was applied to lakes from the other countries, indicating problems with reference values estimated using indicator-specific models calibrated for Swedish lakes.

Different sampling methods complement each other

Over the decade since the development of EQR8 and VIX, alternative sampling methods (e.g. boat electrofishing, vertical and horizontal hydroacoustics, trawling, Nordic coastal multi-mesh gillnetting, and Norden multi-mesh stream survey netting) have been used in larger water bodies in Sweden and elsewhere (reviewed by Holmgren 2016a; Figure 3.14). Method comparisons confirm that different sampling methods are complementary rather than perfectly related and exchangeable (e.g. CEN 2006). Mobile hydroacoustics was recently

Figure 3.14: Coastal multi-mesh gillnets are used in the largest lakes (left) and point electrofishing from boats (right) can be used as a complement in shallow and vegetated areas. *Photos: Alfred Sandström*.

adopted as a European standard for estimating fish abundance (CEN 2014), but additional methods are needed to monitor species composition and age structure. The development of Swedish methods for status assessment in larger water bodies lags the increased monitoring of fish using alternative methods. Within WATERS we initially aimed at improving the current site-specific multimetric fish indices, for example, by estimating various sources of uncertainty, to develop indicators better adapted to larger lakes and streams. We also intended to explore and describe complementary and/or alternative sampling and analytical methods, to provide guidance for status assessment in large or otherwise distinctive lakes and streams.

Fish data from whole-lake sampling used

In WATERS we used fish data from NORS and SERS, i.e. whole-lake sampling with Nordic multi-mesh gillnets and electrofishing by means of upstream wading in defined stream reaches. Thousands of lakes and stream sites have been sampled with these standard methods. Considerably smaller subsets of the total fish datasets were used in specific analyses. Uncertainty analyses focused on specific time periods and/or on water bodies sampled repeatedly at multiple sites. Datasets for ongoing and completed studies of indicators and index responses to pressures were always limited by incomplete data on relevant pressures for most sites for which there are fish data. In contrast to the general use of only one standard sampling method in small lakes and streams, one recent study combined available data from gillnetting and/or hydroacoustics for 21 of the 67 water bodies defined in the four largest lakes of Sweden (Sandström et al. 2016).

A library of variance components for fish was created

An important part of WATERS was to create a library of variance components for all biological indicators used in Swedish waters (Bergström and Lindegarth 2016), including the fish indices EQR8 and VIX. With 772 whole-lake estimates of EQR8 and its eight subindices, we accounted for variability among samples within lakes and among years within a six-year assessment period. We found that the uncertainty of EQR8 and its subindices is slightly reduced by sampling in two or more years within a six-year assessment period. Another dataset included 2330 estimates of VIX and its side indices VIXsm and VIXh, estimated in multiple years at two or more electrofishing sites per stream. Here it was possible to estimate variance among electrofishing sites within streams and the interaction between years and sites. We found that the precision of VIX and its side indices increases more by increasing the number of sites sampled in a stream than by sampling more years in the six-year assessment period. Another study used data from 45 lakes and focused on uncertainty in two indicators of fish abundance (Balsby and Holmgren, in manuscript), measured as biomass and number of fish in individual gillnets. For both indicators, the variation between years was much smaller than the variation among gillnets within a lake. Therefore, assessing the status of a six-year period by sampling only one year, appears to be largely as precise as sampling two years if the number of nets sampled during that single year is doubled.

Common fish indicators for Swedish and Norwegian lakes

EQR8 responds more strongly to acidity than to high nutrient levels (Holmgren et al. 2007), but acidity as a pressure was not considered in the previous intercalibration (Olin et al. 2014). The current Norwegian Fish Index (FCI) also responds to acidity, and it failed (as EQR8) to be inter-calibrated with the Finnish and Irish fish indices for eutrophication. FCI relies on knowledge of temporal change in local species occurrence and dominance relationships, rather than deviance from reference values derived from recent samples at least-impacted sites. EQR8 and FCI are conceptually different, and the indices were not correlated when applied to both Swedish and Norwegian lakes within the same catchment (Schartau et al. 2012). A project was therefore initiated to find common indicators for assessing lakes on both sides of the Swedish – Norwegian border (Holmgren 2016b). Within WATERS, we acquired physical and chemical data for use in a revised pressure filter, i.e. to split biological datasets between minimally disturbed (i.e. reference) and stressed sites, respectively (Johnson et al. 2014). Out of 598 non-limed lakes in the Swedish fish dataset, only 21 reference lakes and 290 stressed lakes were identified. The Swedish – Norwegian project decided to use a less restrictive reference filter developed at the European level (Caussé et al. 2011). We first identified a preliminary set of 170 Norwegian and Swedish reference lakes, along with 332 lakes subject to one or more known pressures. The plan is to test fish responses to gradients of 1) pH and 2) total phosphorus and % agricultural land use. The responses will be expressed as deviations from site-specific reference values, as in the previous Swedish and European fish indices (Holmgren et al. 2007; Argillier et al. 2013).

Promising fish indicators for assessment in large lakes

In contrast to whole-lake gillnet sampling in smaller lakes, sampling with benthic gillnets has been performed in sub-areas of large Swedish lakes, including Sweden's four largest lakes Vänern, Vättern, Mälaren, and Hjälmaren. Pelagic areas of these lakes have also been sampled with hydroacoustics and trawling at multiple sites (Sandström et al. 2014; Figures 3.15 and 3.16). The four largest lakes are divided into a total of 67 water bodies, with recent fish data representing 21 of them (Sandström et al. 2016). Seventeen fish indices were calculated from catches obtained with either Nordic coastal multi-mesh gillnets in the benthic habitat (data from 19 water bodies) or hydroacoustics in the pelagic habitat (data from nine water bodies). Eutrophication pressure was indicated by chlorophyll-*a* (derived by remote sensing), total phosphorus (derived from water samples), and % agriculture in the catchment. Density of pelagic fishes, benthic biomass of planktivorous species, and benthic proportion of cyprinids (excluding roach) responded most strongly to eutrophication pressure (Sandström et al. 2016). All of them increased significantly with each of the three measures of eutrophication. The two benthic fish indices were most strongly correlated with chlorophyll-*a*, possibly reflecting the higher resolution in time and space of remote sensing versus fewer measurements of total phosphorus made within each water body. After modelling relationships between fish samples and remote-sensing data, the latter can be used to map the spatial distribution of fish density in large lakes (Figure 3.17). Such predicted distribution maps may be used to optimize future monitoring programmes.

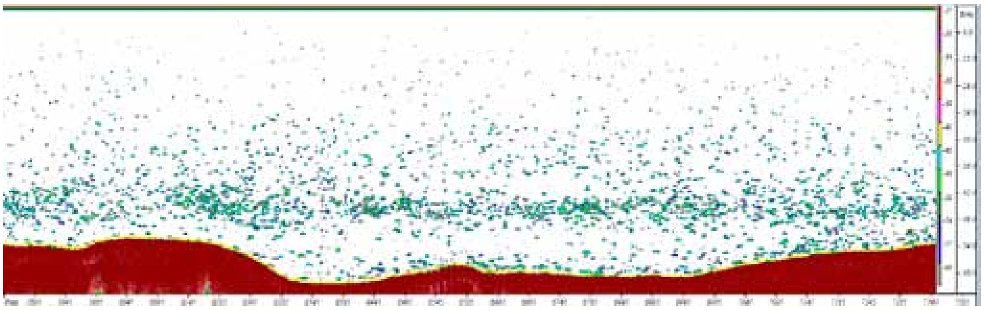


Figure 3.15: Echogram of a transect from vertical hydroacoustics, indicating that most fish reside in the deep and cold hypolimnion at night. *The figure was provided by Thomas Axenrot.*

Figure 3.16: Smelt (Osmerus eperlanus, le*ft; photo: Eva Kyhlberg*) and vendace (Coregonus albula, *right; photo: Alfred Sandström*) are the most abundant fish species in the deep pelagic areas of the largest lakes.

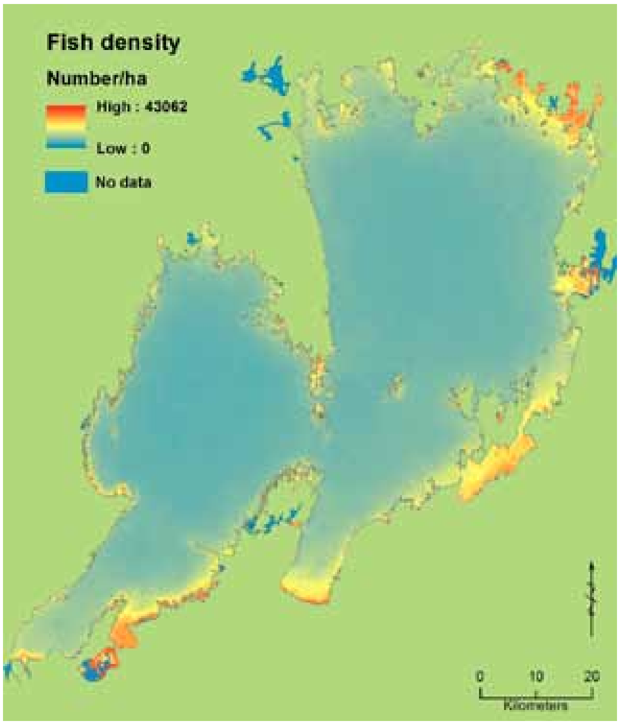


Figure 3.17: Predicted distribution of the total density of pelagic fishes in Lake Vänern, using a linear regression model with one of the remote-sensing variables (coloured dissolved organic matter) as the main predictor.

Pressure-specific fish indices complement VIX

As previously mentioned, the stream fish index VIX was initially complemented with subindices responding more specifically to certain pressures (Beier et al. 2007), i.e. acidity/morphological pressure (VIXsm) and hydrological pressure (VIXh). At that time there were, however, just a few sites in the dataset representing high hydrological pressure. A regional study found that VIXh responded reasonably well to water level regulation at the local scale, but a new index called RIX had higher precision than VIXh in detecting sites affected by water level regulation (Degerman et al. 2013). In RIX, reference values are modelled for the density of salmonids and fish species favoured by regulation, respectively, and the observed residuals were combined to form the RIX index. Altered hydrological regime assessed at sub-catchment scale was, however, a poor predictor of direct hydrological pressure at the local level of electrofishing sites (Spjut and Degerman 2015a), preventing us from refining a national fish index responding specifically to hydrological pressure. A refined fish index of morphological pressure (VIX_{MORF}) was instead developed using data from river habitat surveys in three counties (Spjut and Degerman 2015b). Similar to RIX, the improved VIX_{MORF} was based on fish species favoured or disfavoured by morphological pressure at the local scale.

SUGGESTED REVISIONS OF ASSESSMENT METHODS

New estimates of uncertainties for fish indices should be included in updated guidance

In the present state, we have no obvious reasons to completely replace EQR8 or VIX as an official Swedish assessment method. The upcoming calibration of reference values for Swedish and Norwegian lakes might facilitate fish indicators or a new fish index that can be accepted as intercalibrated for use in smaller lakes of neighbouring countries. VIX is already accepted as an intercalibrated method, which might be more successfully used in combination with refined indices addressing specific pressures. Similar to VIX, EQR8 (and/or some Swedish – Norwegian alternatives) will still need complementary methods to address certain pressures or to be used in certain lake types. Advice on how and when to apply alternative or complementary fish assessment methods might be included in an updated official Swedish guidance on assessment of ecological status.

For fish in lakes, the national host of the NORS database (http://www.slu.se/sjoprovfiskedatabasen) has provided an Excel application called "EQR8 beräkning.xlsx". It can be used to test the effect on EQR8 of manual changes in estimated reference values of included fish subindices or the sensitivity to adjustment of the environmental factors used to estimate site-specific reference values. This tool was developed at the request of end users, for example, to handle cases when reliable information indicates very low historical fish species richness. Another potentially useful piece of advice is to use the inter-calibrated Finnish EQR4 (Olin et al. 2013), for example, for non-acidified Swedish lakes in the Torne river basin (Sairanen et al. 2008).

For fish in wadeable streams, advice on how to use the new index VIXMORF might easily be included in an updated guidance. Just as for the previous subindices VIXsm and VIXh, the observed cut-off value for best distinguishing between reference and impacted sites may be the boundary between good and moderate status (i.e. VIX_{MORF} = 0.467, Spjut and Degerman 2015b).

The updated guidance should include new estimates of uncertainty for EQR8, VIX, VIX-sm, and VIXh. Uncertainty estimates of VIX-MORF might be estimated in a similar way as for VIX. The general key messages of the uncertainty studies should also be emphasised, i.e. to optimize the numbers of gillnets used in lakes and the number of stream sites per water body when conducting assessments for the relevant six-year periods in the water management plans.

REMAINING CHALLENGES

Boundaries between good and moderate status urgently needed

For lakes small enough to be sampled for whole-lake estimates using multi-mesh gillnets, results from the ongoing Swedish – Norwegian project might suggest revision of the current assessment method. Based on recent analyses at the European scale, we expect abundance indices to increase with eutrophication, but responses of diversity and size-based metrics may be insignificant or weaker (Argillier et al. 2013; Brucet et al. 2013; Emmrich et al. 2014; Arranz et al. 2015; Mehner et al. 2016). Based on the Swedish experience, we expect abundance and some other metrics to respond in opposite directions to decreasing pH and increasing total phosphorus, respectively (Holmgren et al. 2007). We will further explore age structure responses, as previously calculated for Swedish lakes (Holmgren 2013).

The finding of significant fish responses to eutrophication gradients in large lakes (Sandström et al. 2016) is an important first step in indicator development for those systems. This suggests that several fish indicators from different habitats, derived using different sampling methods, might be used separately or combined in the integrated assessment of water bodies representing sub-areas of large lakes. An interesting question is how similarly collected data from other water bodies in other large lakes in southern Sweden would fit the observed relationships between fish indicators and pressures. Another, more challenging issue is how to define the reference conditions needed to calculate ecological quality ratios (EQRs) and to set class boundaries. To be useful tools in water management plans, boundaries between good and moderate ecological status are most urgently needed.

Regular monitoring data are needed for indicator development

VIX has preliminarily been applied to 169 shallow electrofishing sites in 14 of 17 large Swedish rivers (catchment >10,000 km^2), and the geographical pattern of assessed status reasonably reflects the known extent of high pressures from water power plants and channelisation (Erik Degerman, unpublished results). In this case, the fish indicators representing hard-bottom habitats in large rivers seemed to be useful for status assessment. For more complete assessment of deep and slow-running river reaches, lack of regular monitoring data is at present a factor limiting indicator development. We suggest increased use of boat electrofishing and/or Norden multi-mesh stream survey nets, in addition to traditional electrofishing by wading along the shore. Data using such methods need to be collected from reference as well as impacted sites, in order to define reference conditions of fish indicators responding to certain pressures. It is also important to collect relevant pressure data representing the monitored sites. The available fish indices VIX, RIX, and VIX_{MORF} were developed for use at wadeable electrofishing sites, but similar concepts might be used for river habitats sampled using other methods. RIX and VIX_{MORF} can be more rigorously tested with relevant hydrological and morphological data for more stream water bodies, for example, collected with river habitat surveys (SEPA 2003).

3.7 Integrated assessment and responses in inland waters

Precision and sensitivity of different BQEs to pressures evaluated in WATERS gradient studies

As mandated by the WFD, multiple BQEs are to be used in monitoring programmes. Proponents of using multiple taxonomic groups and approaches argue that a multiple-lines-of-evidence approach strengthens inference-based models, resulting in lower assessment error (e.g. Stevenson et al. 2004). Moreover, the use of multiple indicators may help to distinguish the effects of multiple stresses. Simply put, different taxonomic groups may respond similarly to some stressors (high redundancy, strengthening inference of change) but differently to other stressors (low redundancy, helping elucidate different stressor effects). However, responses in one BQE arguably might indicate biological conditions for other BQEs, so increasing sampling efforts may result in unwarranted cost increases. Few studies have, however, compared the discriminatory power of different taxonomic groups to detect change (e.g. Johnson et al. 2006; Johnson and Hering 2009).

As monitoring efforts have moved from single sites to monitoring whole catchments (European Commission 2000; Hering et al. 2010), consideration should also be given not only to the taxonomic groups or microhabitats within systems best correlated with disturbance but also to the systems (e.g. streams and lakes) that provide the most precise signal of activities occurring in the catchment (e.g. Johnson et al. 2014). Lakes and streams are often perceived as structurally and functionally different ecosystems, with major differences related to water residence times, the importance of benthic versus pelagic production, and connectivity to the surrounding landscapes (Kratz et al. 1997; Allan 2004). For example, streams are probably more affected by natural disturbance associated with heavy precipitation (e.g. spates) and have strong lateral connections to terrestrial-aquatic edges (e.g. riparian vegetation), while the longer retention times of lakes may result in stronger trophic interactions

Figure 3.18: Representative pictures of four stream sites from the gradient studies of inland waters: (1) an agricultural stream strongly affected by high nutrient and sediment levels; (2) an almost entirely dry stream bed downstream from a hydropower dam. *Photo: Amélie Truchy (1,2).*

(e.g. effects of zooplankton on phytoplankton) (Johnson et al. 2014).

To design and implement cost-effective management programmes and to strengthen the use of a multiple-lines-of-evidence approach to monitoring and assessment, more knowledge is needed of the response trajectories and uncertainty associated with the use of different taxonomic groups and functional measures to detect ecological change. Ideally, the selection of a response indicator should be a knowledge-based decision using stress-response information to select the "best" indicator, that is, the indicator with the highest accuracy and precision (e.g. Johnson 1998; Pocock and Jennings 2008). The main objective of the WATERS freshwater gradient studies was to evaluate the precision and sensitivity of different taxonomic groups to selected pressures. This knowledge can be used to improve our understanding of stress-response relationships and should aid in designing more robust management programmes.

Effects of nutrients, hydromorphological alteration, and forestry in focus

The freshwater gradient studies focused on three main pressures (Fig 3.18): (1) nutrient enrichment (for both lakes and streams), (2) hydromorphological alteration (for streams only), and (3) forestry (for streams only). Each pressure was studied in a region where we expected strong effects on freshwater habitats; for example, nutrient enrichment was studied in an agricultural area in the boreonemoral ecoregion, while forestry was studied in the northern boreal region. Hydromorphological alteration was studied in both the north (e.g. ditching) and south (e.g. small hydropower dams, with sampling conducted within 100 m of the dams at impacted sites), representing the pervasive extent of this impact. For each gradient in each region, 10–16 sites were sampled spanning an increasing gradient of degradation, with an identical set of biological indicators sampled from each. Results are presented presented here for the main BQEs in the WFD: benthic macroinvertebrates (including both littoral and profundal invertebrates from the lakes), diatoms, macrophytes, and fish from both lakes and streams, and phytoplankton from the lakes only.

Further details of the gradient studies, including descriptions and maps of study sites and references for sampling protocols, can be found in McKie et al. (2016).

Figure 3.18: (3) a stream impacted by clear-cutting; and (4) a forested reference stream.
Photo: Dan Evander (3, 4).

Data analysis

Community composition and diversity responses were compared with selected environmental pressures and complex pressure gradients. Pressure-response relationships were assessed separately for each gradient, based on correlations between selected measures of community composition and diversity for each BQE and on selected abiotic parameters for each gradient. The abiotic parameters included:

(1) an axis from a principal component analysis (PCA) ordination summarising the environmental variability associated with the pressure variability among sites, and

(2) three focal abiotic variables that strongly characterise human-induced and/or natural environmental variation along each gradient.

For the lake agricultural gradient, the three additional environmental predictor variables were total phosphorus as a measure of nutrient enrichment, and turbidity and total organic carbon (TOC) as measures of water clarity and colour, respectively. For the stream agricultural gradient, the three additional variables were total phosphorus, turbidity, and the extent of agricultural ditching as a measure of changes in catchment drainage. The hydropower gradient was represented by three variables characterising different aspects of the influence of river regulation on hydrology: the volume of water regulated, the number of rapid flow rises, and the date of maximum water flow (typically delayed in regulated rivers). For the stream forestry gradient, the variables were the percentage of forest clear-cutting and extent of ditched area in the forest catchment, as measures of forest integrity and drainage impact, respectively, and TOC concentration, representing a strong natural gradient of water colour in the region.

Non-metric multidimensional scaling (nMDS) was used to summarize community composition via three community composition axes for each BQE. BQE diversity was summarised by measures of taxon richness and abundance and by two complementary diversity indices: Simpson's (which emphasises abundant species) and Shannon's (which emphasises rare species). Pressure-response relationships were quantified based on Kendall's tau rank correlation coefficient, which is robust to small sample sizes and nonlinearity. The focal indices of community composition (nMDS axes) and diversity were chosen because they are easily compared across all BQEs. Correlations significant at $p < 0.05$ are highlighted, but results significant at $p = 0.05–0.1$ are also flagged as indicative of further potentially important variability.

Note that each pressure gradient represents a correlated suite of stressors. Hence, relationships between particular stressors and the BQEs are presented here for comparative purposes only, and do not imply causality.

All outputs of stressor response analyses are presented in their entirety in the Supplementary information to Section 4.3. The supplementary output includes both graphic plots of the correlation coefficients and statistical significance levels.

Diatoms sensitive to variation in total phosphorus in lakes

Among BQEs, the community composition of diatoms consistently responded to variation in total phosphorus and turbidity, with evidence of shifts in the community composition of littoral invertebrates, phytoplankton, and fish in response to at least some of these parameters (Supplementary information). The diversities of diatoms and phytoplankton all tended to increase with increasing phosphorus, total phosphorus, and TOC. There were also associations between the diversity of profundal invertebrates and total phosphorus concentrations (Supplementary information), especially in combination with variables related to water clarity (i.e. turbidity and TOC). There was also weak evidence of an association between macrophyte diversity and turbidity (Supplementary information).

Turbidity primarily affected composition of diatoms in streams

The community composition of diatoms and benthic invertebrates shifted along the agricultural gradient, with evidence of a macrophyte response as well (Supplementary information). Along with diatoms and invertebrates, the community composition of fish also responded to variation in total phosphorus, while turbidity primarily affected the composition of diatoms (Supplementary information). However, turbidity was strongly associated with fish diversity, while diatom diversity responded more to the overall PCA axis and to total phosphorus concentrations (Supplementary information). Agricultural ditches were associated with variation in the community composition of fish, benthic invertebrates, and macrophytes, with evidence of an association with diatom composition as well (Supplementary information).

Macrophytes were affected the most by hydropower impact

None of the BQEs responded strongly to the overall PCA axis of hydropower impact (Supplementary information); rather, different groups responded more to particular aspects of the hydropower impact. Macrophytes were affected by the largest number of hydropower-related variables, with community composition altered as both the volume of water regulated upstream and the number of flow rises increased, while macrophyte diversity was reduced by the delayed timing of maximal flow peaks in regulated rivers (Supplementary information). Invertebrate community composition also shifted as the number of flow rises and timing of maximal flows increased. Diatom community composition appeared unaffected along the hydropower gradient, but diatom diversity increased with increasing volumes of water regulated. There were indications that fish also responded to some aspects of hydropower, but these effects were not significant at the 5% level (Supplementary information).

Natural gradient of total carbon had greater impact than did forestry

None of the BQEs responded strongly to the PCA axis characterising the overall forestry impact (Supplementary information). Strong negative correlations with the Shannon's and Simpson's diversity indices for macrophytes were apparent, but these were not significant, likely reflecting low statistical power following the exclusion of samples in which no macrophytes were recorded (Supplementary information). The percentage of catchment clear-cuts was associated with variation in the community composition of benthic invertebrates, diatoms, and macrophytes and in the diversity of invertebrates. The extent of forest ditching was associated with altered macrophyte community structure and with changes in the diversity of both macrophytes and diatoms (Supplementary information). However, overall, the natural gradient of TOC was associated with the greatest number of responses in community composition and, especially, BQE diversity in this stream forestry gradient (Supplementary information, note that the TOC gradient was completely orthogonal to the main indicators of forest management, indicating it represents background variability unrelated to forest management).

New specific indicators for hydropower and forestry need to be developed

Overall, primary producers responded most strongly along the pressure gradients. Both diatoms and phytoplankton responded to nutrients and other disturbances associated with agriculture. Macrophytes responded to disturbances associated with hydropower and forest management. Invertebrates frequently responded also, while few responses were observed for fish (e.g. to nutrients in lakes and turbidity in streams). However, the responses of particular groups (e.g. anadromous fish to the presence of dams) or variables (e.g. fish biomass increase with nutrients) were likely missed by the broad comparative approach employed here.

Forestry and hydropower are pervasive pressures in Sweden that have not previously been subject to extensive monitoring. Our preliminary analyses suggest that particular aspects of these pressures can affect the composition and diversity of BQEs, including modifications to the hydrological regime caused by hydropower dams and the extent of clear-cuts and forest ditches associated with forestry. Notably, though, the BQEs in forestry-affected streams appeared to be affected more often by background variation in TOC among sites than by forestry activities. Overall, these findings point to the need to develop particular approaches and indicators for hydropower and forestry. The gradient studies indicate that indicators based on the diversity and composition of macrophytes might be particularly useful for detecting hydropower and forestry impacts, reflecting the sensitivity of particular species to changes in light environments and hydrological regimes.

Selecting robust indicators to detect change with known levels of uncertainty

Selection of BQEs based on which to infer the ecological status of ecosystems is challenging, and multiple indicators may be needed to distinguish the effects of multiple anthropogenic pressures (Dale and Beyeler 2001; Heino et al. 2005; Johnson et al. 2006c). Ideally, the selection of response indicators should be a knowledge-based decision using stress-response information to select the most precise and accurate indicators (Johnson 1998; Pocock and Jennings 2008; Kelly et al. 2016). Indeed, different BQEs are expected to respond differently to different pressures depending on their sensitivity and resilience to stress (e.g. Johnson et al. 2006a; Johnson and Hering 2009). For example, according to first-principle relationships, primary producers such as phytoplankton and benthic diatoms are expected to respond more rapidly to changes in nutrient concentrations than are consumers (e.g. fish and invertebrates), which respond to secondary (i.e. indirect) effects related to food resources, oxygen conditions, and habitat.

In designing biomonitoring programmes, conceptual models are a useful tool for envisaging the responses of indicators to different types of disturbance (Table 3.7). For example, as agricultural land use often results in the eutrophication of streams and lakes, it can be hypothesised that primary producers, such as benthic diatoms, phytoplankton, and macrophytes, due to their first-principle relationship with nutrients, will respond strongly to modest changes in nutrients (Johnson and Hering 2009). This conjecture is based on the assumption that at high concentrations nutrients no longer limit growth and production. Consumers, both primary (many of the invertebrate taxa) and secondary (many of the fish taxa), might also be expected to respond to nutrient enrichment, albeit less strongly than do the primary producers, as their response is more likely related to changes in basal resources. Agricultural land use often results in a number of other effects on aquatic ecosystems. Increased productivity results in decreased oxygen concentrations, increased loads of fine sediment that smother in-stream habitats, loss of riparian vegetation affects incident light, water temperature, and allochthonous sources of energy, and water abstraction alters hydrology and connectivity. Empirical relationships among taxonomic responses to pressures can be used to select single or combined indicators that most accurately measure the desired response with known levels of uncertainty (e.g. Johnson et al. 2006b, 2006c).

Table 3.7. Conceptual model of agricultural pressure-response relationships of BQEs in lakes and streams.

	Phytoplankton	Benthic diatoms	Macrophytes	Benthic invertebrates	Fish
Nutrients	Strong	Strong	Moderate	Moderate	Moderate
Fine sediments	Poor	Moderate	Moderate	Strong	Moderate
Channelisation	Poor	Moderate	Strong	Moderate	Strong
Altered riparian habitat	Poor	Moderate	Moderate	Strong	Moderate
Oxygen	Poor	Poor	Poor	Strong	Moderate

As expected, phytoplankton and benthic diatoms are sensitive to agricultural pressure

From our conceptual model, we predicted that primary producers would display more sensitive responses to elevated nutrients than the indirect, secondary responses of fish and benthic invertebrates. Results from the gradient studies revealed that community composition and diversity responded more strongly to agricultural pressures than to altered hydrogeomorphology and forestry pressures (Table 3.8a and Section 4.3). In agreement with our prediction, phytoplankton and benthic diatoms displayed relatively strong correlations with agricultural land use, and particularly with total phosphorus concentration (Table 3.8a and b and Section 4.3). The responses of macrophyte and benthic invertebrate assemblages were weaker. Interestingly, significant (for invertebrates) or almost significant (for macrophytes) relationships were noted with catchment land use classified as agriculture, but the relationships with total phosphorus concentration were weaker. This finding implies that macrophytes and invertebrates may be responding to other environmental drivers that partly covary with elevated nutrients, such as changes in microhabitat (e.g. increases in fine sediment) associated with land use. Our prediction that fish assemblages would display a weaker relationship with nutrients than would primary producers was not supported, i.e. fish assemblages in both lakes and streams were better correlated with [TP] ($\tau > 0.54$) than were macrophytes and benthic invertebrates. Although our studies focused on detecting responses to degradation, similar principles should often apply when quantifying early responses to improvement.

Stream channelisation and water regulation are widespread, pervasive pressures in landscapes dominated by agriculture, forestry, and urbanisation. The composition and diversity of macrophyte assemblages are often related to habitat heterogeneity; in particular, in-stream (substratum) and edge (riparian) habitats are important predictors. In a study of regulated and unregulated streams, Baattrup-Pedersen et al. (1999) found that that macrophyte diversity was related to habitat heterogeneity; specifically, the spatial distribution of pool/riffle sequences and coverage of coarse substrata were strong predictors of differences in macrophyte assemblages between regulated and unregulated streams. In agreement with earlier studies, we found that water regulation resulted in significant changes in macrophyte assemblages. Contrary to our predictions, benthic diatom assemblages also responded significantly to regulation. This could be a direct response to an abiotic variable such as altered hydrology or indirectly related to a response to another biological variable. For example, macrophytes are known to affect the biodiversity of many other BQEs, as they provide food resources, shelter from predation, and increase habitat heterogeneity (e.g. Johnson and Hering 2010). Hence our finding that the diversity of benthic diatoms was related to water regulation might be a secondary response related to macrophytes.

Effect of forestry on stream ecosystems is strong but poorly understood

An interesting finding of our gradient study and other work within the WATERS programme (e.g. Johnson et al., in manuscript a, b) is that forestry has a strong and so far poorly understood effect on the biological integrity of stream ecosystems. In the forestry gradient study, the taxonomic compositions of three of the four studied groups were significantly correlated with clear-cut logging. Taxon responses of stream assemblages to clear-cut logging were strongest for macrophytes, followed by benthic invertebrate and benthic diatom assemblages. Although the mechanisms are unclear, one potential pressure related to forestry is the alteration or loss of riparian habitats. Johnson et al. (in manuscript b), in a study of multiple pressures on stream invertebrate assemblages, found that loss of riparian integrity was the strongest predictor of taxonomic and trait composition. Likewise, the benthic invertebrate assemblages of lake littoral habitats have been demonstrated to be affected by forestry (Johnson et al., in manuscript a). The alteration or loss of riparian vegetation affects incident light, water temperature, and basal food resources. Riparian vegetation has been demonstrated to influence gross primary production and ecosystem respiration (Naiman et al. 1993; Burrell et al. 2014), decomposition of leaf litter (Lagrue et al. 2011), dispersal of aquatic insects (Carlson et al. 2016), and water quality (Burrell et al. 2014).

Biological responses are context dependent

An important feature of the WFD is its focus on river basins in the management of aquatic resources. Lakes and streams, as mirrors of the landscape, constitute ideal model systems to track changes in catchment land use. Furthermore, as monitoring efforts move from single sites to whole catchments (European Commission 2000; Hering et al. 2010), as discussed below, consideration should be given not only to what BQEs or microhabitats within systems are best correlated with disturbance, but also to what systems (e.g. streams, lakes, and coastal areas) provide the most accurate signal of activities occurring in the catchment (e.g. Johnson et al. 2014) and how many sites are needed to quantify variability within a water body. Indeed, there is increasing awareness that the responses of various BQEs may be context specific, particularly at broad spatial scales (Angeler 2007), resulting in habitat- and trophic-level differences within the same BQE. For example, benthic invertebrate assemblages of riffle habitats of streams have been demonstrated to be better correlated with elevated nutrients than are assemblages of pool habitats (Carlson et al. 2013), benthic invertebrate assemblages of vegetation-free, hard-bottom substratum are often more strongly related to eutrophication than soft-bottom, macrophyte-dominated habitats (e.g. Tolonen et al. 2001; Donohue et al. 2009) and lower trophic levels (e.g. phytoplankton) have been demonstrated to respond more strongly to changes in lake acidity than do higher trophic levels (e.g. benthic invertebrates) (Stendera and Johnson 2008).

Differing in their sensitivity, individual species and BQEs display stress-specific responses to environmental pressures (Figure 3.19). In addition, the concept of selecting not only the appropriate BQE but also the appropriate ecosystem/habitat, exemplified by inland waters, has been illustrated. In this example, organisms in stream habitats are depicted as responding at lower levels of pressure than do organisms inhabiting lakes. The rationale is that streams are probably more affected by natural disturbances associated with heavy precipitation (e.g. spates) and have strong lateral connections with terrestrial-aquatic edges (e.g. riparian vegetation), while the longer retention times of lakes may result in stronger trophic interactions (e.g. effects of zooplankton on phytoplankton). This suggests that streams respond more rapidly to anthropogenic pressures, but that false-positive error rates may be high due to high natural variability. Conversely, lakes may respond more slowly than do streams to environmental pressures, but the associated uncertainty is lower,

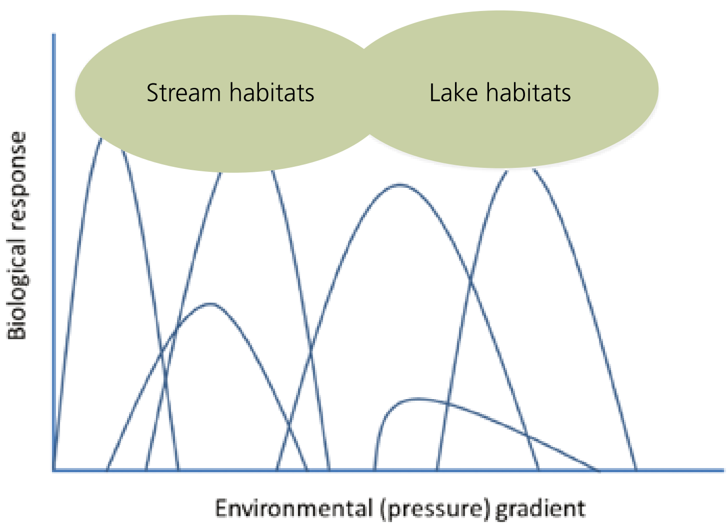


Figure 3.19: Schematic of the responses of species and habitats to an environmental pressure.

Table 3.8. Assemblage composition (nMDS axes 1–3 scores) and Shannon diversity responses of different BQEs in lakes (a) and streams (b) to complex (PCA) gradients characterising agricultural, hydro-geomorphological (HYMO), and forestry pressures and individual stressors (i.e. TP, water regulation, and clear-cut logging). Correlations using Kendall's τ. *** P < 0.05, ** P < 0.01***, • P = 0.05–0.1. Data are taken from the gradient studies described in Section 3.7. Bold text indicates significant relationships. For community composition, absolute values are shown, and if more than one relationship is significant, only the highest correlation is shown. *na* = not applicable, *ns* = not significant.

	Community composition			Shannon diversity		
A.	**Agriculture**	**HYMO**	**Forestry**	**Agriculture**	**HYMO**	**Forestry**
Lakes						
Phytoplankton	*ns*	na	na	***0.7333 *****	na	na
Benthic diatoms	***0.5556 ****	na	na	ns	na	na
Macrophytes	*ns*	na	na	ns	na	na
Invertebrates – littoral	*ns*	na	na	ns	na	na
Invertebrates – profundal	*ns*	na	na	0.4222 •	na	na
Fish	***0.5111 ****	na	na	ns	na	na
Streams						
Benthic diatoms	***0.7333 *****	*ns*	*ns*	***0.6444*****	*ns*	*ns*
Macrophytes	0.4444 •	*ns*	*ns*	*ns*	*ns*	*ns*
Benthic invertebrates	***0.5111 ****	*ns*	0.3167 •	*ns*	0.4222 •	*ns*
Fish	0.4535 •	*ns*	*ns*	*ns*	0.4495 •	*ns*

B.	**[TP]**	**Regulated**	**Clear-cut**	**[TP]**	**Regulated**	**Clear-cut**
Streams						
Phytoplankton	***0.5111 ****	na	na	***0.6889 *****	na	na
Benthic diatoms	***0.6000 ****	na	na	*ns*	na	na
Macrophytes	*ns*	na	na	*ns*	na	na
Invertebrates – littoral	*0.4667* •	na	na	*ns*	na	na
Invertebrates – profundal	*ns*	na	na	0.4667 •	na	na
Fish	***0.5556 ****	na	na	ns	na	na
Streams						
Benthic diatoms	***0.7191*****	*ns*	***0.3905 ****	***0.4944 ****	***0.6445 ****	*ns*
Macrophytes	ns	***0.5968 ****	***0.4833 *****	*ns*	*ns*	*ns*
Benthic invertebrates	0.4495 •	*ns*	***0.4667 ****	*ns*	ns	***0.5167 *****
Fish	***0.5394 ****	*0.4535* •	*ns*	*0.4318* •	ns	*ns*

resulting in low false-negative error rates. Few studies have analysed habitat-specific responses (e.g. Stendera and Johnson 2008; Johnson et al. 2014). Comparing the responses of primary producers in streams and lakes to elevated nutrients, Johnson et al. (2014) found that the responses of stream communities could generally be measured with higher precision than could those of lake communities. For example, approximately 50% of the variability in benthic diatom assemblages in streams was related to total phosphorus concentration compared with 13% for phytoplankton in lakes. Findings from our gradient studies further support the contention that stream ecosystems respond more rapidly than do lake ecosystems to pressures.

Findings from the gradient studies further support the conjecture that responses are not only taxon specific but also context dependent. As discussed above, benthic diatoms and phytoplankton were strongly correlated with total phosphorus concentration. In line with our prediction that the relationships would differ between lakes and streams, we noted stronger responses in streams than in lakes (i.e. diatom assemblages had coefficients of determination of 0.72 in streams compared with 0.60 in lakes) (Table 3.8b). This latter finding supports the concept that responses are habitat (i.e. context) specific. However, caution is advised when interpreting findings from the gradient studies, as the studied lakes and streams were located in different catchments. Nonetheless, findings from WATERS and other recent studies suggest that, in designing monitoring programmes, greater consideration should be given to choice of habitat so as best to detect changes occurring in the catchment, i.e. ecosystems (streams, lakes) and habitats within ecosystems (pelagic, benthic). Combining knowledge of the stress-specific responses of BQEs and habitats should be used to design more robust and cost-effective monitoring programmes to detect degradation and to track recovery after management interventions.

Ecosystem responses to pressures detect early warnings of potential impairment

In summary, our results illustrate how conceptual models and empirical data on taxon- and habitat-specific responses to pressures are useful for designing monitoring programmes. Although needing further study, our results suggest that stream ecosystems better reflect changes occurring in the catchment than do lake ecosystems. Knowledge of not only how BQEs but also ecosystems respond to pressures is needed in order to detect early warnings of potential impairment as well as responses to rehabilitation.

There is increasing awareness that multiple pressures are often prevalent and that interactions among stressors are complex (e.g. Matthaei et al. 2010). Consequently, although not discussed here, management decisions based on individual stressors alone may result in inappropriate programmes of measures, for example, addressing the wrong pressure. For example, Johnson et al. (ms), in studying multiple-pressure effects on lowland streams, found that benthic invertebrate assemblages were responding more to land use effects on riparian vegetation and less to in-stream water quality. This type of knowledge is essential for understanding cause-effect relationships when designing programmes of measures. Use of taxon- and habitat-specific Pressure-response relationships combined with a better understanding of multiple-pressure effects on ecosystem taxonomic composition and processes are needed in order to design future and revise current monitoring programmes to achieve good ecological quality in our inland waters.

4. Coastal WATERS

4.1 General introduction to coastal waters
Sofia A Wikström, Lena Bergström, Mats Blomqvist, Jakob Walve, Antonia Nyström Sandman

4.2 Phytoplankton
Jakob Walve, Helena Höglander, Agneta Andersson, Bengt Karlson, Chatarina Karlsson, Marie Johansen

4.3 Coastal macrophytes
Sofia A Wikström, Mats Blomqvist, Jacob Carstensen, Dorte Krause-Jensen, Susanne Qvarfordt, Antonia Nyström Sandman

4.4 Benthic macrofauna
Marina Magnusson, Rutger Rosenberg, Mats Blomqvist, Kjell Leonardsson

4.5. Coastal fish
Lena Bergström, Leif Pihl, Martin Karlsson, Jacob Carstensen

4.6 Integrated assessment and responses in coastal waters
Leif Pihl, Lena Bergström, Mats Blomqvist, Marina Magnusson, Antonia Nyström Sandman, Jakob Walve, Sofia A Wikström

4.1 General introduction to coastal waters

The Swedish coast is characterised by a strong gradient of salinity, extending from oligohaline conditions (<5‰) in the northern and coastal areas of the Baltic Sea to fully marine conditions below the halocline in the Skagerrak. This has implications for efforts to develop ecological status indicators for assessing anthropogenic impact.

Low salinity is a strong stressor for marine organisms. Coastal areas with low or variable salinity are thus characterised by impoverished marine communities, consisting of a few species with a wide salinity tolerance. However, some freshwater species are also able to extend into the brackish waters of the Baltic Sea, which means that the communities of the Baltic Sea consist of a mixture of marine and freshwater species that all have to cope with suboptimal salinity conditions.

Since species composition changes strongly along the salinity gradient along the Swedish coastline, indicators including species richness or species composition must use a relevant baseline or reference based on the species that can occur at a given salinity. In the current assessment criteria, this is mainly handled through division of the coast into coastal types that differ in salinity and other natural factors. However, the salinity variation, in both space and time, can be considerable within coastal types, which may introduce uncertainty into the measured indicators.

The strong salinity gradient along the Swedish coast poses a challenge when developing indicators for anthropogenic stress. To detect change that can be ascribed to anthropogenic pressures, it is necessary to control for natural gradients, particularly the salinity gradient.

The most widespread anthropogenic stressor in Swedish coastal waters is eutrophication, driven by nutrient discharge from mainly land-based activities. Increased nutrient concentrations have a strong influence on marine ecosystems, stimulating phytoplankton production, which leads to decreased water transparency and causes oxygen depletion in deep areas. This can have large effects on species composition and in some cases lead to nuisance blooms of algae.

Besides anthropogenic nutrient loading, nutrient availability is regulated by a number of factors, such as freshwater inflows, leakage from sediments, upwelling of nutrient-rich deep water, and strong stratification restricting vertical water exchange. Most of the nutrients enter the sea from diffuse sources, making it difficult to define areas of high or low pressure. In addition, the water bodies are strongly connected and offshore areas are also affected by eutrophication. The lack of minimally disturbed areas that can be used as references for natural conditions is another challenge for the development of assessment criteria for coastal areas.

Another important pressure to Swedish coastal areas is environmental toxins, which can have large effects on coastal species and communities.

High variability and poor response to pressures

The current Swedish assessment criteria are based on three taxonomic groups or biological quality elements (BQEs): phytoplankton, macrovegetation, and benthic invertebrates (SEPA 2007). The first management cycle revealed a number of weaknesses in the assessment criteria. For instance, some indicators display high variability that results in high uncertainty in the status assessment and some indicators respond poorly to known pressures.

The criteria were developed based on the data available at the time, and for some indicators and areas it was not possible to conduct appropriate testing of the indicators' response to pressures.

The use of the assessment criteria in the first management cycle also made it clear that the various BQEs sometimes yield very different results for the same water bodies. For instance, the current macrophyte indicator generally indicates a higher status than do the indicators for the other BQEs. It is not surprising that different organism groups respond differently to the same anthropogenic pressure, but the large differences in some Swedish coastal areas may also be driven by differences in how the reference conditions and class boundaries are set. The assessment criteria for different coastal BQEs were developed independently and using different approaches. For instance, the methods used to set reference values and class boundaries differ strongly between the BQEs (Johnson et al. 2014; see further Section 5.1).

Improved knowledge needed to improve Swedish assessment criteria

The overall aim of the coastal part of the WATERS programme has been to produce knowledge that can be used to improve the Swedish assessment criteria for coastal waters, particularly for assessing coastal eutrophication. Specific aims have been to:

- collate existing biological and environmental data
- validate existing indicators
- when necessary, improve existing indicators or identify and suggest new indicators
- improve methods for field sampling and indicator calculation
- improve methods to reduce uncertainty in assessment
- quantify the response of different BQEs in a common pressure gradient (euptrophication)

Fish is not a mandatory element in the Water Framework Directive (WFD) assessment of coastal waters. Fish are to be assessed for transitional waters, which are not present in Sweden. However, coastal fish are assessed as part of the Marine Strategy Framework Directive (MSFD) (EC 2008). WATERS has supported the development of indicators and indicator-based assessment of coastal fish in Sweden, with the aim of facilitating the harmonised overall assessment of the environmental status of the Swedish coastal ecosystem.

4.2. Phytoplankton in coastal waters

OVERVIEW OF CURRENT INDICATORS

Total biomass: currently the only phytoplankton indicator

According to the WFD, phytoplankton status classifications should be based on phytoplankton abundance, biomass, and taxonomic composition, as well as on the frequency and intensity of blooms. Currently, the status of phytoplankton in Swedish coastal waters is classified only according to the chlorophyll-*a* concentration (a proxy for phytoplankton biomass) and the total biomass of phytoplankton (i.e. autotrophs and mixotrophs) measured as biovolume (SwAM 2013) (Table 4.1). When biovolume data are missing, the classification is based on chlorophyll-*a* only. Since chlorophyll-*a* is a relatively easy and low-cost parameter to measure, it is the most commonly used proxy for phytoplankton biomass in assessment systems around Europe. Systems for assessing coastal water that include total phytoplankton biovolume from microscopic analysis are less common, and phytoplankton indicators using species composition have not yet been used in Sweden (e.g. Höglander et al. 2013 and references therein).

Table 4.1. Phytoplankton indicators used in assessing the ecological quality of coastal waters in Sweden and changes suggested by WATERS.

Indicator	Sampling period	Suggested revisions and comments
Biovolume	June – August	Recommended revisions: use new assessment periods[1]; use carbon biomass instead of biovolume
Chlorophyll-*a*	June – August	Recommended revision: use new assessment periods[1]
Index based on taxonomic groups or genera		New potential indicator
Index based on potential harmful genera/taxa		New potential indicator

1) Gulf of Bothnia: July – August or July – September; Baltic Proper: July – August; Kattegat and Skagerrak: May – September.

Main aim:
phytoplankton species composition indicators in relation to pressure

Eutrophication is anticipated to increase chlorophyll-*a* and total phytoplankton biovolume to unacceptable levels. However, an assessment based on only these variables ignores the fact that different species may have different impacts on the ecosystem and on water quality.

Several attempts have been made to find phytoplankton indicators based on classes, orders, genera, or species and their responses to nutrients (reviews in Höglander et al. 2013 and Walve et al. 2016), but so far with limited success for coastal marine waters (see e.g. Carstensen 2015 for review). The main aim of the present work was to evaluate promising phytoplankton composition indicators in relation to anthropogenic pressure, mainly eutrophication, for different geographic areas around the Swedish coast: the Gulf of Bothnia, Baltic Proper, and Kattegat and Skagerrak. Important phytoplankton groups and species for the different Swedish coastal areas and promising indicators from other studies (see Höglander et al. 2013 and references therein) were evaluated according to their responses to elevated nutrient concentrations, and thus their suitability as indicators of eutrophication.

When developing phytoplankton indicators, there are several challenges. Generally, lack of data has so far hampered the indicator development. Although some valuable long time series exist, their spatial coverage has been limited. In this work we were able to include a spatially more extensive dataset, covering a wider range of nutrient conditions.

Natural gradients and season influences phytoplankton

Many natural factors influence the occurrence and growth of various phytoplankton species, such as nutrient availability and variations in salinity, temperature, Secchi depth (i.e. water transparency), and water exchange. There is also a general gradient along the Swedish coast with respect to the limiting nutrient – with phosphorus being limiting in the Bothnian Bay, near balance in the Bothnian Sea and nitrogen being limiting in the Baltic Proper and in the Kattegat and Skagerrak (Graneli et al. 1990; Andersson et al. 1996; Boesch et al. 2006) – which may influence the community composition of phytoplankton. The strong salinity gradient from the Bothnian Bay to the Skagerrak is a general structuring factor affecting the relative occurrence of species of marine and freshwater origin. In addition to the large-scale salinity gradients, gradients from the mouths of rivers to open coastal areas affect species distributions on a smaller scale.

There is also a large-scale gradient of light conditions (often measured as Secchi depth, i.e. water transparency) due to the gradient of coloured dissolved organic matter (CDOM), with the highest concentrations being found in the northern Baltic Sea and lowest in the Skagerrak. As for salinity, local freshwater influence can cause substantial variability in CDOM along the coast of Sweden (Harvey et al. 2016). The CDOM content influences the structure of the plankton communities because of reduced light availability and because CDOM may be a source of energy and nutrients for pelagic bacteria and also for eukaryotic plankton, and high concentrations can promote mixotrophic species (Paczkowska et al. subm.). The effect of CDOM on light conditions can also lead to increased chlorophyll-*a* content in the cells, without corresponding biomass increase (e.g. Andersson et al. 1989).

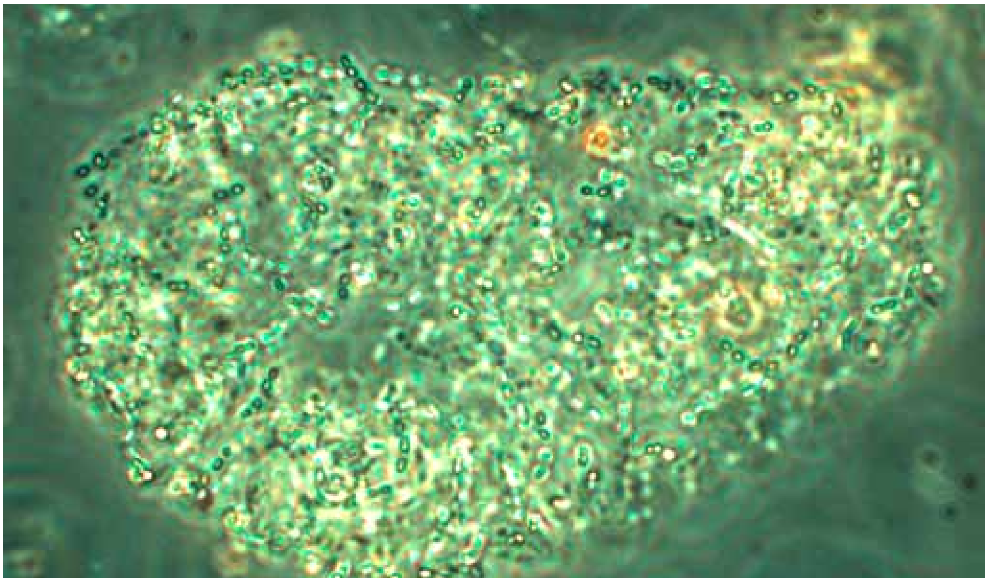
The composition of phytoplankton communities is strongly influenced by seasonal succession. This is a challenge, since variability of the seasonal succession of phytoplankton communities could override nutrient effects. The season starts with a spring bloom, which usually evolves as a community dominated by various species of diatoms, which are successively replaced by dinoflagellates. Normally, after a relatively abrupt termination of the spring bloom, a completely different summer community progressively evolves from May/June until September. The succession of phytoplankton species and groups will affect the uncertainty of nearly any indicator based on species composition, unless the sampling programme is very intensive.

Shortcomings of the current phytoplankton biomass indicators

The current seasonal June – August assessment period was chosen because of its relatively stable conditions and because it is normally an intensively sampled period (Larsson et al. 2005). The use of a common June – August assessment period along the whole Swedish coast, despite large differences in seasonal succession, has received some attention because of observations of high variation in phytoplankton data from June. In some years and areas, particularly in the Gulf of Bothnia, remains of the spring bloom can still be present in June, resulting in high phytoplankton biomass. In other years and areas, a clear-water phase may develop after the spring bloom, resulting in low biomass values in June. For the Swedish west coast and the southern Baltic Sea, this is usually not a problem, since the spring bloom is normally earlier. In this area, however, the current June – August sampling period is not aligned with the May – September sampling period used in neighbouring countries. Moreover, if data are available it may be statistically preferable to include more months in the assessment.

The use of biovolume may overestimate the biomass of certain species, particularly diatoms, which have large volumes but low carbon contents (because of large vacuoles). Because carbon biomass is directly correlated with organic content, the carbon biomass should be a more accurate eutrophication indicator than biovolume.

Reference values for chlorophyll-*a* and biovolume have been defined as fixed values for each type of area (i.e. Swedish west coast and Gulf of Bothnia) or are determined for each sample according to measured salinity (i.e. Baltic Proper) to adjust for the natural background nutrient gradients in coastal areas. A salinity-adjusted nitrogen reference value is used in an empirically determined nitrogen-chlorophyll-*a* or nitrogen-biovolume relationship to calculate the corresponding reference values for chlorophyll-*a* and biovolume. The fixed reference values used in many types of areas may lead to problems for certain water bodies. Where natural nutrient inputs (of both coastal and open-sea origin) contribute to elevated phytoplankton biomasses, it may be difficult to determine an appropriate good/moderate boundary. In other water bodies, the current boundaries may be less strict than is appropriate.



OVERVIEW OF INDICATOR DEVELOPMENT

New valuable datasets available

Our approach to finding eutrophication indicators was to statistically evaluate coastal gradient data in three areas: the Gulf of Bothnia, Baltic Proper, and Swedish west coast. Along with long time series of the national monitoring programme, new datasets were also evaluated. For the Gulf of Bothnia, we studied a campaign dataset collected mainly in the northern Bothnian Sea three to six times in July – August 2011, with a total of 120 samples from 114 stations in 25 water bodies. We used several datasets from the Svealand region, which covers the area from the southern Bothnian Sea to Northern Baltic Proper. The geographically most extended dataset comes from the monitoring and research programme of Svealands Kustvattenvårdsförbund, which sampled 24 – 25 stations for phytoplankton in July – August 2007 – 2013. In addition, we separately evaluated data from a few stations in the Askö – Himmerfjärden area, which are sampled every second week in summer, as well as monthly sampled stations within the Stockholm Vatten monitoring programme (Stockholm Archipelago). The datasets from the Swedish west coast cover the area from the coastal Skagerrak and the Kattegat, including results from the monitoring programmes of Bohuskustens vattenvårdsförbund and Hallands kustkontrollprogram. The analysed datasets from the Swedish west coast include data collected year round, but the focus of the analysis is on June – August 2007 – 2014. The chosen period includes data from regular monitoring as well as a sampling campaign performed in 2011 and 2012 involving an extended number of stations. Additionally, data from two gradient studies performed within the WATERS programme on the Swedish west coast and in the archipelago of Östergötland in the Baltic Proper in 2012 – 2013 are included (see 4.3.3. below). The WATERS gradient study on the Swedish west coast found a gradient of coloured dissolved organic material (CDOM), masking any effects of eutrophication. Therefore the usefulness of this dataset was limited for studying the effects of eutrophication.

Evaluation of length of the seasonal assessment period

The assessment period was evaluated using the existing indicators chlorophyll-*a* and biovolume (Walve et al. 2014). The focus was mainly on the summer period, evaluating May – September data. However, we also looked into the possibility of assessing the spring bloom period. The spring bloom period is dynamic and often short lived, making it difficult to assess. Ecologically it is very important, since the sedimentation of the bloom affects sediment oxygen consumption and food availability for benthic fauna. Nevertheless, to be used as an indicator, it may require a relatively large sampling effort, and we evaluated the usefulness of the common monthly data for spring period assessment. However, to assess the spring bloom, sampling much more frequently than monthly is required. With the unclear improvement from using a long assessment period including the spring bloom, and the principal objections of including the spring bloom if monitoring is only monthly, the recommendation is to exclude the spring bloom and not start monitoring before the spring period is over (Walve et al. 2014). With monthly sampling there is a risk of biased assessment due to changes in the timing of the spring bloom. Therefore, indicators based on the size or succession of the spring bloom must be limited to the very few stations with high-frequency monitoring (Walve et al. 2014).

Our results indicate that, in the Gulf of Bothnia, phytoplankton biovolume data have larger variance and a higher mean concentration in June than in July – September, indicating a spillover effect from the spring bloom. Rather surprisingly, this problem was not evident for chlorophyll-*a*. The low chlorophyll-*a* content of the dominant species at the end of the spring bloom may contribute to this pattern. For the Baltic Proper, our results indicate some systematic differences in chlorophyll concentrations between June and later months, but that these differ between inner stations (higher in June), intermediate stations (similar between months), and outer stations (lower in June). This pattern probably reflects the effects of continued nutrient inputs to inner areas from land runoff after the spring bloom, and indicates that cyanobacteria tend to be relatively abundant in July – August in the outer parts of the Stockholm archipelago. Moreover, inclusion of June data did not reduce the uncertainty.

For the Kattegat and Skagerrak, our results indicate that shortening the evaluation period to July – August may result in a more uncertain assessment. Another possibility is to make the evaluation period longer. The spring bloom often peaks as early as February but may occasionally start in or extend into March and even April. The recommendation is to exclude the spring bloom and thus not start earlier than May. The extended May – September period displayed similar concentrations and variability as did the currently used summer period (June – August). Increasing the length of the evaluation period would ensure more data points for several assessments. Another advantage is that an assessment period extending from May to September would facilitate comparisons with neighbouring waters, as both Denmark and Germany use this period.

Carbon biomass more relevant than biovolume

Because carbon biomass is directly correlated with organic content and biomass (as dry weight), it is a more relevant indicator of eutrophication (as well as a better measure of biomass in food web studies) than is biovolume. We therefore evaluated the performance of carbon biomass instead of biovolume as the biomass indicator. Different species have different carbon concentrations per cell volume (i.e. C/biovolume ratio) according to the biovolume-to-carbon formula in Menden-Deuer and Lessard (2000) commonly used in phytoplankton monitoring. This biases a biovolume assessment towards large vacuolated species with comparatively low carbon content, such as diatoms, while underestimating the biomass of small cells (containing more carbon per biovolume). As a result, the mean carbon-to-biovolume ratio of natural phytoplankton assemblages varies depending on the species composition. The highest C/biovolume ratios were consequently found to be associated with high contributions of various small-sized genera and the lowest with a large proportion of diatoms. We found area specific median C/biovolume ratios for the natural assemblages of species in the Bothnian Sea, Baltic Proper, and west coast areas, respectively. We used these median C/biovolume ratios for these areas to recalculate the reference values and class boundaries based on biovolume to new values based on carbon (Walve et al. 2016).

Evaluation of potential taxonomic phytoplankton indicators

Generally, positive correlations between total nutrients and total phytoplankton biomass are expected and are usually found. This is also the case for many general taxonomic groups (e.g. classes). However, some taxa may be more stimulated by higher nutrient concentrations than others, so that the proportions of their biomass could change with nutrient enrichment. For example, increased nutrient availability is often considered to stimulate the occurrence of diatoms, chlorophytes, and cyanobacteria (Sagert et al. 2008; Phillips et al. 2013; Andersson et al. 2015). The abundance of some taxa may also be negatively affected by nutrients or more affected by factors other than nutrients, leading to weaker correlations with nutrients.

Pros and cons of using general taxonomic groups versus individual species

An advantage of using more general taxonomic groups in an assessment is that the potential error due to differences in skills of taxonomists is reduced. However, species-specific responses within groups may remain undetected. To partly overcome this problem, we divided cyanobacteria into orders, which more or less correspond to different functional groups. We performed correlation studies, i.e. partial least squares (PLS), principal component analysis (PCA), and linear regression, of the biomass of various larger taxonomic groups (mainly classes) and environmental factors within the different datasets from Gulf of Bothnia and Northern Baltic Proper.

With the Gulf of Bothnia dataset, sixteen phytoplankton classes and orders were investigated for relationships with environmental factors (Walve et al. 2016). Unidentified flagellates, Dinophyceae, and Nostocales were positively influenced by salinity, while all the three cyanobacterian orders, i.e. Oscillatoriales, Nostocales, and Chroococcales, and the class Chrysophyceae were promoted by total P concentrations. The relationships were mostly affected by relatively few nutrient-rich stations. Salinity promoted brackish water taxa such as the dinoflagellate *Dinophysis acuminata* and the Nostocales cyanobacterium *Aphanizomenon spp*., and disfavoured freshwater taxa such as Charophyceae spp. and *Pseudopedinella* spp.

Tests on the Svealand dataset (Northern Baltic Proper) indicated that the cyanobacterian orders Oscillatoriales and Chroococcales correlated positively with total N and P, also largely as a result of high abundances and proportions at a few stations with high nutrient concentrations (Walve et al. 2016). Cyanobacteria of the order Nostocales (potentially nitrogen-fixing taxa), and to some extent also Prymnesiophyceae (mixotrophic group), did not correlate with nutrients, but correlated positively with salinity and negatively with inorganic nitrogen. Diatoms and Cryptophyceae correlated positively with inorganic N, which is high in the Stockholm inner archipelago. Here, upwelling deep water containing inorganic N and P stimulates phytoplankton growth. Before the large reductions of P loading to the Stockholm archipelago, there were large biomasses of Oscillatoriales (genus *Planktothrix*) (e.g. Brattberg 1986) in this area. The conditions at that time were similar to the most eutrophic areas in our present dataset, with high biomass of other Oscillatoriales species, although the genus *Planktothrix* is probably limited to areas with low salinities. Chlorophyceae was positively correlated with both total N and total P, but the proportion was still small even at the highest nutrient levels. We cannot exclude that chlorophytes were suppressed by the relatively high salinity (4 – 5) at the highest nutrient levels in the studied dataset.

In conclusion, our studies in the Baltic Sea suggest several possible indicators based on phytoplankton groups, in particular the biomass or the proportion (of total biomass) of the cyanobacterial orders Chroococcales and Oscillatoriales (Walve et al. 2016). Our data suggest that, for these orders, biomasses >25 – 50 µg C L^{-1} or proportions >10–20% of total biomass may indicate excessive eutrophication (Walve et al. 2016). These limits may also apply to other groups, such as the Chlorophyceae.

Phytoplankton taxa were ranked according to nitrogen optima

Usually, the term indicator species refers to species that are either favoured or non-favoured as a result of an increased nutrient supply (eutrophication). A change in the biomass or proportion of the indicator species or taxa is used to assess whether the system is affected by nutrient enrichment. For example, the Swedish freshwater phytoplankton assessment includes a trophic index based on the classification of each included species according to its occurrence in trophic gradients. This disregards whether these species are considered “good” or “bad”, for example, due to potential toxicity.

In the Gulf of Bothnia, twenty phytoplankton taxa were found to have a carbon biomass >10% of the total phytoplankton biomass in any of the single samples (Walve et al. 2016). For ten of the taxa, a significant PLS model was obtained, and for nine of them, driving environmental factors were identified. The Nostocales (nitrogen-fixing) cyanobacterium *Aphanizomenon* spp., the dinoflagellate *Dinophysis acuminata*, and small unidentified flagellates were found to have a positive relationship with salinity, while the diatom *Diatoma tenuis* and the chrysophyceans *Pseudopedinella* spp. and *Uroglena* spp. were found to have a negative relationship with salinity. The cyanobacteria *Pseudanabaena* spp. was found to have a positive relationship with total phosphorus and nitrogen as well as temperature. The mixotrophic ciliate *Mesodinium rubrum* had a weak negative relationship with total nitrogen. High temperature was observed to promote warm-water groups such as Prasinophyceae and *Pseudanabaena* spp., while, for example, *Mesodinium rubrum* was negatively related to high temperature. The latter species is known to occur both in relatively cold spring water and during the summer.

A total of 135 genera were observed in the Svealand area (Northern Baltic Proper) (Walve et al. 2016). For this dataset we choose to rank taxa according to their nitrogen optima, i.e. according to their proportion of total biomass weighted for observed nitrogen concentrations, since nitrogen is the main limiting nutrient in this area. We also conducted the corresponding analysis for salinity and inorganic nutrients. Among the 20 genera having the highest nitrogen optima (i.e. high proportions of total biomass at high total nitrogen concentrations) were ten cyanobacteria genera (six Chroococcales, three Oscillatoriales, and one Nostocales) and six chlorophyte genera (Walve et al. 2016). The genera associated with the lowest nitrogen concentrations were found among the prymnesiophytes (*Chrysochromulina*), chrysophytes, the common Nostocales cyanobacteria *Aphanizomenon* and *Anabaena* (*Dolichospermum*), and the dinoflagellates. Many of the diatom genera were associated with relatively high total nitrogen, but also with high dissolved inorganic nitrogen and low salinity, because of their large contribution to the biomass in the Stockholm inner archipelago.

Successional patterns need to be considered when evaluating phytoplankton indicators

An analysis of the Himmerfjärden dataset (Baltic Proper) indicated that within the June – August period there is a clear succession of some species and groups. For example, the spring species *Mesodinium* is fairly abundant in June, but less common in July – August. This was also reflected in a negative relationship with temperature in the dataset from the Gulf of Bothnia. This highlights that these types of successional patterns must be considered whenever a taxonomic indicator is evaluated.

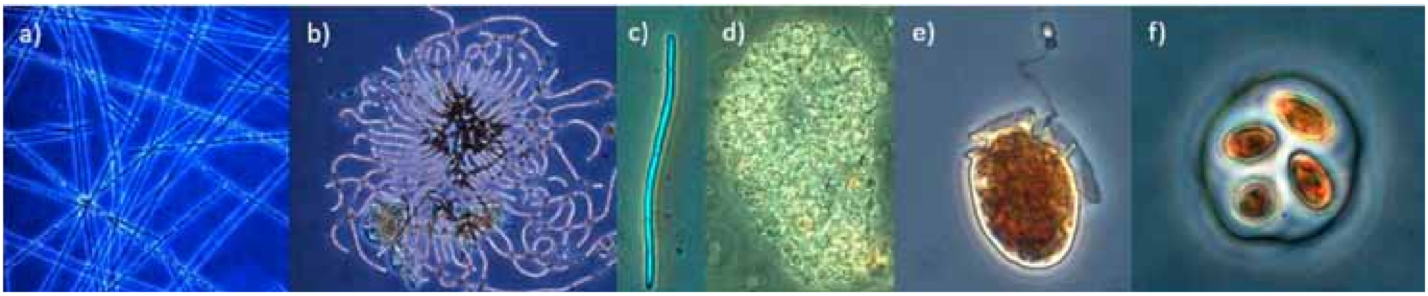


Figure 4.1: Potential phytoplankton indicator taxa: the cyanobacteria genera *Planktothrix* (Oscillatoriales) (a), *Dolichospermum* (b), *Pseudanabaena* (Oscillatoriales) (c), and *Cyanodictyon* (Chroococcales, d); the potentially toxic dinoflagellate *Dinophysis acuminata* (e) and Chlorophyceae (illustrated by the genus Oocystis, f). *Photo: Helena Höglander.*

Certain species may have harmful effects

Some phytoplankton species or genera can be problematic since they have deleterious effects on aquatic animals or on humans, including economic damage, for example, to mussel farming or tourism. Blooms with high densities of cells can discolour the water, drift to shore, and cause oxygen depletion or odour when degrading. Some can produce toxins dangerous to both humans and other species (e.g. Cronberg and Annadotter 2006). The coupling between these harmful algal blooms and anthropogenic nutrient enrichment is, however, not always clear (Gowen et al. 2012) and can differ between areas. Toxic or nuisance species usually develop due to a combination of favourable factors/conditions, such as salinity, temperature, nutrient concentrations and ratios, water stratification, and biological interactions. Due to the many controlling factors, we should not expect a certain species to automatically increase as a result of nutrient enrichment. However, as long as there is a clear elevated risk of increased biomass with nutrient enrichment, a harmful or nuisance species can reasonably be included in the assessment.

Harmful algae at risk levels indicate the need for assessment

We suggest including a selected set of potentially harmful taxa in the assessment (Supplementary Informationa). When considering certain harmful algal bloom species, there are defined warning limits for harmful abundances (e.g. OSPAR 2005; Persson et al. 2014). In our dataset these risk levels were reached on several occasions, indicating the need for such an assessment. For example, there were 18 occasions when *Alexandrium* had a higher biomass than the warning limit of 0.5 µg L^{-1} and 133 occasions when *Dinophysis acuminata* had a higher biomass than the warning limit of 3 µg L^{-1} in a total dataset of 517 samples (Walve et al. 2016). Abundances could also be translated into biovolumes and carbon biomasses for these species.

A possible method for assessing harmful taxa is an index based on the occurrence of selected taxa, which is combined with the existing total biomass indicator. Total biomass is therefore still the foundation of the assessment, reflecting the fact that total biomass is considered a reasonably good indicator of general eutrophication problems (Walve et al. 2016), such as effects on water clarity and oxygen conditions. However, this approach would stress problems directly related to certain problematic species. Briefly stated, ratios of the observed biomasses of harmful algae relative to defined boundaries for each species are calculated, and are then used to modify the ecological quality ratio (EQR) of total biomass (i.e. chlorophyll-*a* and carbon biomass) (Walve et al. 2016). A possibility would be to use the approach suggested for freshwater, i.e. use the indicator only to lower, not increase, the status of a water body.

Reference values and good-moderate boundaries for total biomass vary among (and sometimes within) water body types. It is reasonable also to modify boundaries for individual potentially harmful taxa according to this. A suggestion for species-specific boundaries is to use 50% of the good/moderate boundary of total biomass, unless information supports a different boundary, such as the warning limits discussed above (Walve et al. 2016). An advantage of focusing on harmful species is that it is a pragmatic approach that can be used without considering salinity. If a species is disfavoured because of too low or too high salinity, it will not reach high biomass and will not be a problem.

Species richness and diversity not correlated with nutrients

Species diversity often displays a hump-shaped relationship with system productivity (Smith 2007), with the highest diversity occurring at intermediate nutrient levels, while a few species take over and become dominant at high nutrient concentrations. We thus expected a positive or negative relationship between nutrient concentrations and number of species, depending on the nutrient concentration range.

Generally, there were poor correlations (i.e. low r^2) between diversity indices and salinity and total nitrogen in the tested datasets (Walve et al. 2016). Diversity tended to be higher at the most nutrient-rich stations in the Baltic Proper; this was not found in the Gulf of Bothnia, however, where the diversity instead tended to increase with salinity. Our results support those of Cermeno et al. (2013), who found no relationship between species richness and productivity.

Cell size not correlated with nutrients

Generally, low nutrient concentrations are expected to promote the growth of small phytoplankton and high nutrient concentrations the growth of large phytoplankton (see Walve et al. 2016 and references therein). In our tests of this hypothesis, nutrients and cell sizes were mostly uncorrelated or weakly correlated (Walve et al. 2016). It should be noted that our analysis did not include single-celled autotrophic picoplankton, the smallest (< 2µm) and most abundant phytoplankton, as those are not included in the traditional phytoplankton monitoring of cells or colonies. Picoplankton are currently included only in the national monitoring programme in the Gulf of Bothnia (see Karlson 2014 for more information on phytoplankton monitoring methods).

SUGGESTED REVISIONS OF ASSESSMENT METHODS

Change seasonal assessment periods or correct for missing months

For the Gulf of Bothnia, we suggest changing the assessment period to July – August or July – September (Walve et al. 2014). This will reduce the variance of the phytoplankton biomass by ensuring that the residue of the spring bloom remaining in June is excluded. This will not cause any changes in the classification boundaries. For the Baltic Proper, we suggest changing the assessment period to July – August (Walve et al. 2014). The chlorophyll-*a* concentration in June often deviates systematically from that of the July – August period, but differently in different areas. The use of the July – August period would not increase uncertainty, even if fewer samples are included because monthly sampling is maintained in some areas. Correction factors for the differences between the June – August and July – August periods are suggested (Walve et al. 2014) and can be used as guidelines for new classification boundaries. For the Skagerrak and Kattegat areas, a prolonged May – September assessment period is suggested, as the differences between months are small (Walve et al. 2014). Currently, sampling is generally performed monthly, and an expanded period would generate more data points for assessment. The sampling period would also be aligned with the sampling period used in neighbouring countries.

Use carbon biomass instead of biovolume

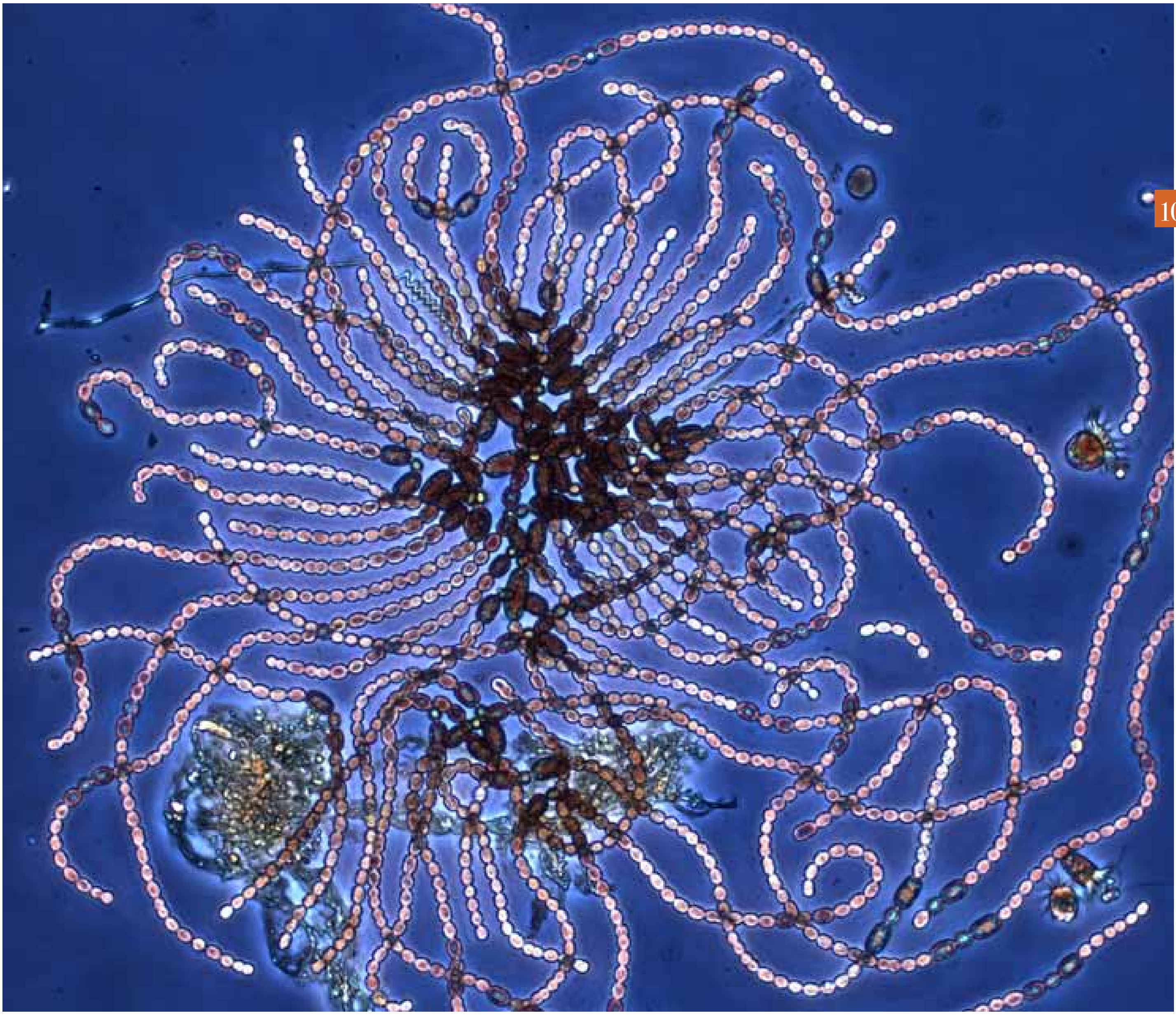
We recommend using carbon biomass (based on the biovolume-to-carbon conversion factors in Menden-Deuer and Lessard 2000) instead of biovolume as a measure of total phytoplankton biomass (Walve et al. 2016). New carbon biomass reference values and class boundaries for different types of areas are suggested based on the median C/biovolume ratios for natural phytoplankton assemblages (Walve et al. 2016).

Use selected taxa or groups to identify the most impacted water bodies

At high nutrient concentrations, there is a clear risk of high abundances and proportions of the cyanobacteria orders Oscillatoriales and Chroococcales. Since these observed eutrophication effects occur at high nutrient levels, changes in species composition will probably not help identify the good/moderate boundary, but will help identify water bodies with the most severe problems. In addition, the Nostocales can be stimulated, provided that inorganic nitrogen concentrations are low. However, contrary to the other orders, Nostocales may often dominate the biomass (as %) even at low total P levels. We suggest that excessive absolute biomasses of the cyanobacteria orders Oscillatoriales, Chroococcales, and Nostocales be used to classify the most affected water bodies in the Baltic Sea (Walve et al. 2016). Other possible indicator taxa are the cyanobacterial genus *Pseudanabaena* and the class Chlorophyceae, but evidence for this was not consistent across datasets (Walve et al. 2016). Moreover, since many combinations of nutrient levels and salinities are not represented in the datasets tested, the generality of our findings is difficult to verify.

Include selected potentially harmful species in the assessment

A possible assessment method is to include an index based on the occurrence of selected harmful taxa, which is combined with the existing total biomass indicator (Walve et al. 2016). Total biomass is then the foundation of the assessment, but problems related to the increased risk of potentially harmful species are stressed. In this report we list several taxa that may be included in the assessment of Swedish coastal areas. Based on available evidence, we suggest that the cyanobacteria genera *Nodularia, Planktothrix*, and *Dolichospermum (Anabaena)* be included as well as the dinoflagellate genera *Alexandrium, Dinophysis, Noctiluca* (Walve et al. 2016), and Karenia and the potentially toxic diatom genus *Pseudo-nitzschia*.



REMAINING CHALLENGES

An index that describes the response to nutrients is a first step in developing a species-based assessment. Once an index has been defined, boundaries for the water types or even water bodies must be determined and evaluated. Since water bodies differ in a number of natural factors, such as the natural level of nutrient input and the water exchange with the open sea, fixed boundaries for each type of area cannot always handle this variation. The salinity-dependent reference values currently used for nutrients (and in the Baltic Proper currently also for chlorophyll-*a* and biovolume) are another way the natural gradients can be taken into account. An important area for further development is to evaluate and develop the system for determining water-body specific reference values for nutrients, chlorophyll-*a*, and total biomass.

Clearly, the search for suitable indicators as well as revisions of current and new proposed indicators must proceed as new datasets are successively available. In particular, time trends in areas with changing nutrient loads are valuable, since comparisons across different areas will inevitably involve many factors affecting the response to nutrients. Of particular interest are threshold nutrient values, where changes are relatively large.

Another important question that needs further attention is how to incorporate taxonomic indicators into the current assessment of coastal phytoplankton. Even though potential indicator groups or species have been suggested, the weighting of indicators will strongly affect the results. Moreover, since we will often have chlorophyll-*a* data only, it is necessary to keep a system that can be used to classify incomplete datasets.

In areas affected by high humic substances and particulate matter (see e.g. Harvey et al. 2016), the class boundaries for chlorophyll-*a* might need further revision, since the Secchi depth (used to set the reference values) is greatly affected by this.

To increase the availability of data for assessment, a future challenge is to evaluate the usefulness of complementary datasets to obtain better data coverage in both time and space. Such methods may include satellite remote sensing (for chlorophyll-*a*) and automated sampling from buoys and ships (for chlorophyll-*a* and phytoplankton).

4.3. Coastal macrophytes

OVERVIEW OF CURRENT INDICATORS

Shortcomings of the current indicator MSMDI

The current Swedish WFD assessment method for macroalgae and soft-substrate macrophytes is the Multi Species Maximum Depth Index (MSMDI; Table 4.2). It is based on the depth limit of three to nine perennial eutrophication-sensitive species. Each species obtains a score based on its depth limit relative to defined scoring depth limits. The MSMDI then averages this information across the involved species to give an overall MSMDI score. This method was developed in 2006 (Kautsky et al. 2006) and implemented in Swedish law in 2008 (NFS 2008:1, EC. 2008). The rationale behind the indicator is the well-documented relationship between eutrophication, leading to reduced water transparency, and the depth limit of vegetation (Duarte et al. 1991; Krause-Jensen et al. 2008).

Table 4.2. Macrophyte indicators used in assessing the ecological quality of coastal waters in Sweden and changes suggested by WATERS.

Indicator	Acronym	Suggested revisions and comments
Multi Species Maximum Depth Index[1]	MSMDI	Not recommended for further use
Depth limit of eelgrass		New suggested indicator for soft substrate in the Kattegat and Skagerrak
Cumulative cover		New suggested indicator for hard substrate (all coasts) and for soft substrate in the Baltic Sea and Gulf of Bothnia
Species richness		New suggested indicator for hard substrate (all coasts)[3] and for soft substrate in the Baltic Sea and Gulf of Bothnia
Macrophyte Sensitivity Index[2]	Mi_c/MI_a	New optional indicator for soft substrate in the Baltic Sea and Gulf of Bothnia

1) Kautsky et al. (2006), 2) Hansen (2012), 3) Not in the Bothnian Bay.

Despite the strong theoretical basis for vegetation depth limit being a good indicator of eutrophication, there are several problems with MSMDI. One major problem is that more than 75% of the existing monitoring transects cannot be used to calculate MSMDI (Blomqvist et al. 2012a). This is because more than 50% of the transects do not extend to a sufficient depth and many of the transects lack suitable substrates below the deepest occurrence of a species or have fewer than the required three indicator species. Ultimately, this is because many transects were originally established for purposes other than measuring depth limits, such as giving a general description of the vegetation along the transects. Furthermore, in many water bodies it can be difficult to find suitable sites for monitoring vegetation according to the demands of MSMDI, particularly in shallow areas and areas dominated by soft substrate.

Another major problem is the methodological challenge of finding the deepest-occurring specimen of a certain species. Since the deepest occurrence is likely to be represented by small and scattered specimens, the depth limits underlying MSMDI will have high uncertainty (Figure 4.2).

In addition to these problems, we have identified mathematical limitations to the index (Blomqvist et al. 2014) and a weak relationship between MSMDI and eutrophication (see further below and in Blomqvist et al. 2014). The evaluation therefore identifies a need for alternative or supplementary vegetation indicators that are useful in all coastal areas, are more responsive to pressures, and can be monitored more objectively and with less uncertainty by divers.

Figure 4.2: The deepest observation of a species is often represented by a single small frond, for example, bladderwrack (*Fucus vesiculosus*) in a mat of *Battersia arctica* in the Bothnian Sea. *Photo: Susanne Qvarfordt.*

OVERVIEW OF INDICATOR DEVELOPMENT

New vegetation candidate indicators identified

The first step towards developing new vegetation indicators was to identify a set of candidate indicators to be tested based on reviews of the MSMDI method, indicators used in other European countries, and existing Swedish vegetation data (Blomqvist et al. 2012a). These indicators represent hard- and soft-substrate vegetation and reflect the distribution, abundance, and composition of the vegetation.

The candidate indicators were then evaluated for their suitability as vegetation indicators in Swedish coastal areas, as described in detail below. This included testing their responses to gradients of nutrient concentrations and Secchi depth, as well as to natural gradients of salinity and wave exposure. We further quantified the spatial, temporal, and methodological variability, or uncertainty, of the indicators.

Based on this evaluation, we suggest a number of new indicators to be used in assessing the ecological status of coastal vegetation. Since the current main field method for sampling coastal vegetation is not optimal for recording all of the new suggested indicators, we also tested new or modified field methods, as described in detail below.

Candidates evaluated by testing response to pressure

We evaluated the candidate indicators by testing their responses to pressures in two sets of analyses. First we established a quality-assured database with large datasets from Swedish vegetation monitoring and inventories (Blomqvist et al. 2014). In this database, large amounts of Swedish vegetation data were compiled and related to environmental data for the first time. The compiled data span the entire 11,500-km mainland coastline and represent both hard- and soft-substrate macrophytes.

We used these extensive data to test the responses of a set of candidate indicators to eutrophication. To identify potential indicator responses to pressures, as much as possible of the variation due to other factors must be accounted for. Therefore, we related the indicators not only to the gradients of eutrophication variables but also to the gradients of other environmental variables. In addition, we accounted for methodological variability as far as possible.

The second set of analyses was performed using data from the coastal gradient study performed within WATERS (see further Section 4.6; Wikström et al. 2016). Here we tested pressure-response relationships for a number of the candidate indicators based on WATERS field studies along well-defined pressure gradients on the Swedish west and east coasts in 2012 and 2013. This helped us overcome the lack of data in the vegetation database for some of the candidate indicators, for example, regarding the depth distribution of eelgrass (Zostera marina; Figure 4.3). In addition, we could also test whether decreased methodological uncertainty would result in clearer vegetation indicator responses to anthropogenic pressures. The main sampling method in the gradient studies was SCUBA diver surveys of large squares (5 × 5 m) placed on either hard or soft substrate within a limited depth range (3 – 5 m) to reduce sampling variability. For soft-substrate communities on the west coast, we surveyed the depth range of eelgrass using video.

In line with accepted theories of vegetation responses to eutrophication, we predicted that reduced eutrophication would lead to deeper vegetation with greater cover, lower representation of opportunistic species, more complex communities, and greater diversity of species and traits. The tested indicators differed in their responses to eutrophication and other environmental variables and in their variability within and between water bodies. We summarise the main findings below.

Depth distribution of eelgrass is a potentially useful indicator

Despite the strong theoretical basis for vegetation depth limits as a good indicator of eutrophication, MSMDI displayed only a weak relationship with eutrophication-related variables (Blomqvist et al. 2014). However, we found that the depth distribution of eelgrass is a potential indicator for soft substrate on the Swedish west coast. This is supported by results from the gradient study, in which the depth distribution of eelgrass on the west coast was strongly related to Secchi depth, and by studies from other areas (review in Krause-Jensen et al. 2008). Eelgrass is a good example of a common and conspicuous species that has a relatively high abundance even at its deepest occurrence. This means that it is possible to follow the species' depth limit and register multiple observations of the deepest-occurring specimens per site, which strongly decreases the uncertainty compared with searching for scattered specimens in a narrow transect that crosses the depth limit.

Vegetation cover increases with declining nutrient concentration

Analyses of existing data indicated that on hard substrate, vegetation cover at a specific depth increased significantly along gradients of declining nutrient concentration and increasing water clarity. Empirical modelling accounting for diver effects, sampling variability in time and space, salinity gradients, wave exposure, and latitude explained 79% of the variation in

Figure 4.3: Eelgrass (Zostera marina) outside Lysekil on the Swedish west coast. *Photo: Mats Blomqvist.*

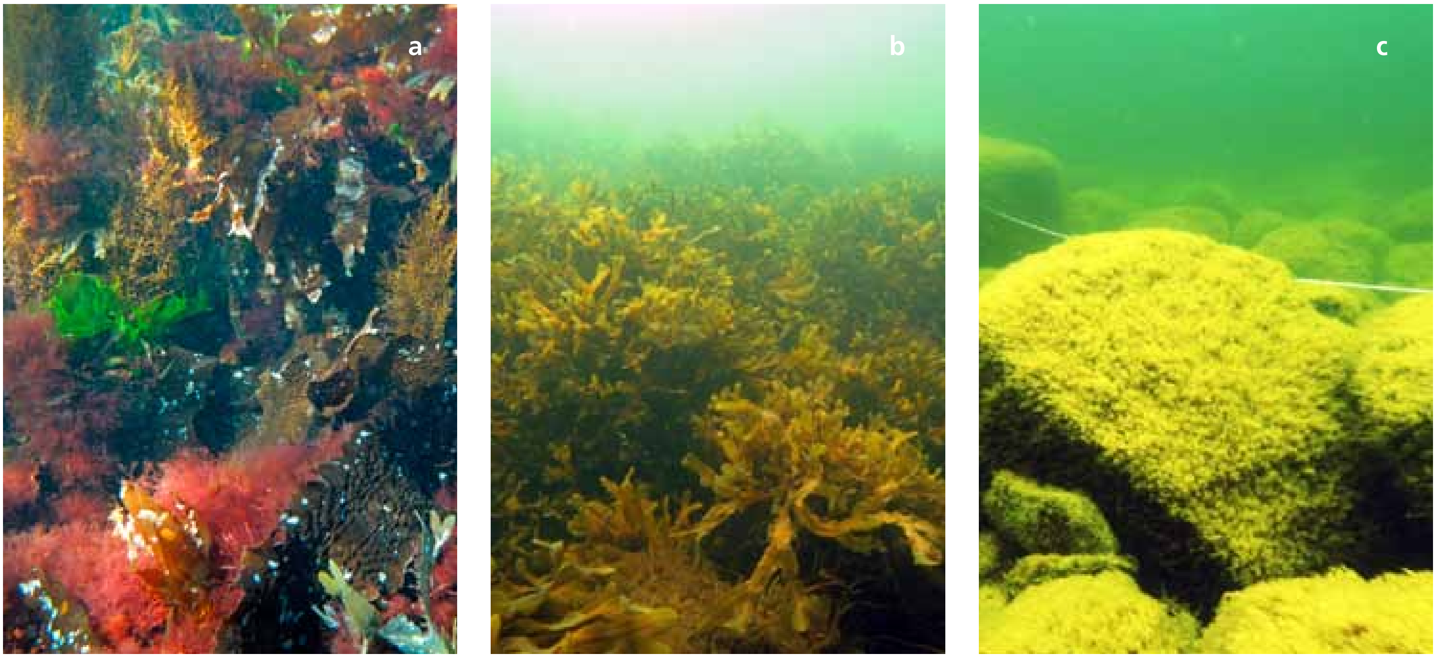
algal cover across 130 water bodies. A parallel analysis of macrophytes living on soft substrates revealed a similar pattern: the cover at a specific depth increased along gradients of declining total nitrogen concentration when accounting for differences in salinity. However, in this case we could explain only 52% of the variation in cover, probably because soft-substrate vegetation is more spatially heterogeneous.

The general response, particularly of hard-substrate vegetation cover, to gradients of eutrophication across wide ranges of environmental settings makes it a promising indicator of ecological quality useful for monitoring and managing marine vegetation in areas with strong environmental gradients (Blomqvist et al. 2014; Wikström et al. 2016). Cover of soft-substrate vegetation is more variable on small spatial scales, which means that a large sampling effort is required for monitoring in order to detect changes.

Effects of eutrophication and salinity difficult to separate

The species richness of hard-substrate vegetation (i.e. macroalgae and aquatic mosses) responded to nutrient concentrations and Secchi depth when accounting for natural gradients of salinity and physical exposure and normalising for sampling effort (i.e. size of sampled area). Also in the gradient study, species diversity displayed a strong response to coastal gradients, though it was somewhat difficult to separate the effects of eutrophication and salinity. This implies that the species richness of hard-substrate vegetation could be used as an indicator of ecological status, except in the Bothnian Bay where species richness is overall low, if it is possible to account for the strong effect of salinity (Figure 4.4).

Figure 4.4: Seaweed communities on hard substrate change strongly along the salinity gradient from the Skagerrak and Kattegat (a, *photo: Mats Blomqvist*) to the southern (b, *photo: Susanne Qvarfordt*) and northern Bothnian Sea (c, *photo: Susanne Qvarfordt).*



The response of the species richness of soft-substrate vegetation was tested only in the gradient study. In the otherwise species-rich communities on the east coast, species richness was low in the most eutrophic inner areas. This impoverishment was driven by a loss of sensitive species from these areas. This indicates that species richness may be a possible indicator for soft-substrate vegetation, if the effect of salinity can be taken into account (Figure 4.5).

Figure 4.5: Diversity could be a useful indicator in the mixed meadows of vascular plants, charophytes, and macroalgae on soft substrate in the brackish Baltic Sea, but the large spatial and temporal variability in these communities is a challenge for monitoring. *Photo: Joakim Hansen.*

Proportion of opportunistic and late-successional algae did not respond as predicted in the Baltic Sea

The proportion of opportunistic algae did not display any strong relationship to eutrophication when analysed across the Baltic Sea, and we were only able to explain a limited part of the indicator variability. Salinity is known to affect species composition, leaving mainly opportunistic species in areas with low salinities (Nielsen et al. 1995), which makes it difficult to detect any eutrophication effect. In the gradient study, the proportions of both opportunistic and late-successional algae responded as predicted in the coastal gradient on the west coast, but not on the east coast. This suggests that the functional composition of macroalgal communities, used as an indicator in other marine areas (e.g. Orfanidis et al. 2003; Juanes et al. 2008), is difficult to use in the Baltic Sea gradient.

Weak pattern of correlation between traits and eutrophication gradients

We conducted a comprehensive review of the traits of soft-substrate macrophytes based on a literature survey (Blomqvist et al. 2012a), and coupled it to the species registrations in the database. We found that some traits correlated with gradients of eutrophication along the Swedish coast, but that the pattern was relatively weak, probably again due to interaction with the effects of salinity. It was evident that a trait-based indicator for soft-substrate macrophytes is relevant only in the most brackish and sheltered areas, where the species diversity is large enough to include many trait combinations. Before drawing any conclusions regarding the use of traits as indicators of ecological status in brackish areas, however, we suggest repeating the analyses in narrower salinity ranges where the potential for identifying the responses of traits to eutrophication would be strongest (Blomqvist et al. 2014). On the Swedish west coast, where the diversity of soft-substrate macrophytes is low and eelgrass dominates, the distribution and abundance of eelgrass, and possibly the abundance of eelgrass relative to opportunistic macroalgae, are better indicators. Such seagrass indicators are already in use for the WFD in several European countries, including more saline areas in the Baltic Sea (Marbà et al. 2013).

Macrophyte sensitivity index also in deeper waters

The macrophyte sensitivity index, proposed by Hansen (2012), is an indicator of the ecological quality of vegetation in shallow coastal bays of the Baltic Sea. The index is based on the classification of soft-substrate species as either sensitive to or tolerant of pressures. In the east coast gradient study, this index displayed a significant relationship to nutrient concentrations, which suggests that the index could be used as an indicator for soft substrates in larger and deeper water bodies as well (Wikström et al. 2016).

Community complexity may be a useful indicator on hard substrate

The cumulative cover of vegetation is the summed cover of all species in the sampling plot, which means that it is affected by both the total area covered by vegetation and the number of vegetation layers. Dividing the cumulative cover by the total cover of vegetation gives a more direct measure of whether the vegetation is multi-layered, i.e. a measure of community complexity. Calculated in this way, community complexity can be expected to reflect species richness as well as the presence of large canopy-forming species, as these increase the cumulative cover by both adding a canopy layer and providing substrate for epiphytes, further increasing the cumulative cover. The indicator could only be tested in the gradient studies, since the existing database contained few data on total cover. The indicator performed better than did the proportion of opportunistic and late-successional vegetation on hard substrate in the east coast gradient, but displayed no relationship with the gradient on soft substrate. It may be a useful indicator on hard substrate, but since it was correlated with both cumulative cover and species richness in the dataset, it is probably redundant to include all three in the assessment criteria.

Field methods used are essential for choice of indicator

The field methods used for sampling coastal vegetation data determine the indicators that can be calculated, the uncertainty of these indicators, and hence the assessments that can be performed. The most common field method for monitoring coastal macrophytes in Sweden is diving transects extending from the depth limit of vegetation to the shoreline (Kautsky 1992).

The substrate cover and macrophyte taxa are recorded in more or less homogenous segments, which can differ in length and span different depth intervals. This is not optimal for measuring depth limits, since each transect gives only one value that can be highly uncertain, particularly for species for which the depth limit is defined by single, isolated specimens. As species richness depends on the size of the sampled area, the variable size of transect segments is problematic for this indicator. In addition, our analyses of existing monitoring data indicated large variability between substrates, sites, sampling occasions, and divers, particularly for the soft-substrate indicators (Blomqvist et al. 2014).

This means that new or improved field methods are required for better assessment of the ecological status of coastal vegetation. We have described and tested a video method for monitoring the depth limit (and possibly also cover) of eelgrass on soft substrate on the Swedish west coast (Blomqvist et al. 2012b; Wikström et al. 2016) and a SCUBA method with a fixed size of sampling area (Nyström Sandman et al. 2016). In the latter method, we also attempted to reduce spatial and methodological variability by restricting sampling to either hard or soft substrate and by using lists of which taxa to include and to what taxonomic level they should be identified.

The video method consisted of transects perpendicular to the depth curves, extending from a depth of about 1 m to the deepest growing eelgrass shoot. The transects ended with a zigzagging stretch parallel to the depth curves in order to obtain an additional 7 – 10 replicate observations of the deepest part of the eelgrass meadow. The camera was mounted on a sledge pulled after a boat. A similar method is used to monitor seagrass in Germany (Fürhaupter and Meyer 2009).

In the SCUBA method (Figure 4.6), the vegetation was surveyed in sampling squares placed on either dominant hard (i.e. rock, boulders, and stones) or soft (i.e. sand, clay, and mud) substrate. The vegetation cover was recorded in relation to the cover of the dominant substrate (i.e. hard or soft), using a predetermined taxonomic resolution for species determination to reduce the variation between divers.

The results from the method studies indicate that using a fixed-size sampling area, substrate-specific sampling, and a defined taxonomic level together reduced data variation in the proposed hard-substrate indicators compared with the current field method. In contrast, we were unable substantially to reduce the variation in the soft-substrate indicators. This likely reflects a large spatial variability in soft-substrate vegetation cover, together with difficulties in estimating the cover of species composed of long, slim stalks without leaf canopy.

Uncertainty analyses using the data from the tested field methods indicated that the temporal variation (between years) was generally smaller than the spatial variation (within and between sites in a water body). This indicates that to decrease the uncertainty in the status assessment, it is more important to increase the number of sites or samples than to sample each year during a six-year cycle.

Figure 4.6: A diver surveying a 5 × 5-m square of hard substrate in Marstrandsfjorden on the west coast of Sweden. *Photo: David Börjesson.*

SUGGESTED REVISIONS OF ASSESSMENT METHODS

We suggest that monitoring is stratified to either hard or soft substrate

The current macrophyte indicator MSMDI is difficult to monitor in many coastal areas and displays a weak response to eutrophication. We therefore suggest that new assessment criteria to be developed based on a number of new indicators. We also suggest suitable field methods to monitor these indicators.

We suggest that the monitoring of coastal vegetation be stratified to either hard or soft substrate to reduce variation and we propose using different indicators for the different substrates.

For vegetation on hard substrates, our studies identified cumulative cover and species richness as the indicators displaying the clearest response to eutrophication. These indicators have the additional advantage of also being applicable to shallow water bodies because they do not require the identification of species' depth limits. Species richness is probably not a useful indicator in the Bothnian Bay where species richness is overall low due to low salinity.

For soft-substrate vegetation on the Swedish west coast, the depth limit of eelgrass was identified as the most promising indicator to be included in a revised Swedish assessment method. Eelgrass cover, as well as eelgrass cover relative to the cover of opportunists, can also be obtained using the video method tested in the project. These may form supplementary indicators of these communities.

For the species-rich soft-substrate vegetation in the Baltic Sea and Gulf of Bothnia, cumulative cover at a specific depth together with species richness or the macrophyte index could be used as indicators. However, our results indicate that these communities are highly variable on small spatial scales, which means that a large sampling effort is required for monitoring in order to assess status with sufficient precision. The monitoring of soft-substrate vegetation will still be important in order to assess the status in areas dominated by soft substrate and in areas with very low salinity and impoverished hard-substrate vegetation.

For all indicators, our analyses highlighted the need to take salinity as well as other environmental variables into account when interpreting the ecological status.

SCUBA diving is not always the best method

The most common field method for monitoring coastal macrophytes in Sweden is SCUBA diving transects. The method is not well suited for monitoring species richness and is not optimal for registering the depth distribution limits of species. We therefore suggest new or modified monitoring methods for the suggested indicators.

The depth distribution of eelgrass can be monitored using underwater video transects of seagrass meadows, with multiple depth registrations along each transect. The method is described by Blomqvist et al. (2012b) and Wikström et al. (2016).

Monitoring species richness requires SCUBA diving and fixed-size sampling areas, for instance squares. Furthermore, our field studies suggest that a fixed-size sampling area can reduce the uncertainty of other indicators. The cumulative cover indicators also require SCUBA diving. In addition, they require data from a range of depths, including depths where the cover is regulated by light availability rather than physical exposure, which can be used to model the cover at certain depths. This implies using a method incorporating the SCUBA inventory of squares at different depths within a water body. Further analyses are needed in order to devise an optimal sampling design, for example, to determine the size of the sampling squares and whether the squares can be placed along a depth gradient within one or a few sites in a water body.

The large uncertainty in the cover-based indicators on soft substrate is problematic. We suggest evaluating whether the variation between divers could be reduced by providing a more detailed method description. One aid for cover estimations could be reference photos showing vegetation with different defined coverages, particularly including species for which cover is difficult to estimate.

REMAINING CHALLENGES

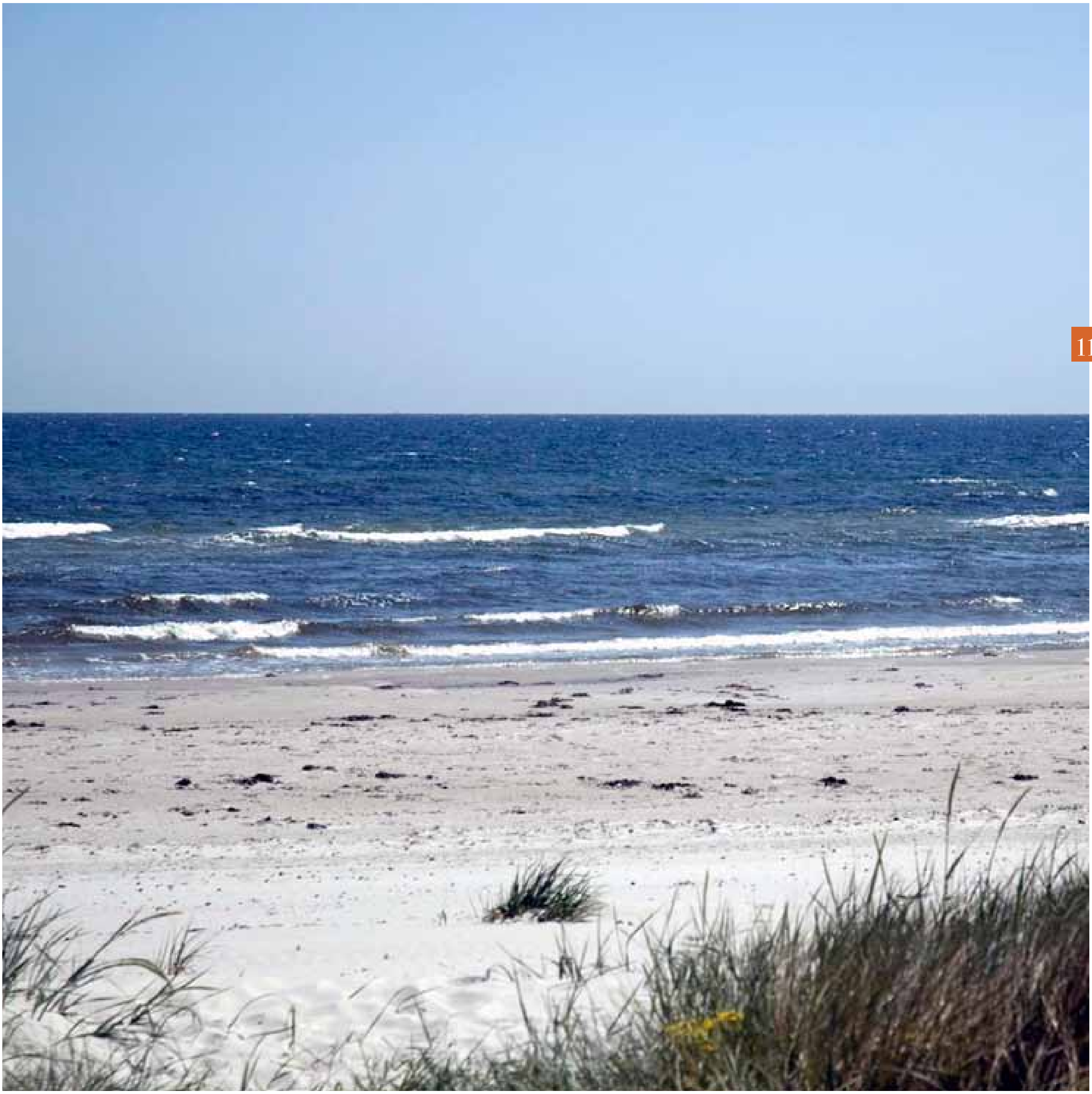
Methods to account for natural variation and for defining reference conditions needed

To implement the suggested indicators, one important task is to develop a method to account for variation due to natural gradients, most importantly salinity but also wave exposure. Some of this variation is captured by the water types, but both salinity and wave exposure vary within water types and, particularly for wave exposure, within single water bodies. We therefore suggest that the reference value, and class boundaries, be adjusted for wave exposure at the sampling sites and for the salinity of the water body. This can be done using empirical models of the responses of the indicators to pressures and natural gradients. For some of the indicators (e.g. cumulative cover on hard and soft substrates), such empirical models are already available in Blomqvist et al. (2014) based on a large dataset from the entire Swedish coast. For the remaining indicators, such models have to be developed when appropriate field data are available.

A second prerequisite for full implementation of the candidate indicators is that reference levels and class boundaries be defined. We see two possible ways to do this. One way is to model the current relationships between the indicators and Secchi depth and to use historic Secchi depth data to set reference values, as is done for the phytoplankton indicator. This would be a simple solution, since reference values for Secchi depth are already available for all water bodies. It would also ensure that the assessment criteria for phytoplankton and macrophytes are harmonised in coastal areas. We have demonstrated that all the suggested indicators display a clear relationship to Secchi depth in Swedish coastal areas, when comparing areas of different Secchi depths. This suggests that modelling reference levels from Secchi depth could be a useful method, although it would be even better to model the relationship over time. This should ideally be done in the future, when longer time series are available.

An alternative approach is to use a reference filter as is done for the BQEs in freshwater. Since minimally disturbed sites are difficult to find in Swedish coastal areas, an alternative approach is to delineate areas that are in "at least good status" and to compare those with sites with known impacts, as is done for the benthic fauna indicator. This could be done for cumulative cover on hard and soft substrates using the large vegetation dataset assembled in WATERS. However, this requires substantial work to define and choose sites to be used as references and sites where the vegetation is expected to be impacted by human pressures. This is complicated by the fact that high nutrient discharge to coastal areas often coincides with large freshwater input. To separate the effects of nutrients from the effects of salinity, we need areas with high freshwater input but low nutrient input and areas with little freshwater input but high nutrient concentrations, which are difficult to find in Swedish coastal areas.

Finally, the individual vegetation indicators have to be combined to give a common vegetation-based assessment of ecological quality. For this purpose, a combination rule is needed and we suggest using the method for integrated assessment developed in WATERS (see Chapter 5.2).



4.4. Benthic macrofauna in coastal waters

OVERVIEW OF CURRENT INDICATORS

Sediment-dwelling animals mirror long-term environmental conditions

Assessing environmental impacts in coastal and marine waters can effectively be done by analysing the composition of benthic communities. Sediment-dwelling animals are fairly stationary and live for months to several years; they therefore integrate long-term environmental conditions of particular places. Through evolution, benthic species have adjusted to cope with variable environmental conditions, but a pronounced impact will induce changes in the community composition. In general, diversity will decline with increasing stress, and the relationships between magnitude of disturbance and temporal and spatial changes in the benthic faunal composition (e.g. species numbers, abundance, and biomass) can be predicted according to the Pearson and Rosenberg (1978) model and paradigm. These successional changes in the benthic communities along the pressure gradient form the basis for most benthic habitat quality assessments worldwide (Figure 4.7).

BQI has influenced the first Swedish status classification according to the WFD

Based on this model, the Benthic Quality Index (BQI) was developed by Rosenberg et al. (2004) to assess the ecological status of coastal water bodies (Table 4.3). The initial index was later refined by adding an abundance factor together with descriptions of the procedures for developing sensitivity values and a framework for assessment based on BQI values from multiple sites (Leonardsson et al. 2009).

The sensitivity values are based on ES50, the estimated number of species among 50 individuals as interpolated from the rarefaction method of Hurlbert (1971) (on the west coast) or from the literature/expert judgement (on the east coast). Taxa lacking a sensitivity value are excluded from the sensitivity factor but included in the total number of species and individuals considered when calculating BQI. BQI is generally well established and has had a great influence on the first Swedish status classification according to the WFD. BQI was also recently introduced as an indicator of good environmental status of coastal and marine waters for several descriptors within the framework of implementing the MSFD in Sweden (SwAM 2012).

New possibilities for improvement of the current method

Over time, a lot of new information about faunal distributions has been gained, new assessments have been conducted, and some issues with BQI have been revealed. For example, evidence has been recorded that the occasional strong recruitment of a particular species can considerably reduce the sensitivity value. This leads to inaccuracy in the quality assessment, as has been noted by Labrune et al. (2006) and Grémare et al. (2009). Another problem has been spatial variation in the BQI values, even within a single water body. The existing method to deal with some of this variation has been to apply different WFD status class boundaries in shallow and deeper waters on the Swedish west coast (Rosenberg et al. 2004;

Figure 4.7: Model of the faunal successional stages along a gradient of increasing disturbance from left to right (after Pearson and Rosenberg 1978). Sediment profile images (colours enhanced) are shown on the top, where brownish colours in the sediment indicate oxidised conditions and dark grey reduced conditions. The general conditions in terms of number of species (S), species abundance (A), and biomass (B) along the gradient are shown in the lower panel.

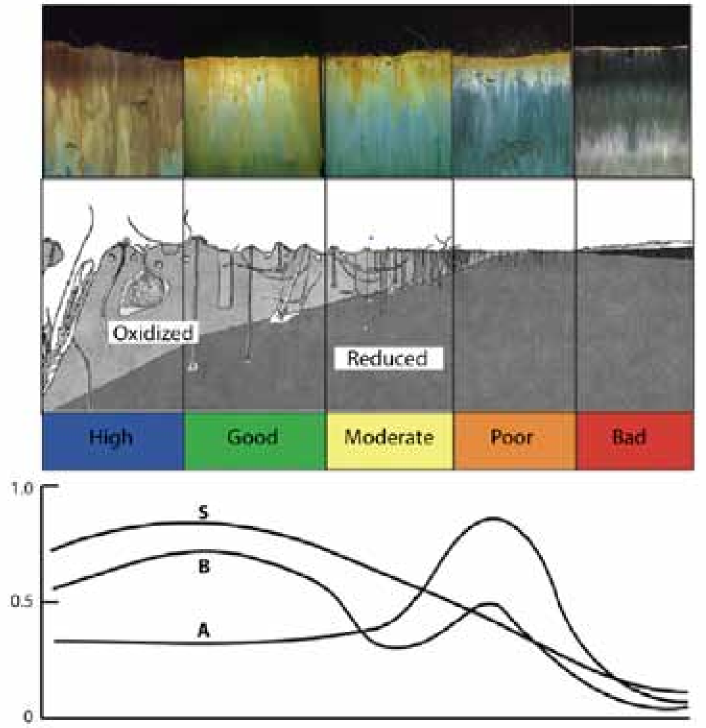


Table 4.3. Benthic macrofauna indicators used in assessing the ecological quality of coastal waters in Sweden and changes suggested by WATERS.

Coastal area	Indicator	Acronym	Suggested revisions and comments
Kattegat and Skagerrak	Benthic Quality Index[1, 2] (West coast)	BQI_{2015}	New methods for calculating the sensitivity value are suggested: remove the abundance factor from the BQI calculation; adjust for the depth, thereby eliminating the earlier need to have two sets of WFD boundaries for above and below a depth of 20 m.
Baltic Sea and Gulf of Bothnia	Probability-based Benthic Quality Index[3] (East coast)	pBQI	New index for the east coast where pBQI is based on the mean of four probability-based components (i.e. number of species, number of individuals, biomass, and abundance-weighted mean species-sensitivity values) from each sample. The probability of each component is calculated in relation to the baseline suggestion from national or regional reference areas. More thorough analyses are needed of how pBQI responds to anthropogenic pressure.

1) Leonardsson et al. (2015), 2) Leonardsson et al. (2016), 3) Blomqvist and Leonardsson (2016).

Leonardsson et al. 2009). An improvement of the index to explicitly account for variation due to environmental factors such as salinity would reduce the sampling effort needed for accurate boundary setting and assessment. Our objectives have therefore been to develop more knowledge of species sensitivity values and of how to account for spatial variability in BQI.

OVERVIEW OF INDICATOR DEVELOPMENT

Refining indicators for the Swedish east and west coasts

To refine indicators, different approaches were used for the west and east coasts of Sweden due to great differences in the number of species and in salinity. On the west coast, one issue was the calculation of sensitivity values, where the results can depend on the proportion of samples from disturbed and undisturbed environments. We have therefore developed a method to find the optimum mixture of the distribution of samples from disturbed and undisturbed areas for calculating species sensitivity values. In addition, we have gone from using ES50 values (i.e. the estimated number of species among 50 individuals) when calculating the sensitivity value to using the observed number of species in each sample (Leonardsson et al. 2015). Efforts were also made to reduce spatial variation by taking environmental factors into account. As depth alone explained much of the variation, using a regression model to account for the depth variation reduced the uncertainties. We also developed a method to establish the boundary between good and moderate ecological status within the WFD (Leonardsson et al. 2016). For the east coast, we have developed a new way to combine the components of an indicator based on converting the components to probabilities, producing a new index called the probability-based BQI (pBQI).

In our work on refining the BQI indicator for the Swedish coast, benthic fauna data were obtained from the Skagerrak and Kattegat areas from 1965 to 2013 and from coastal areas in the Bothnian Bay and Northern Baltic Proper from 1995 to 2015. The analysed benthic communities ranged from the phase of increasing environmental degradation during early successional stages in disturbed environments, to more mature successional stages in comparatively undisturbed conditions. Fishing pressure was not included in our analysis as we currently lack suitable data on this pressure and its effects. Our results are discussed in more detail below.

A modified method for calculating sensitivity values improves accuracy

BQI (version 2009) is calculated based on three factors, i.e. sensitivity, species abundance, and species richness, where the sensitivity factor has the greatest influence on BQI. The species sensitivity value is commonly used in various indices for assessing the marine environment. It is of great importance since it indicates species' tolerance of disturbance as well as their ability to coexist with other species. A high sensitivity value means that a particular species mainly occurs in communities of high diversity and is seldom found in disturbed areas. A low sensitivity value indicates, on the other hand, that the species is mainly found in species-poor and often disturbed areas. One of the most challenging aspects of benthic quality indices (e.g. AMBI: Borja et al. 2000; BQI: Rosenberg et al. 2004) has been compiling reliable measures of species' sensitivity to and tolerance of various magnitudes and kinds of disturbances. Marine benthic fauna encompasses thousands of species, most of which occur at low densities.

Scientific knowledge of the ecology of many species is limited, which makes it difficult to assign sensitivity values to many species based on documented knowledge.

For the west coast, two potential improvements to the calculation of the species sensitivity values were tested in Leonardsson et al. (2015). First, we changed the basis of the sensitivity value from expected number of species among 50 individuals (ES50) to the observed number of species (S) since the ES50 approach could be problematic giving unexpected low values in samples with many species in which one or two species were very dominant. Second, we developed a method for objectively calculating species sensitivity values to be more accurate and robust against how the underlying samples are distributed between disturbed and undisturbed areas.

Assessment precision increased by taking environmental factors into account

Despite the improvement made by calculating the sensitivity values in a new way for the west coast, there are still considerable discrepancies in the BQI values due to spatial variation. To reduce uncertainty in the assessment due to this spatial variation, the contributing factors need to be identified and explained. We used a regression model approach to find environmental variables that will reduce the variation in the index. A regression model that successfully takes the environmental variables into account would simplify sampling design and assessment since there would be no need to divide the water types into subtypes. In our new approach, the final assessment will not be based on the BQI values per se, but on the residuals from the regression model. By using the residuals we can clearly see how different environmental factors help to improve assessment precision (see section 5.3.2 for a general description of how status assessment can be done using regression models). Three environmental variables were included in the analysis: depth, salinity, and sediment type. All these variables have the potential to directly or indirectly affect benthic community structure. We demonstrated that depth can be used as a proxy for both salinity and sediment type, since both these variables are related to depth (Leonardsson et al. 2016). Depth is the best proxy possibly because depth covaries with salinity to depths of 20 – 30 m and because depth could be associated with food availability. Food production and distribution are among the primary factors affecting the composition and biomass of benthic communities (Pearson and Rosenberg 1987).

We also analysed the relationship between different factors in the BQI and depth to further reduce the variability. It was clear that there were obvious relationships between the different factors and depth, with the exception of the abundance factor, i.e. N/(N+5), which we therefore suggest be removed from the index. According to the regression model between depth and BQI, the predicted BQI values were higher for samples from deep areas than from shallow areas (Leonardsson et al. 2016). However, the residuals could differ between shallow and deep bottoms, which is why transformation to obtain the same range of residuals independent of depth is necessary. Depth-adjusted values would considerably reduce uncertainty in the assessment compared with unadjusted values. By using the residuals from a regression model between BQI and depth, the variance in BQI between samples was reduced by 50 – 75% in most situations. Another major improvement made by the depth adjustment is that we no longer need to have two sets of WFD boundaries, above and below a depth of 20 m, within each water body. This will significantly facilitate the design of monitoring programmes and the assessment of status (Leonardsson et al. 2016). We also suggest a method to define the WFD good-moderate (GM) ecological status boundary and the MSFD boundary for open-sea good environmental status, by using transformed residuals from undisturbed areas.

We suggest a probability-based index for assessment in the Baltic Sea

While BQI seems to function well in species-rich environments such as the Swedish west coast, it has turned out to be sensitive to abundance variation in areas of low diversity, such as the Baltic Sea. Another issue with BQI in the Baltic Sea has been the poor status classification of the unpolluted northernmost part of the Bothnian Bay. For these reasons, there was a need to revise the BQI for the Baltic.

To reduce variation and remove as much as possible of the variation due to depth, salinity, and sediment type, we reanalysed the components of BQI for the Baltic Sea. We also included biomass as a potential candidate for the index, since it has been successfully used as the basis for a benthic indicator in eastern parts of the Baltic Sea (Lauringson et al. 2012). Furthermore, we investigated the potential of increasing the taxonomic resolution of chironomids, which turned out not to improve any of the indices substantially.

The main component of BQI is the sensitivity factor, and including more components in the index for the east coast may be one way of stabilising the index. However, it is not obvious how to combine different benthic fauna variables such as biomass, abundance, and species richness into a single index, since they have different measurement scales. One of the difficulties is how to weight the different variables in relation to each other. A way to solve this problem is to convert the fauna variables to probabilities. A low probability for any of the variables means that few observations with lower values than the observed are present in the baseline data. If the probabilities for all the variables are low in a single sample, then the average will be low and a sample with such a composition is unlikely to be classified as coming from an environment with an ecological status comparable to that of the baseline data. Hence, unlike the original BQI, there is an intuitive interpretation of the value of this new probability-based index, pBQI. The assumption made here is that a deteriorated environment will generally result in low values (Blomqvist and Leonardsson 2016).

The most essential part of a well-functioning pBQI is a baseline dataset that is comparable in environmental terms to those samples to be compared with the baseline data. Providing that proper baseline data are available for assessing a given water body, the assessment will indicate how likely it is that the ecological status is similar to that captured by the baseline data. For this to work properly, the baseline dataset needs to reflect a benthic community that is expected for the type of environment from which the samples originate.

When comparing BQI with pBQI, many of the results of our study speak in favour of pBQI. As an assessment method, pBQI is much more general than BQI since pBQI can be directly applied to other quality factors as long as there are plenty of data to form appropriate baseline data (Blomqvist and Leonardsson 2016). However, an important aspect of a useful index is that it should respond to anthropogenic pressure. This is so far the weakest aspect of the indices, as no independent data on anthropogenic pressures are available for testing. The analyses of the pressure-responses made for this report were based on the assumption that the anthropogenic pressure decreases from the inner to the outer parts of recipient spatial gradients.

SUGGESTED REVISIONS OF ASSESSMENT METHODS

For marine coastal benthic fauna, we have presented new indicators, new methods for their calculation, and a method for setting the boundaries between different environmental classifications. The species' sensitivity values are used to calculate the sensitivity factor, a crucial component of most WFD benthic quality indices, and this component contributes the most to the quality assessment. We suggest that future assessments of communities of benthic macrofauna in the Skagerrak and Kattegat follow the methods described in Leonardsson et al. (2015) and Leonardsson et al. (2016). This means changing how to calculate the species sensitivity value by changing the base in the Swedish index to the observed number of species, rather than the previously used expected number of species, and by removing the abundance factor. It is also important to use sensitivity values derived from the suggested stratification method. New suggested sensitivity values can be found as supplementary data to Leonardsson et al. (2015).

The new, updated BQI formula for the west coast is the same as the original formula of Rosenberg et al. (2004), but with sensitivity values based on species richness according to Leonardsson et al. (2015) instead of ES50 and with the abundance factor removed (Leonardsson et al. 2016).

To improve the status assessment without stratification by habitat/subtype, we suggest a method to reduce spatial variation in the benthic quality index, using a regression model to account for the spatial variation by adjusting for depth. We also propose a method to establish the boundary between good and moderate status and for deriving EQR values according to the WFD. The ultimate goal of all indices in European coastal waters is, with as high precision as possible, to assess the good/moderate boundary.

For the Baltic Sea we have developed a new index, pBQI, which is suggested as an alternative to the traditional BQI, over which it is believed to have several advantages. In this case, pBQI is the mean probability of finding a certain species richness, sensitivity, composition, biomass, and total abundance within a sample. However, pBQI has been tested in only a few areas and we do not know how it responds to anthropogenic pressure, as no independent data on the benthic pressure load are currently available for testing the indicator.

REMAINING CHALLENGES

Assessment methods need to be adaptive as new data and knowledge becomes available

To implement the suggested changes in the interest of improved assessment, and before new class boundaries can be suggested, we need new data with improved spatial coverage from the upcoming national monitoring programmes. Long-term climatic trends are also likely to affect the faunal communities, and such impacts need to be reflected in the assessment boundaries since a return to the original state may no longer be possible (Duarte et al. 2009). Assessment work therefore needs to be adaptive in the sense that indicators should be improved as new knowledge is obtained and in that class boundaries should be adjusted over time.

For the Baltic Sea, we need further testing and more knowledge of how the indices respond to anthropogenic pressures. This could not be done within WATERS due to lack of data on such pressures. One positive aspect of pBQI is that it is easy to include more components in the index. It could, for example, be of interest to evaluate the potential of supplementing the index with a size structure measure, which could be created by using the average size (i.e. biomass) of individuals in the sample (e.g. calculated from biomass divided by number of individuals).

4.5 Coastal fish

OVERVIEW OF CURRENT INDICATORS

Eutrophication and habitat deterioration are the strongest pressures on coastal fish

Several human-induced pressures affect coastal fish communities, including eutrophication, habitat deterioration, and fishing. Eutrophication may worsen the quality of recruitment and feeding areas for certain fish species, or alter competition among species in that some species are favoured at the expense of others (Sundblad et al. 2013). Habitat deterioration may directly affect the availability of suitable recruitment areas (Sundblad et al. 2014). In addition, many coastal fish species are targeted by commercial as well as recreational fisheries, which may affect their abundance and size structure (Karlsson 2014; ICES 2014). Fish may also be affected indirectly by these pressures, via effects on the food web (Östman et al. 2016). In addition, natural environmental factors such as temperature and salinity may strongly influence fish abundance and species composition (Olsson et al. 2012). Habitat-related factors, such as topography, wave exposure, and the availability of suitable recruitment habitats, may affect the natural capacity of an area to support fish populations.

The development of coastal fish indicators has been coordinated by HELCOM

The development of coastal fish indicators in Sweden has largely been channelled internationally via HELCOM and geographically focused on the Baltic Sea region (Ådjers et al. 2006; HELCOM 2006, 2012; Bergström et al. 2016). This work has aimed at achieving regionally agreed-on indicators and methodological standards for following up the goals of the Baltic Sea Action Plan (HELCOM 2007) and for reporting in relation to the MSFD (EC 2008; HELCOM 2013).

A comprehensive set of indicators of the environmental status of coastal fish was identified by HELCOM (2012), based on multivariate analyses and a set of formal selection criteria. These indicators were identified to cover the following key categories: species composition, size structure, trophic structure, and species diversity. These categories are also used in Sweden for presenting results of national and county-level assessments of coastal fish. The same indicators are used in all coastal areas to the extent justified by the local species composition; in other cases, indicators that are as comparable as possible are used. The results are presented in fact sheets, reporting trends over time without any assessment of status. (www.slu.se/faktablad-kustfisk).

Indicators corresponding to those of the Baltic Sea proposed for the North Sea

Three of the indicators proposed by HELCOM (2012) were recently agreed to be used in regional assessments of the status of biodiversity in the Baltic Sea ("Core" indicators, HELCOM 2013, 2015a). The Swedish implementation of the MSFD is to use the same indicators for the Baltic Sea marine region. Corresponding indicators are proposed for the North Sea marine region, which includes the Skagerrak, Kattegat, and Öresund (HVMS 2012; SwAM, 2012b). In addition, one indicator of size structure and one of trophic structure are proposed by

SwAM (2012b). However, the two national indicators and the proposed North Sea indicators require further development and evaluation before being operational. All these indicators are summarised in Table 4.4.

Table 4.4. Proposed Swedish indicators for the environmental status of coastal fish.

Indicator	Sampling period	Comments
Abundance of coastal fish key species (CPUE)	August	HELCOM Core indicator; the key species is perch (Perca fluviatilis) in the central and northern Baltic Sea, and flounder (Platichtys flesus) or cod (Gadus morhua) in the southern and western Baltic Sea
Abundance of piscivores (CPUE)	August	HELCOM Core indicator
Abundance of cyprinids (CPUE)	August	HELCOM Core indicator
Abundance of mesopredators (CPUE)	August	Proposed for assessment in areas where abundance of cyprinid fish is not applicable
Size structure of coastal fish communities	August	Proposed indicator, to be developed further
Trophic structure of coastal fish communities	August	Proposed indicator, to be developed further

Little effort so far to investigate responses to pressure gradients

Indicator development for fish has been initiated relatively recently, and so far little effort has been made to investigate the responses of indicators to natural and human-induced pressure gradients, or to their temporal and spatial variability. Other important gaps in Sweden are that the current coastal fish monitoring areas do not provide sufficient geographical coverage for MSFD reporting (Fredriksson 2014), and that indicators applicable to the more marine areas of the Öresund, Kattegat, and Skagerrak are less developed. WATERS has helped fill these knowledge gaps, also aiming to harmonize the development of the MSFD indicators and assessment protocols with the reporting requirements of the WFD.

We have particularly focused on studying how different indicators reflect human-induced changes and how they change along natural gradients. This information is important in order to support the geographical delineation of indicators, identifying boundaries for good environmental status and ultimately in order to identify suitable management measures where needed. In addition, the information will contribute to refining the assessment protocol by improving comparability among geographical areas and by increasing the number of areas that can be assessed. One part of the indicator development was to test the different fishing methods in use to be able to suggest the optimal ones for different coastal regions.

We focused on the same selection of indicators as proposed by SwAM (2012a) for reporting in relation to the MSFD. In addition, we have revisited some of the early indicators proposed by HELCOM (2012) to identify potential indicators reflecting aspects of species diversity. Our aim was to assess indicators representative of different functional groups and to cover a wider range of aspects.

OVERVIEW OF INDICATOR DEVELOPMENT

The salinity gradient along the Swedish coast strongly affects coastal fish

As with other organism groups, the decrease in salinity from the Skagerrak to the Bothnian Bay in the inner Baltic Sea strongly affects fish species composition. These differences affect what indicators may be applicable to different geographical areas. We studied how the species composition of coastal fish assemblages and their functional attributes change with salinity, based on data from fish surveys at 129 sites sampled during 1988 – 2011 (Karlsson et al., in manuscript).

Coastal fish assemblages consisted of a mix of freshwater and marine species in all areas of the Swedish coastline, but freshwater species were most common on the brackish Baltic Sea coast, whereas marine species dominated on the Swedish west coast (i.e. the Skagerrak, Kattegat, and Öresund). In parallel, there was a shift from mainly near-bottom species in the Skagerrak and Kattegat to species living mainly in the water column in the Baltic Sea sub-areas. The share of plankton feeders and omnivores also increased towards the inner Baltic Sea, away from being dominated by bottom-fauna feeders in more saline areas. In all areas, the coastal fish assemblages had a strong component of species with limited dispersal distances, indicating that all Swedish coastal fish assemblages are likely to respond to changes in local environmental factors as well as to environmental changes at larger geographical scales. The observed geographical variation in both species richness and functional traits provides guidance for the suitable geographical delineation of assessment units for indicators, and emphasises that good environmental status should be defined separately in the Skagerrak, Kattegat, and Öresund compared with the Baltic Sea coastal areas, in order to encompass differences in natural environmental preconditions.

Figure 4.8: Marine- and freshwater-dominated fish assemblages. *Photos: Martin Karlsson.*

Evaluation and improvement of methods for monitoring coastal fish

In response to the differences in species composition, different monitoring methods are used in different parts of the Swedish coastline. Coastal fish are monitored using gill nets in the Baltic Sea and fyke nets in the Skagerrak, Kattegat, and Öresund. These two methods sample slightly different parts of the fish assemblages. The fyke nets are more selective for near-bottom species since they are positioned near the sea floor, whereas gill nets extend up to around 2 m from the bottom.

We compared data from areas where fyke nets and gill nets had been used in parallel, to assess how well the two methods meet the needs of indicator-based status assessment (Bergström et al. 2013). Gill nets clearly performed better than did fyke nets in the Baltic Sea, in terms of the number of species sampled and in precision, and fyke nets generally performed poorly. On the Swedish west coast, the two methods estimated species number equally well if the values were standardised against the size of the catch. However, fyke nets required less time in the field to achieve a certain sampling precision. Fyke nets were also concluded to be more environmentally friendly, since the catch can be released live after being registered, which is important in sensitive coastal areas of the Swedish west coast. One additional factor that affects coastal fish monitoring in these areas is the shore crab, which is common from the Skagerrak to the Öresund. Considerably more time is required to handle shore crabs when these are caught in gill nets than in fyke nets, and they cause less damage to gear and catches in fyke nets.

The monitoring standards for test fishing in coastal areas have now been updated, based inter alia on experience from the WATERS gradient studies (see Section 4.3). For gill nets, the existing standard was updated with improved technical information (Karlsson 2015). For fyke nets, the WATERS field surveys followed a newly developed protocol aiming for improved spatial representation in the sampled areas (Bergström and Karlsson 2016). The new standard is compatible with the older standard in shallow areas, but covers a larger spatial area and a wider depth range in order to represent a wider part of the sampled coastal habitat. Like the monitoring standard for Nordic coastal multimesh gillnets in the Baltic Sea, it is based on randomised depth stratification and recommends using more stations than does the earlier fyke net standard, in order to improve monitoring precision. Although the earlier standard is still being used in ongoing monitoring programmes (Andersson 2015), the new standard (Bergström and Karlsson 2016) is recommended for any newly established monitoring programmes and for inventory studies.

Figure 4.9: Top: a typical coastal fish monitoring area in the Baltic Sea includes 45 stations sampled using Nordic coastal multimesh gillnets, spanning depths of 0–10 m based on randomised depth stratification (Karlsson 2015). Bottom: monitoring using fyke nets on the Swedish west coast is typically performed near shore at depths of 0–6 m at fewer stations, but the sampling design varies among areas (Andersson 2015). *Photos: Martin Karlsson.*

Natural environmental conditions must be taken into account

An important part of indicator development is to clarify to what extent the indicators reflect real changes in the coastal fish communities and whether they can identify changes due to human impact. We compared large-scale differences in species composition in a geographically extended dataset from the Swedish Baltic Sea coastline (from the northern Bothnian Bay to Bornholm Basin) and parallel changes in indicators (Bergström and Olsson 2015). The variability in indicators could largely be explained by changes in natural environmental variables, showing that it is important to take natural environmental conditions into account when identifying boundary values for good environmental status.

The indicators assessed in WATERS were the two HELCOM core indicators abundance of perch and abundance of cyprinid fish (Bergström et al., in manuscript 1). In addition, the proportion of large perch was included in order to assess an indicator of size structure. Generally, spatial variability was considerably higher than temporal variability. At the local scale, the strongest relationships were to water depth and water temperature during fishing, whereas wave exposure was relatively less important. This small-scale natural variability can be corrected for when comparing different areas with each other. The variability among monitoring areas was mainly related to differences in water transparency, which was used as a measure of eutrophication. In particular, the abundance of cyprinid fish increased with decreasing water transparency. Furthermore, the results indicated that commercial fishing did not seem to control the coastal fish populations in the dataset. The studied gradients may not have been long enough to permit detection of such responses, since the level of commercial fishing pressure was fairly similar in all studied areas. In addition, information on the geographical distribution of recreational fisheries was not available, although this is the main type of fishing in Swedish coastal areas. The relative effects of salinity and seasonal temperature had little influence on the variation in indicator values across areas at the range of the study (latitude 56 – 66oN and salinity 2 – 8).

The studies were based on data from reference areas for coastal fish monitoring, that is, areas not subject to direct anthropogenic influence, as well as from areas affected by human impacts ("disturbed") in order to evaluate the environmental gradients. Information from such disturbed areas is normally scarce in the national coastal fish database, which has a predominance of data from reference areas. The data from disturbed areas were mainly obtained from an extended sampling campaign on coastal fish conducted in 2011, which covered, for example, areas near commercial harbours, industries, urban areas, and agricultural land (Söderberg and Mattsson 2011), and from field sampling conducted as part of the WATERS gradient studies.

Eight potential and established fish community indicators tested in WATERS gradient studies

The WATERS gradient studies addressed one coastal gradient in the Skagerrak and one in the Baltic Sea, and were designed to compare the responses of indicators representing different biological quality elements (see Section 4.3). The data were evaluated in detail to assess the responses of various indicators (i.e. potential coastal fish status indicators of eutrophication; Pihl et al., in manuscript). Eight potential and established fish community indicators were tested, representing the following properties of the coastal fish community: fish abundance, species diversity, abundance within functional groups, size structure, and trophic level. The indicators were assessed focusing on aspects of eutrophication, and their importance relative to natural gradients remains to be explored. In the Skagerrak gradient, the indicators total abundance, species richness, abundance of piscivores, and abundance of mesopredators correlated with water transparency and nutrient concentrations, increasing in areas less affected by eutrophication. No corresponding relationships were seen for the size-based indicators or indicators of trophic level. In the Baltic Sea gradient, several indicators correlated with water transparency and chlorophyll-a. However, the indicators also correlated with salinity, which changed simultaneously with the eutrophication variables in the studied gradient. The indicators total abundance and abundance of cyprinid fish responded most strongly to the studied gradient in the Baltic Sea, both increasing in more eutrophic areas.

SUGGESTED REVISIONS OF ASSESSMENT METHODS

Abundance of cyprinid fish is a suitable indicator of eutrophication in the Baltic Sea

The results support the current national selection of indicators and indicate that these should be developed further. In addition, it is proposed that species richness could be used as a supporting indicator to estimate diversity where needed, but that results should be standardised against the abundance of the catch if they are to be compared among geographical areas.

Abundance of cyprinid fish was confirmed as a suitable indicator of eutrophication in coastal areas in the Baltic Sea.

A mesopredator indicator for habitat quality in the North Sea

The results also support the use of a mesopredatory fish indicator in the North Sea management unit (i.e. the Skagerrak, Kattegat, and Öresund). However, its response to environmental gradients is generally not similar to that of the corresponding indicator in the Baltic Sea (i.e. abundance of cyprinids, see above). The abundance of mesopredators in the Skagerrak correlated primarily with environmental variables attributable to good habitat quality in the studied gradients. In both sea regions, the indicators' responses to fishing pressure need to be explored further and be assessed in more pronounced pressure gradients than the ones available here (Bergström et al. 2016b).

Good environmental status should be defined differently along the Swedish coast

The studies of fish community composition along the Swedish coasts suggest that good environmental status should be defined separately in the Skagerrak, Kattegat, and Öresund compared with the Baltic Sea coastal areas, to encompass differences in natural environmental preconditions. Finally, we recommend that the new monitoring standards for fyke nets suggested by WATERS (Bergström and Karlsson 2016) be used in newly established monitoring programmes and in inventory studies.

REMAINING CHALLENGES

Data on human-derived pressures limit indicator development

The analyses of the indicators' responses to environmental pressures were restricted mainly by the available data, although the scope for such analyses was significantly enhanced by the extended sampling campaign in 2011 and by the WATERS gradient studies. Most importantly, the number of monitored areas was small, but in some respects pressure data were lacking as well. The most important gap in this respect was a lack of information on recreational fisheries, the main source of fishing mortality for many species of coastal fish in Sweden (Karlsson et al. 2014).

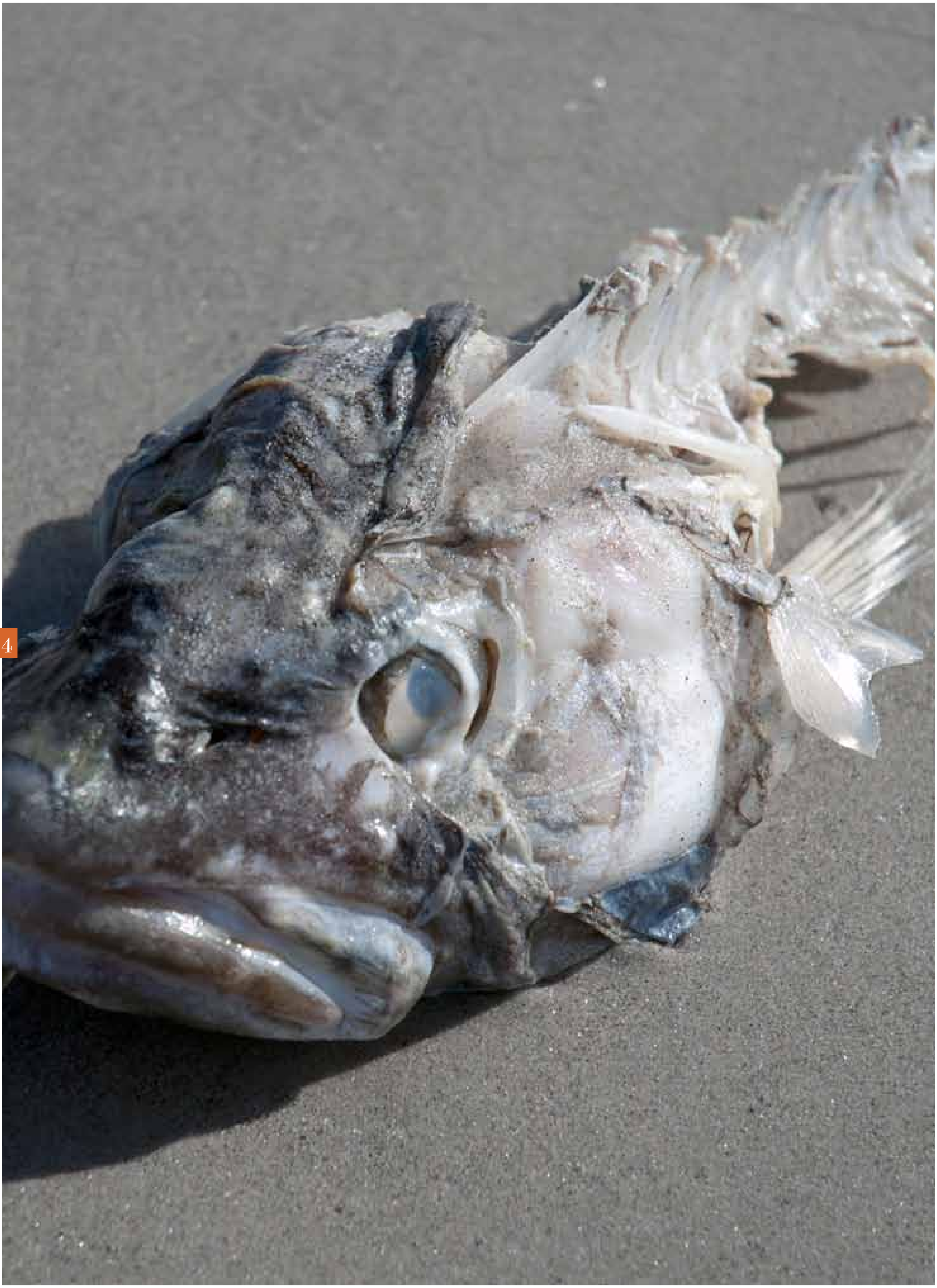
The scarcity of monitoring data also affects the confidence in the status assessments. Bergström and Olsson (2015) summarised the 2014 state of indicator-based status assessments of coastal fish communities in Sweden, focusing on potential connections between the MSFD and WFD in order to facilitate the harmonisation of assessments. As also reported by Fredriksson (2014), one main conclusion was that the current network of coastal fish monitoring in Sweden is too scarce to enable the representative and relevant monitoring of coastal fish in all required assessment units for the MSFD (and even more so for the WFD). This need is emphasised by strong differences in species composition among areas (Karlsson et al., in manuscript).

Method development needed to make increased use of inventory studies

The monitored coastal fish communities are represented by species with limited dispersal, which are potentially influenced by several environmental factors at both local and larger spatial scales. To encompass this variation, the currently applied assessment protocol states that threshold values for good environmental status should be defined locally using time-series-based approaches (HELCOM 2012b,c). However, this requires several years of data. To complement the time-series-based approaches, spatially based assessment approaches would enable better use of data from inventory studies and any newly initiated monitoring programmes. This conclusion is supported by analyses of the relative importance of temporal and spatial variation in the Baltic Sea, which indicate that temporal variation in the studied indicators was minor relative to spatial variation (Bergström et al., in manuscript 1). The study also illustrated how natural local variables can be corrected for in order to focus specifically on potential differences due to human impact. These results should preferably be developed further with the aim of proposing a spatially based assessment protocol to support status assessments in areas where time-series-based approaches are not yet possible.

Increased density of areas for long-term monitoring would improve assessment

One key improvement of the national coastal fish monitoring programme would be to increase the density of areas subject to coastal fish monitoring, as some coastal water types currently go completely unmonitored. In addition, options for how to harmonize existing coastal fish monitoring programmes on the Swedish west coast should be considered, to facilitate comparisons of assessment results. This effort could also be supported by results from a recent evaluation of national coastal fish monitoring (Leonardsson et al. 2016).



4.6 Integrated assessment and responses in coastal waters

During 2012 – 2013 we conducted a coastal gradient study in which all the coastal BQEs (i.e. phytoplankton, benthic flora, benthic fauna) along with fish were sampled in the same coastal eutrophication gradients. The study was performed in two geographic areas, i.e. the west and east coasts of Sweden. On the west coast, we sampled five fjord areas in one gradient inside the islands of Orust and Tjörn in the Skagerrak. On the east coast, we sampled seven archipelago areas comprising three gradients from inner to outer parts of the archipelago of Östergötland in the Baltic Sea.

Phytoplankton were sampled at depths of 0 – 10 m at one station per area on three occasions per year from June to August. Benthic macrofauna were sampled with a grab sampler at 10 – 15 stations in each area. Macrophytes were surveyed by scuba divers in 5 × 5-m squares at depths of 3 – 5 m at 10 stations per area, except for soft-bottom macrophytes in the west coast study area, which were surveyed using underwater video. Fish were sampled according to Swedish national standards for coastal fish monitoring (i.e. using fyke nets in the Skagerrak and Nordic coastal multi-mesh gill nets in the Baltic). In addition to the biological data, we also measured a number of physicochemical variables, including nutrient concentrations, Secchi depth, and salinity.

The Skagerrak: transparency the strongest gradient; salinity and nutrients highly variable over time

The aim of the study was to test the responses of the BQEs to gradients of anthropogenic eutrophication. However, the eutrophication gradients in coastal areas are typically intertwined with gradients of salinity, river-borne and resuspended particles, and humic substances. The different BQEs can be expected to respond to different components of such complex coastal gradients.

In the Skagerrak, the studied gradient was mainly formed by a strong decrease in water transparency (measured as Secchi depth) from the open sea to the innermost area, driven mainly by an increase in concentration of CDOM. Salinity and nutrient concentrations were highly variable over time in all study areas but did not display any clear differences between study areas. This gradient thus differs from the most typical eutrophication gradient, in which anthropogenic nutrient enrichment results in decreased Secchi depth due to increased phytoplankton production. This illustrates how the relationship between nutrients and Secchi depth is often complex in Swedish coastal waters due to the strong contribution of land-derived particulate matter and CDOM to the optical properties of the water.

The four BQEs differed greatly in their response to the water transparency gradient. Both tested macrophyte indicators (i.e. cover and species richness of macroalgae on hard substrate) displayed a strong correlation with the gradient. The relationship between vegetation cover at a certain depth and the Secchi depth (i.e. the depth penetration of light) has a strong theoretical basis and has been demonstrated in both Swedish and other coastal areas (Krause-Jensen et al. 2007, 2008; Blomqvist et al. 2014). In addition, species richness has been documented to be correlated with water quality variables, including Secchi depth (Blomqvist et al. 2014). The fish indicators also responded to changes in water transparency, with an increase in both species richness and mesopredator abundance. These results could be explained by the higher availability of vegetation, providing suitable habitats for coastal fish in areas with high water transparency in the Skagerrak gradient.

The benthic fauna indicator (BQI) displayed a different pattern, with low values in the two innermost areas (i.e. Byfjord and Havstensfjord) and high values in the other areas. Both Byfjord and Havstensfjord are known for their seasonal and almost permanent hypoxic basins with long residence times for the deep water. The response of the benthic fauna indicator in the inner areas is a response to low-oxygen conditions at the bottom, fuelled by eutrophication but primarily due to limited water exchange in combination with high organic loading.

In contrast, the tested phytoplankton indicators (i.e. chlorophyll-*a* or biomass) varied strongly within areas and displayed no correlation with the main gradient across the areas. The lack of response of the phytoplankton indicators to the main gradient is not surprising, given that this is mainly a gradient of Secchi depth driven by CDOM concentration. Phytoplankton chlorophyll-*a* and biomass here measured as organic carbon) respond mainly to nutrient concentrations, which did not differ clearly between the study areas.

The Baltic Sea: gradients of nutrients, Secchi depth, and salinity – different responses between BQEs

In the Baltic Sea, the study areas formed clear gradients of nutrient concentrations (mainly total nitrogen), Secchi depth, and salinity, where areas with high total nitrogen concentrations also had low salinity and vice versa. This means that we were unable to test the response of the BQEs to an isolated eutrophication gradient but that the results have to be interpreted as the response to a combined gradient of eutrophication and salinity.

In this complex gradient, the response also differed between the BQEs. Phytoplankton chlorophyll-a and biomass were very high in the area with highest nutrient concentrations and low salinity, indicating poor status using the current assessment criteria. The other areas differed relatively little in the phytoplankton indicators and were all assessed as of moderate status using the current assessment criteria.

The indicators for the other BQEs did not display the same pattern. As expected, the cumulative cover and species richness of vegetation and BQI were relatively low and the abundance of cyprinid fish was relatively high in the area with the highest nutrient concentrations, where phytoplankton chlorophyll-*a* and biomass were high, resulting in low water transparency. However, cyprinid abundance was equally high in some of the other areas. Likewise, vegetation cover and species richness as well as BQI were also low in other of the study areas. There were also large differences between the BQEs within single areas. For instance, one of the two areas with the highest BQI (indicating good ecological status for benthic fauna using the current assessment criteria) had very low cumulative cover and species richness of vegetation and high abundance of cyprinids. This area had low water transparency due to high concentrations of suspended material discharged from land, which could explain the responses of the vegetation and fish indicators. At the same time, the area was relatively open to the sea, with no shallow sills or narrow straits hindering water exchange with the open Baltic. This is likely to result in good oxygen conditions at the sea bed, favouring benthic fauna.

Different responses of indicators to the same coastal gradient

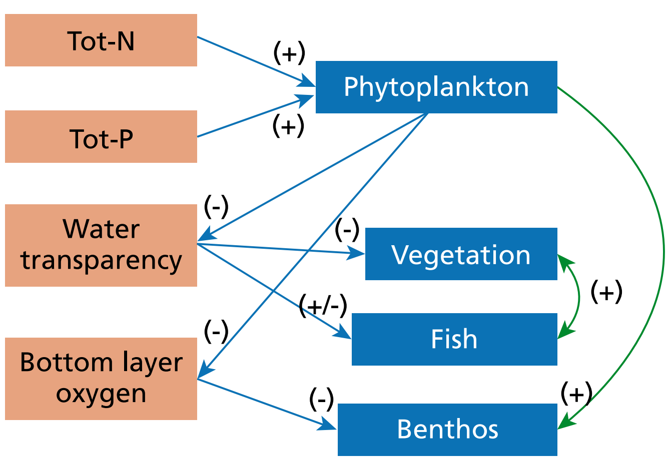
The results of the gradient studies illustrate the diverse responses of different indicators to the same coastal gradient. This is not surprising given that different organism groups occur in different habitats and respond to different factors in the same gradients.

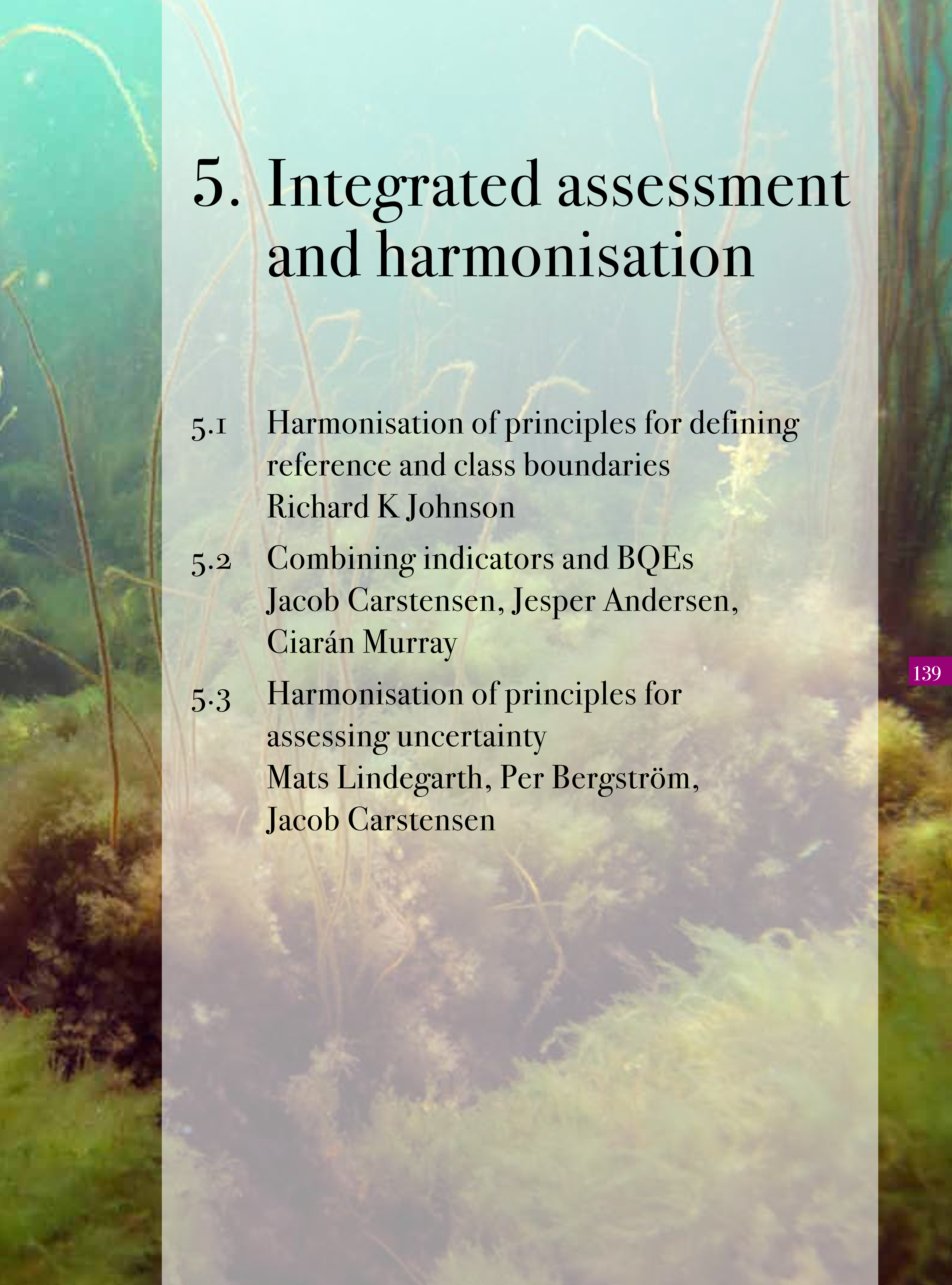
Coastal species respond to several environmental factors and biotic interactions, and some of the important regulating factors in coastal eutrophication gradients are shown in Figure 4.10. While phytoplankton respond directly to the concentrations of nutrients, fish and benthic fauna respond to the production of phytoplankton, which provide food for higher trophic levels but also decrease water transparency and can reduce oxygen concentrations in the bottom layer due to increased deposition of oxygen-consuming organic matter.

Macrophytes can respond directly to increased nutrient concentrations, but our results suggest that the indirect effect of nutrients via phytoplankton, which influence water transparency, has a stronger effect on both macrophyte cumulative cover and species richness in the range of nutrient concentrations that occur in Swedish coastal waters. Macrophyte vegetation and fish are also interlinked through biotic interactions that are not fully understood. Macrophytes constitute an important habitat for coastal fish, and studies demonstrate that the extent of suitable spawning habitat can affect fish populations (Sundblad et al. 2014). The density of large predatory fish, on the other hand, can affect macrophyte communities through cascading effects on small herbivores (Eriksson et al. 2009, 2011), although no studies have examined how this affects the indicators tested here.

An important explanation for the difference in response between benthic fauna, on one hand, and vegetation and fish, on the other, is that they respond to environmental change in different habitats within the same water body (deep aphotic versus shallow photic). High nutrient concentrations and phytoplankton production do not always result in poor oxygen conditions if the deep water exchange is large. Conversely, low oxygen levels can occur in depressions with high sediment accumulation as well as in areas with relatively low production.

Figure 4.10: Schematic of direct and indirect links between nutrient enrichment and the BQEs.





5. Integrated assessment and harmonisation

5.1 Harmonisation of principles for defining reference and class boundaries
Richard K Johnson

5.2 Combining indicators and BQEs
Jacob Carstensen, Jesper Andersen, Ciarán Murray

5.3 Harmonisation of principles for assessing uncertainty
Mats Lindegarth, Per Bergström, Jacob Carstensen

Suggestions for harmonised assessment methods

A key challenge facing WFD implementation has been to ensure that the EQR scale is used consistently for different ecological indicators, across different water categories and types, such that the ecological indicators on which the ecological status assessment is based use reference conditions and class boundaries representing similar levels of anthropogenic disturbance. The development of approaches to set harmonised reference and boundary values is described in Section 5.1.

Second, the number of ecological indicators used for status assessment is large, so there are many ways this information can be aggregated to derive the final status assessment. The only support found in the WFD is that the status of the biological quality elements (BQEs) should be integrated using the one-out, all-out principle. A challenge is that principles for aggregating indicators within the different BQEs have not been specified, so there is a need to develop a harmonised approach to integrating information for the overall ecological status assessment. This challenge will be addressed in Section 5.2.

Third, once a harmonised approach to aggregating information has been determined, it is also necessary to determine the confidence of the status assessment. The information obtained from ecological indicators is inherently uncertain and that uncertainty propagates to the overall ecological status assessment as well. However, quantifying the uncertainty of ecological indicators is not as straightforward as it may seem, since many uncertainty components are frequently overlooked. Therefore, a framework for quantifying the uncertainty of ecological indicators from monitoring data will be presented in Section 5.3.

Finally, an ecological assessment can be biased by the available data. In many cases, not all BQEs are monitored, which may bias the assessment towards those BQEs that are monitored. Given that Sweden has several thousand water bodies, it is unrealistic to monitor all BQEs in all of them. Consequently, water body assessments have to be based on a reduced set of information or alternatively use information from other water bodies of a similar type. Extrapolating information from other water bodies has consequences for the assessment, in particular, for the confidence in the assessment. These issues will also be addressed in Section 5.3.

5.1 Harmonisation of principles for defining reference and class boundaries

Systems A and B account for regional drivers of aquatic biodiversity

The importance of biogeographic drivers for species distribution patterns is well established for inland and marine water bodies, as is the use of spatial typologies for partitioning natural variability when establishing reference conditions and setting ecological targets for restoration (e.g. Hawkins et al. 2000; Johnson et al. 2007; Johnson et al. 2013, 2014). Cognizant of the importance of regional drivers for aquatic biodiversity, the European Commission specified two spatial approaches to partitioning biological variability and assessing the ecological quality of inland and coastal water bodies (European Commission 2000). For inland water bodies, System A consists of four categories (e.g. ecoregion, altitude, catchment area, and geology for streams), and System B consists of a mixture of obligatory as well as optional factors. For coastal water bodies, System A consists of spatial coordinates (i.e. longitude and latitude), tidal range, and salinity, while system B lists eight optional variables (e.g. current velocity, wave exposure, mean water temperature, mixing characteristics, turbidity, retention time of enclosed bays, mean substratum composition, and water temperature range). According to Annex II of the WFD, type-specific reference conditions may also be established using modelling: "Type-specific biological reference conditions may be either spatially based or based on modelling, or may be derived using a combination of these methods".

INLAND WATERS: advantageous to include natural variability associated with reference conditions

In Sweden, classifications of reference conditions using modelling and *a priori* pressure criteria have a relatively long history of use for inland surface waters. For example, early attempts to quantify the ecological integrity of lakes and streams included the modelling of nutrient loads and acidification (Fölster et al. 2014). Building in large part on work done in Sweden, the European common strategy for defining reference conditions and for setting ecological status class boundaries (CIS Working Group 2.3, REFCOND), with Sweden (i.e. the Swedish University of Agricultural Sciences) as the lead partner, recommended a number of approaches to establishing reference conditions (Wallin et al. 2003). In areas with many pristine or minimally disturbed sites, spatial and modelling approaches are considered the most applicable. A major advantage of using existing sites is the possibility of including the natural variability associated with reference conditions (e.g. weather- or climate-induced variability). The CIS guidance document also lists a number of *a priori* parameters and thresholds for use in screening for potential reference sites.

Ten pressures in proposed reference filter

As part of the WATERS programme, the screening criteria (i.e. pressure/reference filter) for establishing reference conditions using minimally disturbed sites were revised. In refining the pressure-filter approach currently used in Sweden (SEPA 2007; SwAM 2013) consideration was given to the inclusion of variables that are readily available (e.g. from Water Information System Sweden) or easily obtained (e.g. land use and water chemistry). Revision included the use of improvements in classifying water quality using chemical criteria for eutrophication and acidification. Furthermore, the reference classification of inland water bodies has been refined by including information on forestry (e.g. clear-cut logging), hydromorphological alterations, and invasive species. The proposed reference filter comprises 10 variables (Table 5.1).

Table 5.1. Pressure-filter variables and threshold values used for classifying the reference conditions of inland surface waters.

Coastal area	Indicator	Suggested revisions and comments
Natural, artificial, and heavily modified waters	Excluded	Water Information System Sweden
Total phosphorus concentration	Good or high status using chemical criteria	Chemical analysis and Water Information System Sweden
Acidification	Good or high status using chemical criteria	Chemical analysis and Water Information System Sweden
Liming	Excluded	Liming database[1]
Point source pollution, priority substances, and specific synthetic pollutants	No significant effect or not present	Water Information System Sweden
Agriculture	<10% of catchment	CORINE Land Cover
Clear-cut logging	<10% of catchment, five years	Swedish Forest Agency[2]
Artificial surfaces	<1% of catchment	CORINE Land Cover
Hydromorphology	Good or high status using selected habitat criteria	Water Information System Sweden
Invasive species	If recorded as present, excluded	Swedish Species Information Centre

1) Nationella Kalkdatabasen. http://kalkdatabasen.lansstyrelsen.se/, 2) Skogsstyrelsen 2013. Utförda avverkningar – Skogsstyrelsen http://skogsdataportalen.skogsstyrelsen.se/Skogsdataportalen/

Coastal and marine waters: openness makes it difficult to find undisturbed areas

In contrast to inland surface waters, finding minimally disturbed areas in marine systems is difficult due to the openness and connectivity of the ecosystems, and to the relative importance of diffuse pressures (e.g. excess nutrients). Consequently, approaches used for coastal/transitional waters differ markedly from and are more diverse than those used for inland surface waters. Other challenges to establishing reference conditions for coastal systems stem from the fact that pressures are difficult to quantify and pressure criteria are difficult to define. Therefore, the pressure-filter approach prevalent in inland waters was not considered a practical way to establish reference conditions in coastal areas. Instead, a range of approaches is used, based on the best available data and ecological knowledge for each of the BQEs.

In the current Swedish assessment criteria for coastal waters, reference conditions are established using a combination of methods, such as using minimally disturbed sites, historical data, modelling, and expert judgment, with a particular focus on spatial representativity and functional responses to impairment. Measures of water transparency (e.g. Secchi depth) that date to the first half of the twentieth century, i.e. before the proliferation of modern agricultural practices and the widespread use of fertilizers, are used in establishing reference conditions for phytoplankton (Larsson et al. 2006). Empirical relationships between phytoplankton and Secchi depth are used to hindcast expected reference conditions (Larsson et al. 2006). In contrast, in the absence of strong empirical relationships between water transparency and benthic invertebrates, contemporary data from areas not affected by local pressures or point discharges have been used to define indicator values of at least good status (Blomqvist et al. 2006). This means that the emphasis is on determining the good/moderate boundary, with no attempt to establish reference conditions. For macrovegetation, the depth distribution of indicator taxa based on empirical relationships and/or expert opinion has been used to establish reference conditions. However, results from WATERS suggest that there is a clear relationship between Secchi depth and the suggested new indicators for macrovegetation, which means that reference conditions could be derived from Secchi depth using an approach similar to that used for phytoplankton (see Section 4.2.2).

Typology- and model-based approaches to establishing reference conditions: a comparison

Theoretically, predictive models should perform better than typology-based approaches due to the use of continuous as opposed to categorical variables when characterising species–environment relationships. Besides the use of continuous variables, another major advantage is that many potentially important predictor variables can be included in models. In typology-based approaches, however, the number of categories used is often limited due to constraints related to data availability, since finding an adequate number of sites to estimate within-type variability is often difficult in areas strongly impacted by land use. Other advantages are that models can be used over large spatial scales, and if calibrated using variables characterising seasonal or among-year variability (e.g. number of degree days), can provide more accurate measures of the intrinsic dynamics of natural communities. Earlier work, as well as insights from research within WATERS on both freshwater and marine ecosystems, has shown that modelling is often better than typology at accounting for natural sources of variability when establishing biological reference conditions and class boundaries. The main modelling approaches used in

WATERS included generalised linear models (GLMs), generalised additive models (GAMs), machine learning (ensemble) methods such as Random Forest, and methods for predicting the probability of taxon occurrence and taxonomic completeness (e.g. RIVPACS) (e.g. Moss et al. 1987; Breiman 2001; Elith et al. 2006; Marmion et al. 2009). In the following we compare the efficacy of spatially and model-based approaches to establishing reference conditions, and the use of models to account for natural variability when calculating biological response metrics.

Using a dataset compiled from the national lake survey conducted in 2000, we compared typology- and model-based approaches to establishing reference conditions for littoral invertebrate assemblages in lakes (Hallstan and Johnson, unpublished). Reference lakes (n = 464) were divided into calibration (approximately 80%) and validation (approximately 20%) datasets. Models were trained using the calibration lakes and performance was evaluated using the validation lakes. For modelling taxonomic completeness, one model was developed for each taxon. All analyses were conducted using R software (R Core Team 2015).

Typology- and model-based approaches tested with three performance criteria

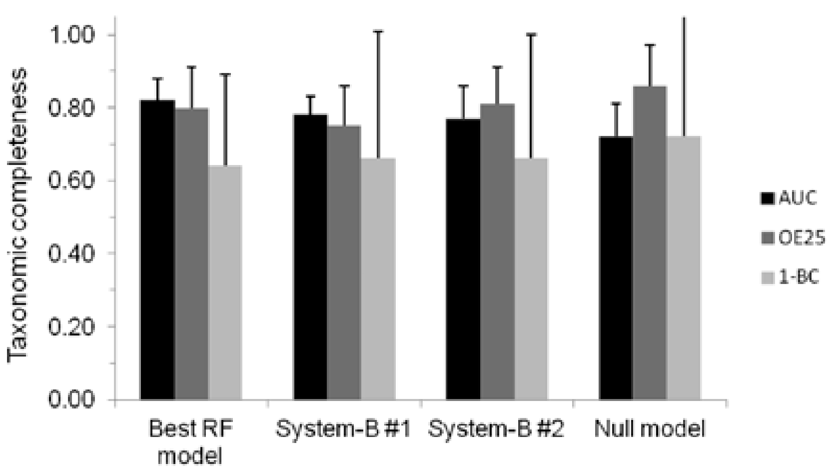
Three indices describing taxonomic completeness were used to compare the efficacy of typology- and model-based approaches to partitioning natural variability. In addition, we compared the performance of two modelling (i.e. linear and Random Forest) and two typology-based approaches to partitioning natural variability associated with using stress-specific response variables. The typology-based approaches used here are: (i) reference values calculated using five variables (i.e. water colour, alkalinity, altitude, mean depth, and the binary variable south of Limes norrlandicus), here referred to as System-B #1; and (ii) reference values calculated using the recently proposed lake-typology approach of Drakare (2014), here referred to as System-B #2. In addition, when comparing models and typology, we also used the typology approach currently used in national assessments (SEPA 2007; SwAM 2013). The current typology for benthic invertebrates in lakes is based on the three major ecoregions, with no attempt at partitioning variability associated with ecotypes within the ecoregions. In contrast, the approach proposed by Drakare (2014) categorises Swedish lakes into 48 types based on region, mean depth, alkalinity, and humic content.

Taxonomic completeness refers to the similarity of a specific taxonomic assemblage to what would be expected without anthropogenic change, and models of taxonomic completeness are obtained by calibration using only minimally disturbed sites and predictor variables unaffected by human activities. The three indicators of model performance used here are: (i) Area under the receiver operator curve (AUC) is an index that does not require the inclusion of a probability threshold and is commonly used in species distribution model studies (Fielding and Bell 1997). AUC ranges from 0 to 1, where 1 indicates a perfect prediction and 0.5 indicates a prediction no better than what would be expected by chance. (ii) Observed-to-expected (O/E) is the most common index used for assessing the performance of RIVPACS-type models (Moss et al. 1997). O/E is calculated by first summarising all occurrence probabilities above a specified threshold (here we used 0.25), which constitutes the expected taxa richness, E. E is then compared with O, the number of the expected taxa actually found in the sample. (iii) An alternative measure to O/E is BC (Van Sickle 2008). BC is similar to the Bray-Curtis dissimilarity measure and is calculated as $BC = \sum |O - P| / \sum |O + P|$, where O is 1 for presence, 0 is for absence, and P is the predicted probability of taxa occurrence. Here we use 1-BC, with low values indicating better performance. In addition, a null model (Van Sickle et al. 2005) was used to obtain a minimum level of accuracy for model comparisons, i.e. the null model assumes no variation across environmental gradients.

Random Forest models more accurate than typologies

All three indicators of performance (i.e. AUC, O/E, and 1-BC) indicated that Random Forest models were generally more accurate and precise than were the typology and null models (Figure 5.1). On average, validation indicated that Random Forest models were the most accurate (e.g. AUC = 0.82 ± 0.06, mean ± 1SD, for the best Random Forest model, n = 31 predictor variables). The "best" Random Forest model comprised five water quality variables (e.g. water colour, alkalinity, and total phosphorus), 14 climate (e.g. temperature and precipitation) and geographic (e.g. lake area and catchment area) variables, six variables characterising land use (e.g. forest type, wetlands and lake shores having coniferous forest), and five variables characterising substratum (e.g. stones and woody debris). Random Forest models were also more accurate than the typologies based on WFD System-B #1 (AUC 0.78 ± 0.05) and System-B #2 (AUC 0.77 ± 0.09, Drakare 2014) typologies.

Figure 5.1: Mean (±1 SD) of three indicators of model performance at estimating taxonomic completeness using a Random Forest model (best model, n = 31 predictor variables), System-B #1 typology (n = 5 variables), System-B #2 typology (n = 4 categories; Drakare 2014), and a null model. Note that the low values of 1-BC indicate high performance. For clarification of abbreviations, see text.



To detect stress-specific biological response

To assess the ability of typology-based approaches and models to account for natural variability in stress-specific indices, we used lake littoral benthic invertebrates as a model assemblage and calibrated models for the MILA index (Johnson and Goedkoop 2007) used to assess acidity and the ASPT index (Armitage et al. 1983) used to assess general degradation (for descriptions of these indicators, see Section 4.2). Models were calibrated using linear least squares regression and Random Forest (Breiman 2001). The two modelling approaches were then compared with two typology-based approaches: the current typology for benthic invertebrates in lakes based on the three major ecoregions (SEPA 2007; SwAM 2013) and the approach proposed by Drakare (2014). A null model was also included in model comparisons.

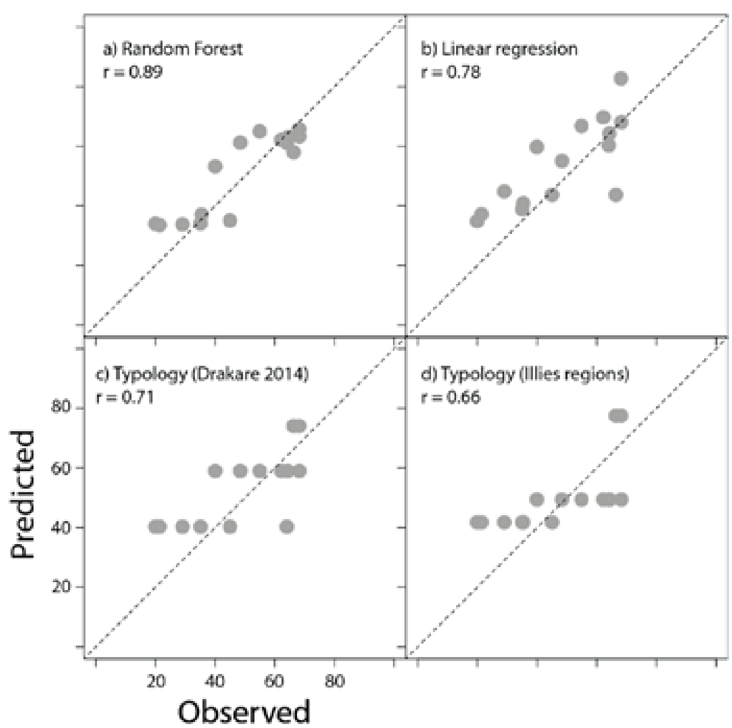
Linear regression and Random Forest were better at partitioning natural variability than was the typology approach currently used in national assessment (based on biogeographic zones, i.e. Illies regions) and the typology approach recently proposed by Drakare (2014) (Figure 5.2).

The lake typology proposed by Drakare (2014) was better than the coarse typology (comprising three ecoregions) currently used in national assessments. The most important predictor variables varied between models and indicators. For example, the mean temperature of the warmest month was an important predictor in Random Forest models for ASPT and MILA, whereas precipitation was an important predictor in linear regression but not Random Forest models. Most geographic variables, except for altitude, were not important predictor variables in either of the models. By contrast, catchment land cover was an important predictor of both ASPT and MILA; for example, broad-leaved forest was an important predictor of ASPT, whereas sparse vegetation was an important predictor of MILA.

Model-based approach supported by the literature

Random Forest modelling was superior to the other three approaches to partitioning natural variability, with a correlation between the observed and predicted MILA values of 0.89. By comparison, typology-based approaches had correlations of 0.71 (proposed typology

Figure 5.2: Comparison of two modelling (a, b) and two typology (c,d) approaches to establishing reference conditions for the MILA acidity index.



of Drakare 2014) and 0.66 (current typology). This finding implies that site-specific, model-based reference values can improve the assessment of ecological status. These findings support the growing literature advocating the use of models as opposed to typology in characterising reference conditions (e.g. Davy-Bowker et al. 2006; Hawkins et al. 2010; Aroviita et al. 2009). Site-specific (model-based) approaches to estimating reference conditions and ecological targets have been successfully incorporated into regional and national assessments in Great Britain, the USA, and Australia.

Work within the WATERS programme has demonstrated that model-based, site-specific reference values can improve assessments of the ecological status of freshwater and marine ecosystems. In the examples discussed above, we focused on comparing two modelling techniques, linear regression and Random Forest. However, several other modelling techniques have been found to work well at predicting taxa presence/absence, such as Multivariate Adaptive Regression Splines (MARS), GAM, GLM, and maximum entropy (Elith et al. 2006; Marmion et al. 2009). A modelling approach that has recently attracted attention is community-level modelling (Maguire 2015), in which taxa co-occurrences are included in the modelling process. Including potential biological interactions may improve predictive performance if biotic interactions are important predictors of community composition (e.g. Johnson and Hering 2010).

Setting class boundaries: a critical step in the ecological classification scheme

Once the reference condition or ecological target has been established, the next step is usually to describe the biological responses to putative pressures. This is done by defining classes representing different stages of ecological integrity and setting class boundaries, which is one of the most critical steps in establishing an ecological classification scheme. Four methods are frequently used separately or in combination to establish class boundaries: (i) breakpoints or discontinuities in the pressure-response relationship (Figure 5.3a), (ii) the distribution of sites of high status (e.g. the lower 25th or 10th percentile of reference sites) (Figure 5.3b), (iii) dividing the pressure-response gradient into equidistant classes, and (iv) plots of paired re-

sponse variables (e.g. variables capturing different functions of the biological community) (Figure 5.3c).

In developing classification schemes for assessing ecological status and setting class boundaries, knowledge of the variance associated with the pressure-response relationship is crucial for quantifying the uncertainty associated with using the response indicator (see Section 2.2). In the WFD, much importance is attributed to the good/moderate boundary, as European legislation requires the implementation of programmes of measures for water bodies failing to achieve good ecological status. Accordingly, inferences of impairment need to be made in the context of model error, and ideally threshold levels (or class boundaries) should be set with *a priori* knowledge of type 1 (false-positive) and type 2 (false-negative) error estimates (Johnson et al. 2006). In other words, classification schemes must explicitly address the question: What is the probability that a site is assigned to the wrong class? If a site is incorrectly placed in a class denoting poorer ecological status than the actual condition, this would be a type 1 (false-positive) error, whereas incor-

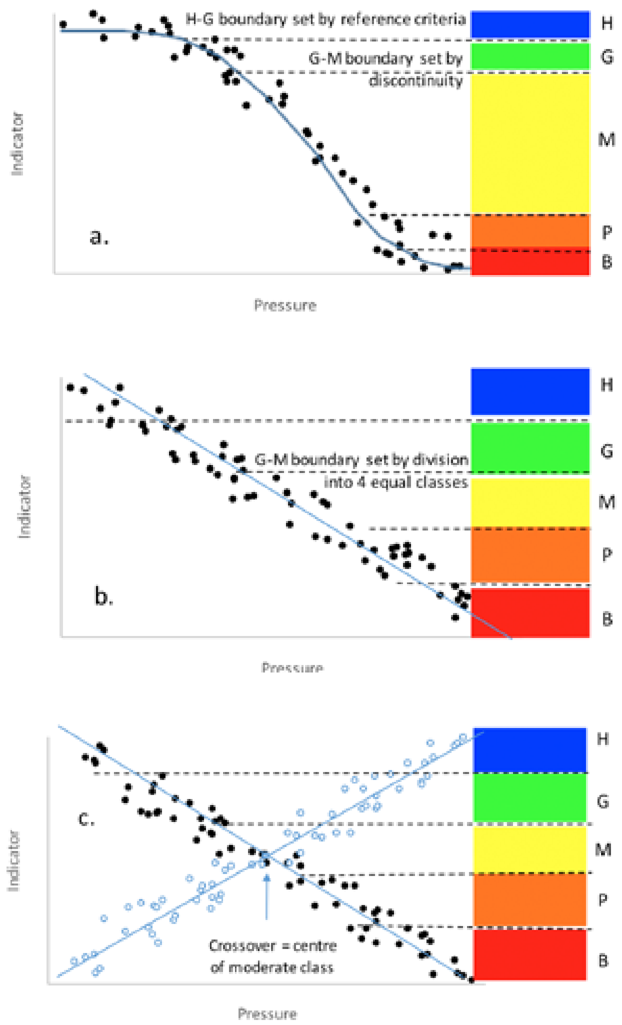


Figure 5.3: Examples of three methods to establish the boundary between good and moderate status; see text for further explanations. Modified from Schmedtje et al. (2009).

rect placement in a class denoting higher ecological status than the actual condition would be classified as a type 2 (false-negative) error. In biomonitoring and assessment, false-negative errors are often considered more serious than false-positive errors, since ecological degradation may then proceed undetected (e.g. Johnson 1998; Johnson et al. 2006).

Rapid response to impairment is important

Organism response to human-induced stress is not always linear, and selecting indicators that respond rapidly at the onset of impairment is important in order to quantify the early effects that humans may have on ecosystem integrity. At high levels of stress, dose-response relationships often display a levelling-off (e.g. low variance around the regression line), typically resulting in funnel-shaped response curves. Taxa that display rapid responses (high slopes, low uncertainty) to low or moderate levels of stress may be used together with taxa that display responses at high levels of stress in order to cover broad gradients of impairment (Johnson and Hering 2009). Knowledge of the pressure-response relationships of various groups can improve our understanding of the use of early-warning indicators for detecting ecological change and for designing cost-effective monitoring programmes (Johnson et al. 2007).

Ideally, the number of classes and placement of class boundaries should be based on well-established pressure-response relationships and measures of uncertainty associated with the response relationships. In practice, however, distinct breakpoints justifying the five ecological classes (i.e. four breakpoints) stipulated by the WFD seldom exist. Consequently, in developing classification schemes, boundaries are often established using single breakpoints or a percentile of reference sites, followed by dividing the remainder of the pressure-response gradient into equidistant classes.

Insights gained from work within the WATERS programme using newly developed statistical approaches, such as Threshold Indicator Taxa ANalysis (TITAN; Baker and King 2010) and Random Forest (Breiman 2001), have helped in more objectively identifying potential breakpoints of pressure-response relationships, and these methods been found to be useful for establishing important class boundaries.

Possible ways towards further improvement

To further develop harmonisation and possibly adjust reference values within and between inland and marine ecosystems, we recommend:

1) that the pressure filter used for characterising the reference conditions of inland surface waters be revised to include recent developments in classifying status using chemical water quality criteria, land use classifications, hydromorphological alteration, and invasive species;

2) improving our knowledge of anthropogenic pressures in coastal areas according to annex 4 (EC 2011) and, if possible, developing alternative benchmarks and “virtual references” for assessments in line with current EU guidelines; and

3) further developing and testing models in order to partition natural variability and estimate reference conditions for inland surface and marine ecosystems.

5.2 Combining indicators and BQEs

Remaining issues for integrated assessment for WFD

To carry out an integrated assessment of ecological status according to the WFD, two requirements have to be fulfilled: first, WFD-specific indicators representative of more than a single BQE and with a numerical target value expressing the good/moderate boundary have to be monitored and, second, methods for combining indicators within BQEs as well as methods or principles for combining BQE-specific classifications of ecological status into a final integrated assessment have to be agreed on and operationalised.

Given that the WFD was adopted in 2000, these criteria were expected to be fulfilled today, but there are still a number of unresolved issues that need to be addressed. Do we have the required indicators? On a pan-European scale, hundreds of indicators or indices have been developed over the past 15 years, but a WATERS study of the Nordic countries reveals that only a reduced set of these is used in practice (Andersen et al. 2016). Do we have methods for combining indicators and BQEs? Several multi-metric indicator-based tools are available, but so far none of these operational tools has been used in the context of national WFD Initial Assessments, neither directly (as they are) nor indirectly (specifically modified for WFD purposes) (Korpinen et al. 2015). Accordingly, WATERS has focused on principles of indicator integration, both within and between BQEs. The findings are summarised in the following section leading to recommendations on how to carry out future integrated assessments of ecological status.

5.2.1 Combining indicators and BQEs

Poor guidance on integration causes confusion at the indicator level

The framework for conducting an integrated assessment is in principle a simple hierarchical system: (1) an assessment is based on the monitoring and application of multiple indicators, for example, from lakes, rivers, transitional waters, and coastal waters; (2) indicators are grouped, for example, as BQEs or supporting elements; and (3) BQEs are combined using the one-out, all-out principle. Hence, aggregation and integration take place at three levels.

However, the WFD does not provide any guidance on how to aggregate and integrate at the indicator level. Indicators have been developed through a bottom-up process and methods for aggregation and integration vary substantially. Furthermore, the WFD does not provide guidance on how to aggregate and integrate BQEs. However, the WFD Common Implementation Strategy does provide some guidance in this regard, and several approaches have been developed. Combining BQEs into a final classification is straightforward and the WFD provides clear guidance on this, i.e. the one-out, all-out principle.

Several approaches to combining indicators exist (Andersen et al. 2016), ranging from averaging between stations to complex index calculations. For aggregation at the BQE level, a wide range of methods is used, for example: (1) simple averaging across indicators, (2) weighted averaging across indicators, (3) conditional rules, (4) scoring or rating, (5) multi-metric approaches, (6) multi-dimensional approaches, (7) decision trees, and (8) probabilistic approaches. For more information, see Moksnes et al. (2013) and Borja et al. (2015).

The WFD was adopted in 2000 and now, more than 15 years after its adoption, numerous indicators and indices developed and applied in the classification of ecological status (Birk et al. 2012). Despite the pan-European Common Implementation Strategy (CIS) providing comprehensive guidance and three rounds of indicator intercalibration, apparently no tools for the integrated assessment of the ecological status of surface waters have been subjected to joint and harmonised pan-European development.

Within WATERS, we have analysed the use of indicators in integrated assessments and, in particular, the principles and tools for aggregation and integration (Borja et al. 2015; Andersen et al. 2016). The main conclusions are: (1) a bottom–up process has resulted in considerable variation in principles and methods, especially as pertains to indicators, BQEs, and supporting elements; (2) very few tools for integrated assessment have been developed, tested, and applied; and (3) consequently, comparisons between BQEs and water types can be complicated and uncertain.

5.2.2 WATERS contribution

Existing tool modified for assessment of ecological status

Given the above conclusions, WATERS has carefully considered how to improve the fundamentals for carrying out integrated assessments and enabling cross-cutting comparisons, mostly between water types. Aiming for harmonisation and coordination at the index level at this stage of implementation is not possible. However, aiming for harmonisation at the level of quality elements and for the harmonisation of methods for combining BQEs and supporting elements into a whole system assessment has been in focus, as this could be attained through the development and application of harmonised multi-metric indicator-based assessment tools, such as the HELCOM Eutrophication Assessment Tool (HEAT) or the Nested Environmental Status Assessment Tool (NEAT) (Borja et al. 2016). The benefits of using such tools are likely to include: (1) ability to make cross-element comparisons, (2) improved ability to make cross-water-category comparisons, and (3) potentially also an improved understanding of downstream changes in ecological status as well as their upstream causes.

The development of WFD- and Marine Strategy Framework Directive (MSFD)-specific assessment tools like those currently used by HELCOM and OSPAR for the assessment of eutrophication status could represent a step forward in the assessment of ecological status, especially if the tools are not water-category specific. We do not envisage a pan-European process leading to the next generation of tools, but rather national and regional testing and application. Ultimately, Member States and Regional Marine Conventions (e.g. HELCOM and OSPAR) have an interest in using tools that not only allow for comparisons between water categories and sub-basins but also allow for comparisons between downstream and upstream water bodies.

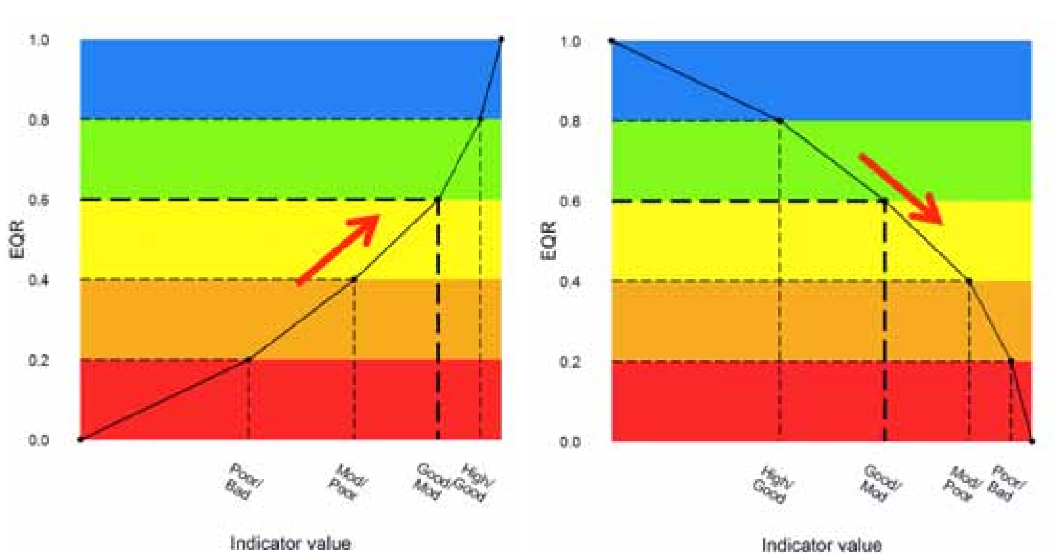
In WATERS, an existing tool has been modified for the assessment of ecological status according to the WFD. The prototype tool adapted and tested by WATERS is based on the DEVOTES Nested Environmental Status Assessment Tool (NEAT; see Andersen et al. 2016 for details), which is a modular and flexible indicator-based tool for the assessment of good environmental status according to the EU's MSFD (EC 2008).

Normalisation of indicators permits comparison

The aim is to compare and aggregate multiple indicators having different measurement scales. This can be achieved when all indicators are operating on the same assessment scale. For this, all indicators must be normalised to a scale from 0.0 to 1.0, equal to the ecological quality ratio (EQR) known from the EU's WFD (EC 2000). The indicator scale is transformed to the normalised scale by continuous piecewise linear transformation. Indicator values can increase or decrease with improved ecological status, so the piecewise linear function can be increasing or decreasing with increasing indicator value though, in the present version of the tool, it must be monotonic.

A key step in defining an indicator is to specify the breaks in the piecewise linear function (Figure 5.4). These breaks are chosen such that they occur at the boundaries between status categories. NEAT uses indicators having an assessment scale with five distinct classes. This is achieved by specifying the indicator values on the original measurement scale that correspond to transformed EQR values of 0.2, 0.4, 0.6, and 0.8, respectively indicating the bad/poor, poor/moderate, moderate/good, and good/high status boundaries. The EQR values of 0.0 and 1.0 correspond to the ultimate range that can be expected in the measured indicator values.

Figure 5.4: The piecewise linear transformation of indicators to an EQR scale with equidistant class boundaries, employed for an indicator with increasing quality (left) and decreasing quality (right).



Having defined the breaks in the piecewise linear function, the EQR value for an indicator's observed value is then obtained by linear interpolation between the two closest boundary values. Normalisation to a common scale for indicators increasing with improving status is done according to Eq. 1:

$$EQR = \begin{cases} 0 & \text{if } \hat{\mu} \le B_{min} \\ \dfrac{\hat{\mu} - B_{min}}{B_{PB} - B_{min}} \cdot 0.2 & \text{if } B_{min} < \hat{\mu} < B_{PB} \\ \dfrac{\hat{\mu} - B_{PB}}{B_{MP} - B_{PB}} \cdot 0.2 + 0.2 & \text{if } B_{PB} < \hat{\mu} < B_{MP} \\ \dfrac{\hat{\mu} - B_{MP}}{B_{GM} - B_{MP}} \cdot 0.2 + 0.4 & \text{if } B_{MP} < \hat{\mu} < B_{GM} \\ \dfrac{\hat{\mu} - B_{GM}}{B_{HG} - B_{GM}} \cdot 0.2 + 0.6 & \text{if } B_{GM} < \hat{\mu} < B_{HG} \\ \dfrac{\hat{\mu} - B_{HG}}{B_{max} - B_{HG}} \cdot 0.2 + 0.8 & \text{if } B_{HG} < \hat{\mu} < B_{max} \\ 1 & \text{if } \hat{\mu} \ge B_{max} \end{cases} \quad (1)$$

where for a given indicator B_{min} is the minimum value, B_{PB} is the indicator value corresponding to the poor/bad status boundary, B_{MP} is the indicator value corresponding to the moderate/poor status boundary, B_{GM} is the indicator value corresponding to the good/moderate status boundary, B_{HG} is the indicator value corresponding to the high/good status boundary, and B_{max} is the maximum possible value.

The WATERS tool has a hierarchical structure that corresponds to WFD requirements: (1) phytoplankton, (2) benthic vegetation, (3) benthic invertebrates, (4) fish (only relevant to lakes, rivers, and transitional waters), and (5) supporting elements. The classification of ecological status for a specific water body is a simple four-step process:

Step 1 Indicators are normalised according formula 1 and an indicator EQR value ranging between 0.0 and 1.0 s is calculated.

Step 2 Indicator EQR values are aggregated within BQEs and supporting elements using a hierarchical structure of sub-elements.

Step 3 The BQE or supporting element is classified in one of the following five classes: high (1.0–0.8), good (0.8–0.6), moderate (0.6–0.4), poor (0.4–0.2), and bad (0.2–0.0).

Step 4 BQEs are combined to give a final classification using the one-out, all-out principle.

The added value of this approach developed by WATERS is: (1) it is scientifically sound and transparent; (2) the prototype tool is easy to use; (3) it is flexible and can be used for a range of related assessments (Table 5.2), not only WFD-related assessments but also for those that are MSFD related; and (4) it has been tested in a limited number of pilot study areas, as described on the next page.

Table 5.2. Summary of possible classification outcomes in the prototype WATERS tool. MSFD: pass or fail. Hybrid: "traffic light", one passed, two failed. WFD: five classes, two passed, three failed.

MSFD	Hybrid	WATERS	EQR	Deviation range
Good (Unaffected	Good	High	0.8 ≤ EQR ≤ 1.0	No deviation from background values
		Good	0.6 ≤ EQR < 0.8	Slight deviation below target value
Not good (Affected)	Moderate	Moderate	0.4 ≤ EQR < 0.6	Slight deviation above target value
	Bad	Poor	0.2 ≤ EQR < 0.4	Major deviation above target value
		Bad	EQR < 0.2	Significant deviation above target value

Interesting results from test of tool prototype in different surface water types

A prototype WATERS tool has been tested in a range of pilot study areas including rivers, lakes, and coastal waters. However, only the tests in coastal waters and lakes can be regarded as fully integrated assessments involving multiple indicators and BQEs (Table 5.3). It should be stressed that these results are simply examples demonstrating the integrated assessment tool and are not based on official boundaries and indicator values.

Table 5.3. Test classifications for pilot studies in coastal waters using multiple quality elements, i.e. phytoplankton, benthic vegetation, benthic invertebrates, and supporting elements (e.g. nutrient concentrations). An asterisk indicates the quality element with the lowest classification. Please note that the one-out, all-out principle is applied for biological quality elements only, and that the values do not represent an official assessment.

Pilot study area	Habitat	Quality elements				BQE-based assessment
		Phyto-plankton	Benthic vegetation	Benthic invertebrates	Supporting elements	
Inre Bråviken	Coastal	0.532	0.493	0.626	0.371*	Moderate
Inre Slätbaken	Coastal	0.380	0.478	0.000*	0.293	Bad
Kaggebofjärden	Coastal	0.462	0.603	0.532	0.336*	Moderate
Yttre Bråviken	Coastal	0.524	0.751	0.677	0.473	Moderate
Trännöfjärden	Coastal	0.472	0.672	0.148*	0.322	Bad
Lindödjupet	Coastal	0.447	0.710	0.546	0.348*	Moderate
Kärrfjärden	Coastal	0.445	0.810	0.603	0.367*	Moderate
Ryssbysjön (2009–2014)	Lake	0.371*	0.516	0.757	0.598	Poor
Stensjön (2009–2014)	Lake	0.671	-	0.698	0.356*	Poor
Vässledasjön (2009–2014)	Lake	0.689	0.610	-	0.489*	Moderate
Stora Nätaren (2009–2014)	Lake	0.557	0.623	0.764	0.525*	Moderate
Försjön (2009–2014)	Lake	0.883	-	0.856	0.497*	Moderate
Landsjön (2009–2014)	Lake	0.340*	0.399	-	0.685	Poor
Glimmingen (2009–2014)	Lake	0.874	0.793*	0.850	-	Good

The tests of the prototype WATERS tool for the classification of ecological status in seven Swedish coastal waters bodies reveal that: (1) none of the pilot study areas has a good ecological status; (2) phytoplankton was never the lowest-ranking quality element in a coastal area; and (3) benthic fauna determined the final classification of ecological status in two cases (i.e. Inre Slätbaken and Trännöfjärden). For all seven water bodies, the assessments were based on two indicators for phytoplankton (chlorophyll-a and biovolume), two indicators for benthic vegetation (cumulative cover and species richness), one indicator for benthic fauna (BQI), and two indicators for supporting elements (total nitrogen and total phosphorus concentrations).

Unlike for coastal areas, testing lake assessments identified one study area having a good ecological status, i.e. Lake Glimmingen. Other interesting results were: (1) fish determined the final classification in four out of six cases; (2) phytoplankton determined the final classification in two cases; and (3) other quality elements did not determine the final classification of ecological status. The integrated assessments for the lake examples were based on five phytoplankton indicators, one benthic vegetation indicator, two benthic invertebrate indicators, and one fish indicator.

Information on uncertainty one of the largest advantages

Perhaps the greatest added value of using the prototype WATERS tool is its data-driven confidence assessment of the classification results. The confidence of the classification results at different aggregation levels in the integrated assessment tool is quantified by uncertainty propagation starting from the indicator level. Hence, an uncertainty estimate for each of the indicator status values is required for assessing the confidence. Such uncertainty estimates can be calculated using the uncertainty framework (see Section 5.3) developed within WATERS (Carstensen and Lindegarth 2016). The propagation of indicator uncertainty to the overall integrated assessment is calculated by Monte Carlo simulation, typically using 1000 or 10,000 simulations and a normal distribution for the indicator. The distributions of these simulations are compared with the class boundaries to derive probabilities for each of the status classes. For example, in Inre Bråviken (Table 5.2), the probability of achieving bad status is 14%, poor status is 80%, and moderate status is 6%. Hence, the most likely status is poor, which also contains the median of the distribution given as the final outcome. Similarly, for Lake Glimmingen (Table 5.3), the probability of achieving moderate status is 2%, good status is 56%, and high status is 42%, yielding good status as the overall assessment. The testing of the integrated assessment tool, using as realistic values as possible, demonstrated that the final assessment in most cases spanned three status classes. Ecological status assessment is therefore inherently uncertain, since it is based on uncertainty information from the indicators.

5.2.3 Summary and suggested next steps

The integration of BQEs/QEs using the one-out, all-out principle is considered simple and transparent. In contrast, the integration of indicators within BQEs/QEs uses methods ranging from simple averaging to complex aggregation, with substantial variation between them (e.g. Borja et al. 2014). The absence of a harmonised approach complicates comparisons between different groups of indicators, not only within water bodies but also across water bodies and water categories. Considering specific indices, the variety of integration principles and methods is considerable, probably because indices have been developed through a bottom-up process.

The prototype WATERS tool has been developed and tested and the preliminary results are promising. With further development and testing based on a significantly larger dataset, it would be possible to develop a national multi-metric indicator-based assessment tool enabling not only WFD-specific assessments but also comparisons across BQEs and water types as well as coordination with other relevant directives, primarily assessments under the MSFD.

5.3 Harmonisation of principles for assessing uncertainty

The WFD and associated guidance documents (e.g. CIS Guidance Documents #7 and #13) stress that estimates and classifications of ecological status need to be accompanied by assessments of uncertainty. Although guidance is provided concerning general definitions of precision and confidence, analyses have demonstrated that these concepts are addressed in very different ways and with different levels of strictness among BQEs in the Swedish assessment criteria (e.g. Lindegarth et al. 2013a, 2013b) as well as in a European context (e.g. Noges et al. 2009; Hering et al. 2010; Birk et al. 2012). Thus, in the current assessment criteria, definitions of uncertainty vary among BQEs and guidelines for estimating uncertainty in classification are lacking for the most important scale, i.e. within water bodies in six-year assessment periods (SEPA 2010; SwAM 2013). As a result, the uncertainties associated with classifications of individual BQEs and integrated assessments lack transparency. To address inconsistencies among BQEs and to provide a comprehensive methodology covering the full requirements of the Directive, WATERS has developed and tested a general framework for uncertainty assessment, which can be implemented for all BQEs (e.g. Lindegarth et al. 2013a, 2013b; Bergström and Lindegarth 2016; Carstensen and Lindegarth 2016). The framework is based on well-known and robust statistical theory and is completely coordinated with tools and methods for integrated assessment developed in other parts of the programme.

5.3.1 A framework for assessing the status and uncertainty of water bodies

Spatial, temporal, and methodological sources of uncertainty

The ecological status of BQEs is assessed using monitoring data for selected indicators. Despite using standardised sampling and lab protocols, most estimates are associated with some degree of uncertainty.

Uncertainty is caused mainly by factors related to spatial or temporal variability and by sampling and measurement procedures (Figure 5.5). The WATERS framework for uncertainty assessment is based on the facts that (1) the importance of these factors can be estimated and (2) estimates of variability can be combined into useful estimates of overall uncertainty.

$$
\begin{aligned}
y = \mu + & \underbrace{year + YEAR + season + SEASON \times YEAR + DIURNAL + IRREGULAR}_{temporal\ sources\ of\ uncertainty} \\
& + \underbrace{gradient + GRADIENT + PATCHINESS}_{spatial\ sources\ of\ uncertainty} \\
& + \underbrace{YEAR \times GRADIENT + SEASON \times GRADIENT}_{spatio-temporal\ interactions} \\
& + \underbrace{sampling\ devices + PERSON + instrument + REPLICATE}_{sampling\ and\ measurement\ uncertainties}
\end{aligned}
\qquad (2)
$$

The importance of different sources of variation can be partitioned using a linear mixed model, where each measurement, y, can be expressed as a combination of the mean, μ, and a set of relevant factors and interactions (Eq. 2). These factors are either "RANDOM", in which case they represent unpredictable variability and their importance is estimated as variance components (s^2), or they are *"fixed"*, implying that each level of the factor is described by a separate parameter. These fixed factors represent predictable effects that can measured and accounted for in the model. Several methods can be used to estimate the model parameters, i.e. variance components and fixed-effect parameters, of linear mixed models (e.g. Bolker et al. 2009). The important thing, however, is to use a comprehensive model and to combine data from many water bodies and years to achieve robust estimates of the relevant components.

Quantitative estimates of the variance components associated with spatial, temporal, and methodological factors are used for estimating the overall uncertainty of status assessments. In combination with information on monitoring designs, i.e. number of years sampled, number of sites per water body, and number of samples per site and year, it is possible to calculate the overall precision (or uncertainty) of an indicator in a given water body during a six-year assessment period. This is done using general procedures for uncertainty (or error) propagation (e.g. Cochran 1977; Taylor 1997). Several examples of uncertainty propagation exist in relation to assessment methods in general and to the WFD in particular, presented by, for example, Clarke et al. (2002, 2006a, 2006b) and Clarke and Hering (2006). These studies have demonstrated the need for the combined assessment of various sources of uncertainty, of the spatial and temporal context of uncertainties, and of the benefits of reducing uncertainty by optimising sampling designs.

Estimating uncertainty: a typical example

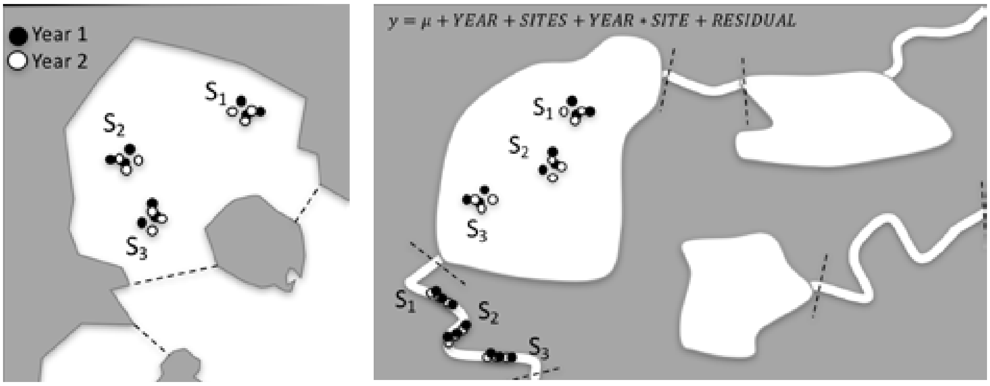
For example, consider a marine or freshwater water body where an indicator has been monitored for two years, at three sites, with three samples collected per site per sampling occasion (Figure 5.5). The uncertainty of an indicator, which for example is estimated as the average of observations ($\bar{y}$), in a water body and a six-year assessment period is affected by several variance components, i.e. among years, s_Y^2, among sites, s_S^2, due to interactive variability, s_{Y*S}^2, and among samples within sites and times, s_e^2. The total variance of the average, $\bar{y}$, can be calculated as follows:

$$V[\bar{y}] = \frac{s_Y^2 * \left(1 - \frac{a}{Y}\right)}{a} + \frac{s_S^2}{b} + \frac{s_{Y*S}^2}{ab} + \frac{s_e^2}{abn}, \qquad (3)$$

This formula for error propagation illustrates how individual uncertainty components are combined into an overall variance estimate and, importantly, how the numbers of samples, sites, and years affect the variance and uncertainty. Increasing the number of samples *(n)* reduces the uncertainty due to small-scale variability within sites and years, but does not affect the uncertainty caused by variability among years or sites. Monitoring at many sites *(b)* reduces the uncertainty associated with sites and samples, but does not result in any reduction of the uncertainty associated with years. Similarly, sampling for a number of years (a) reduces the temporal uncertainty, interactive variability, and patchiness, but not among sites. Note also that if all years within an assessment period are sampled, i.e. $a = Y = 6$, all possible levels of that factor are sampled, which implies that the distribution over the six years (constituting the entire relevant population) is estimated and therefore does not contribute any random variation (e.g. Cochran 1977; Clarke 2012).

This simple example with balanced sampling illustrates the relevant concepts necessary to estimate the uncertainty of status assessments within water bodies during a WFD assessment cycle (for a more general formulation, see Carstensen and Lindegarth 2016). The framework can be applied to any indicator and to all BQEs within the WFD. Similar formulae for error propagation can also be derived and used to assess uncertainty at other spatial and temporal scales of interest (e.g. for overall uncertainty within water body types as required within the MSFD; Lindegarth et al. 2014). Finally, it is also worth clarifying that calculating the confidence in estimation and classification requires that sampling variances be transformed into standard errors (SEs), which are computed as the square root of overall variances, i.e. $SE[\bar{y}] = \sqrt{V[\bar{y}]}$.

Figure 5.5: Schematics of orthogonal monitoring designs in a coastal water body (left) and in a lake and stream (right). In the examples, the number of sampled years, *a* = 2, the number of sites per water body, *b* = 3, and the number of samples per site and year, *n* = 3. The generic mixed model of corresponding spatial and temporal factors is inserted in the right panel (modified from Lindegarth et al. 2013a).



5.3.2 Reducing uncertainty

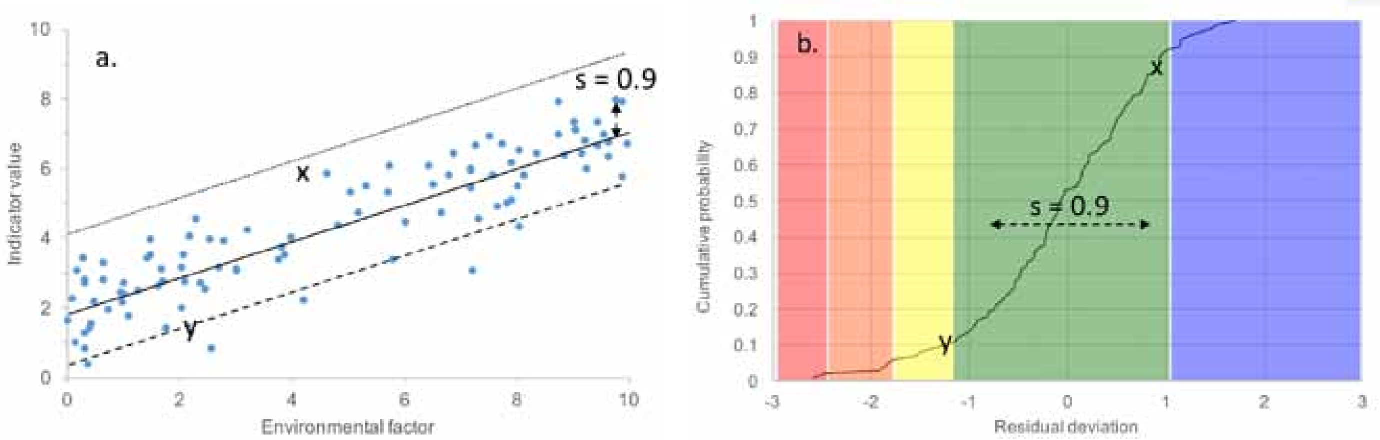
The framework for uncertainty assessment fulfils the fundamental requirement of the WFD, namely, that the uncertainty of monitoring programmes and the status of individual assessments should be quantified. Importantly, it also provides a tool for reducing uncertainty in status assessments and for optimising the costs and benefits of monitoring programmes. This can be achieved by two principal approaches: (1) by accounting for fixed factors in the mixed linear model and (2) by modelling overall precision under different monitoring designs.

Removing variability due to predictable factors

As described briefly above, the general framework using mixed linear models enables us to incorporate fixed factors, with predictable effects on any selected indicator. Such procedures mean that the variability can be explained and the uncertainty quantified and partitioned. This property can be very useful in accounting for natural spatial and temporal variability within water bodies and water body types (e.g. Johnson et al. 2014; Carstensen and Lindegarth 2016; Leonardsson et al. 2016). Importantly this approach is also compatible with procedures for setting class-boundaries and status assessment, where deviations (residuals) from the fitted model are used. Although the metods for calculation and testing procedure may vary slightly, this procedure can be conceptually described in a number of general steps (Johnson et al. 2014):

1. Fit a regression model (linear or non-linear) between the indicator and potentially important environmental factors (Figure 5.6a).
2. Define reference values and class-boundaries in relation to the model. Use the fitted model as an average reference condition across the environmental gradient if the data come from reference sites, or an alternative benchmark at any selected level of deviation from the fitted model if data include impacted sites.
3. Calculate deviations from the fitted model, construct a distribution of deviations and select appropriate class boundaries (Figure 5.6b).
4. Assess status of a sample by calculating the deviation of the observed values from an expected value calculated according to the model (Figure 5.6b).

Figure 5.6: Schematic illustration of status assessment using deviations from a model with a fixed factor: a. fitted model (solid line) of indicator and environmental factor with reference (dotted), good/moderate boundary (dashed) and standard deviation; b. cumulative distribution of deviations with class-boundaries. "x" and "y" are observations on the indicator and residual scale which are classified as good and moderate respectively.



Designing monitoring programmes with the desired precision and minimised cost

Apart from reducing uncertainty by accounting for predictable sources of variation, estimates of the variability due to random components can be used to model and predict the overall uncertainty. By inserting estimates of variability into the formulae for error propagation, it is possible to calculate the expected overall uncertainty under a range of monitoring scenarios (Figure 5.7). This is particularly useful when designing monitoring programmes in order to ensure that programmes achieve the desired precision in accounting for spatial and temporal scales relevant to the WFD, i.e. within water bodies and assessment periods, which is affected by several components of variability. Such methods can be used to allocate samples among stations and years in order to minimise uncertainty and costs. Several analyses within WATERS have indicated that this trade-off between spatial and temporal representativity has fundamental consequences for the uncertainty of status assessments for many BQEs.

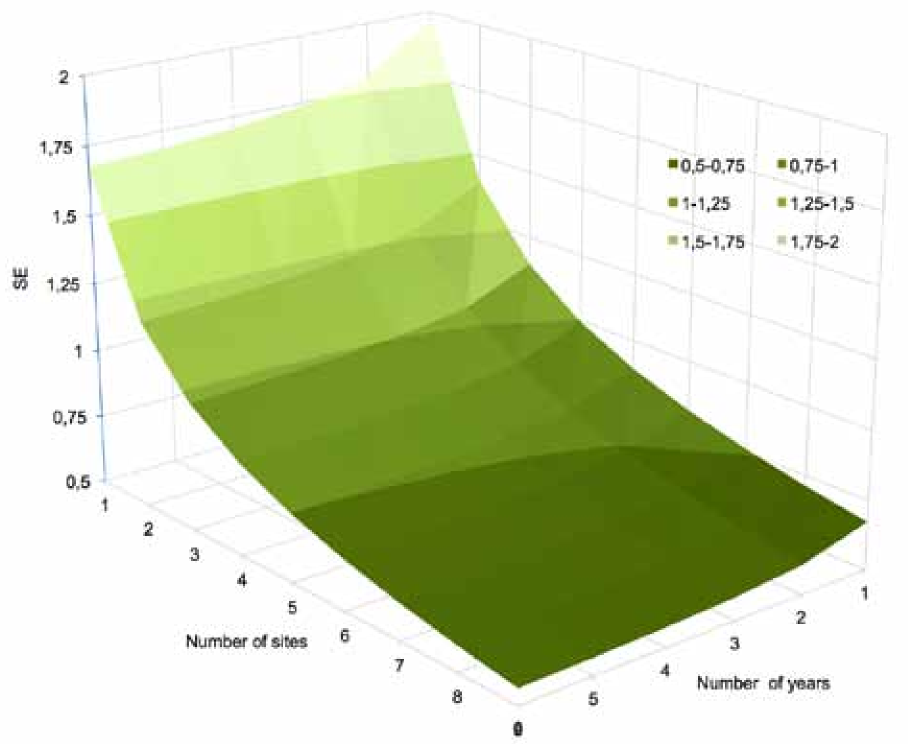


Figure 5.7: Example of modelling overall uncertainty (i.e. SE) for a varying number of sampling years and sites. Data represent uncertainty of BQI on the Swedish west coast (estimates of variability from Bergström and Lindegarth 2016). Number of samples per station and year, $n = 1$.

5.3.3 Uncertainty assessment when monitoring requirements are not fulfilled

The current Swedish assessment criteria define monitoring requirements for each of the BQEs (SwAM 2013). These requirements are designed to ensure sufficient quality in terms of precision and representativity. WATERS has documented that the precision of monitoring according to these guidelines differs among BQEs and that the precision is generally better at the scale of six-year assessment periods than within individual years (≈5 – 20% and 5 – 45% of the mean within assessment periods and years, respectively; Bergström and Lindegarth 2016).

Table 5.4. Number of assessed coastal and inland water bodies complying with requirements in the Swedish assessment criteria, determined by expert judgement and not assessed for each BQE. Data from VISS (preliminary assessment in 2015). Table modified from Lindegarth et al. (2016).

Habitat	BQE	No. assessed using complete data	No. assessed using expert methods	No. not assessed
Coastal	Benthic fauna	103	248	302
	Macrophytes	61	152	440
	Phytoplankton	109	425	119
Lakes	Benthic fauna	94	39	6873
	Macrophytes	196	75	6725
	Phytoplankton	347	156	6498
	Fish	364	53	6586
Streams	Benthic fauna	580	307	12960
	Fish	656	222	12969
	Phytobenthos	761	0	1386*

** For 5101 streams VISS reports that an expert assessment has been made but no assessment is given.*

Aggregation of data can make up for incomplete monitoring

Despite the guidelines, it is clear that status assessments of many inland and coastal water bodies are based on data that are not fully compliant with the monitoring requirements (Table 5.4; Lindegarth et al. 2016). For example, data may have been sampled at fewer stations or times than prescribed and the sampling scheme may be unbalanced in various ways (e.g. different numbers of samples or sites per year). As with compliant data, the proposed framework for uncertainty assessment provides a way to use such incomplete data for status assessment, including the estimation of overall uncertainty (Carstensen and Lindegarth 2016). The uncertainty of a status assessment based on incomplete monitoring is greater than that based on

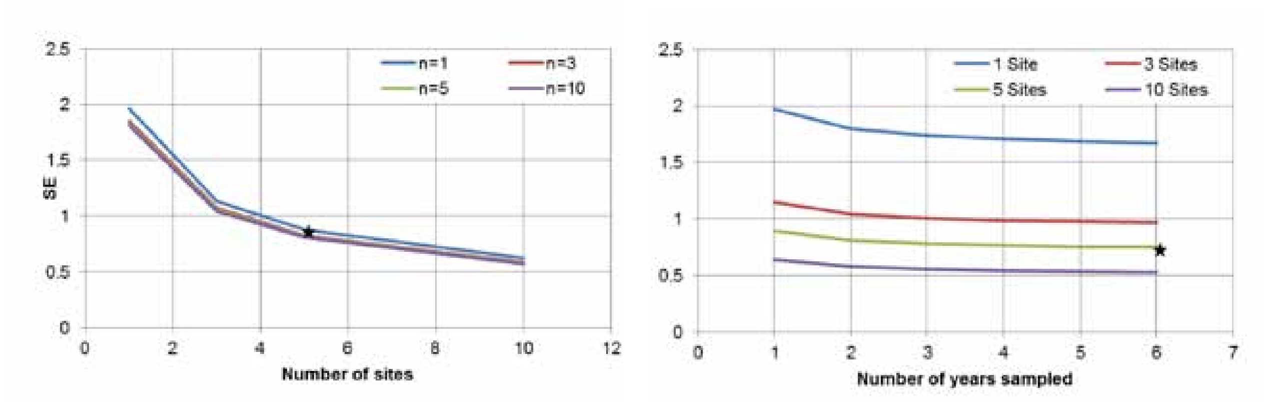


Figure 5.8: Illustration of the overall uncertainty of the estimated status of a coastal water body, i.e. BQI on the Swedish west coast. Stars indicate the monitoring design prescribed in the assessment criteria. The overall standard error (SE) within years (left panel) increases substantially if fewer than three stations are monitored (n=number of samples per site and year) and the SE within assessment periods (right; n=1) increases marginally when samples are taken only at 1–2 years.

recommended protocols, but because the data are aggregated across sites and years, they often have a precision comparable to that obtained with the recommended protocol (Figure 5.8).

A lack of appropriate spatial and temporal replication means that uncertainty components cannot be estimated. For example, if the status of one water body is based on one site only, variability among sites cannot be estimated. Ignoring such spatial variability often results in biased estimates of status and underestimation of the uncertainty. WATERS has developed a methodology in which relevant sources of variability can be accounted for even if they cannot be estimated accurately from data. This may be because the structure of the sampling programme does not allow the estimation of variance components or because insufficient replication causes any estimates to be very uncertain themselves (Lindegarth et al. 2013b, Bergström and Lindegarth 2016).

Existing datasets used to estimate relevant spatial and temporal variability

The methodology is based on the general framework for uncertainty assessment (Carstensen and Lindegarth 2016), which allows for the incorporation of uncertainty components estimated using large datasets. Such estimates should be determined with high precision and be representative of a broad range of spatial and temporal contexts. In fact, in many instances, it is arguably preferable to assess uncertainty using tabled estimates of variance components (determined from large datasets) rather than to estimate uncertainty from a small sample size within the individual water body. To provide practical tools for assessing uncertainty when sample sizes are small or when the monitoring does not allow the estimation of all uncertainty components, we have used large datasets from Swedish monitoring to estimate the relevant spatial and temporal variability for all indicators used in the current Swedish assessment criteria (Bergström and Lindegarth 2016). These estimates are compiled in tables ("uncertainty libraries") and combined with appropriate formulae for error propagation, ensuring that all relevant sources of uncertainty are properly accounted for (Bergström and Lindegarth 2016; Carstensen and Lindegarth 2016).

5.3.4 Predicting status and assessing uncertainty for unmonitored water bodiess

Swedish surface waters: a challenge for monitoring and assessment

Assessing the ecological status of the surface waters of Sweden, whose 2400-km coastline has a strong salinity gradient and is divided into approximately 650 coastal water bodies, and whose inland waters comprise approximately 7000 lakes and 13000 stream water bodies, is challenging. It is not surprising that not all water bodies can be rigorously monitored for all BQEs and that assessments may need to be made based on expert judgement methods or remain not assessed for individual BQEs (Table 5.4).

Uncertainty can also be assessed for expert judgement methods

In practice, various expert judgement approaches are used, approaches that vary among BQEs and routines in different river basin districts, such as (1) non-quantitative evaluations from incomplete data, (2) spatial extrapolation of data from neighbouring water bodies, and (3) various alternative monitoring methods, indicators, pressure analyses, and other proxies. None of these methods has been rigorously documented nor have appropriate approaches for quantifying their uncertainty been developed. As described above, WATERS has developed and tested methods for uncertainty assessment using complete and incomplete data, and has extended the framework for spatial extrapolation (Lindegarth et al. 2016).

Assessing BQEs by spatial extrapolation, in which the status of unmonitored water bodies is estimated in line with the measured status of one or several neighbouring water bodies, is a common strategy for some BQEs, and one that has been suggested as an alternative method in the WFD and Swedish assessment criteria.[1] Such grouping of water bodies can only be done within water body types and when water bodies are exposed to similar pressures. Formally, the state of the target water body is predicted by the observed value of the group, i.e. $\bar{y}_{target} = \bar{y}$: and the uncertainty of the predicted value can be estimated as:

$$V\left[\bar{y}_{target}\right] = \frac{s^2_{WB} * \left(1 - \frac{b}{G}\right)}{b} + V[\bar{y}] \qquad (4)$$

where $V[\bar{y}]$ is the variability within water bodies, s^2_{WB} is the variability among water bodies, b is the number of water bodies used to calculate the mean, and G is the total number of water bodies in the group. The uncertainty of assessments based on extrapolation therefore depends on the variability within and among the water bodies, and also on the sampled proportion of the group. With this expression, it is possible to evaluate the potential of grouping and extrapolation for different BQEs and types, and to compare their uncertainty in relation to that of monitored water bodies (Figure 5.9). An example of benthic fauna from the Swedish west coast indicates that the uncertainty of extrapolated assessments is always greater than

1) Classification of ecological status in individual water bodoes shall be based on integration of the biological quality elements. The classification can also be based on data from a group of water bodies". Translated from SwAM (2013, p. 4).

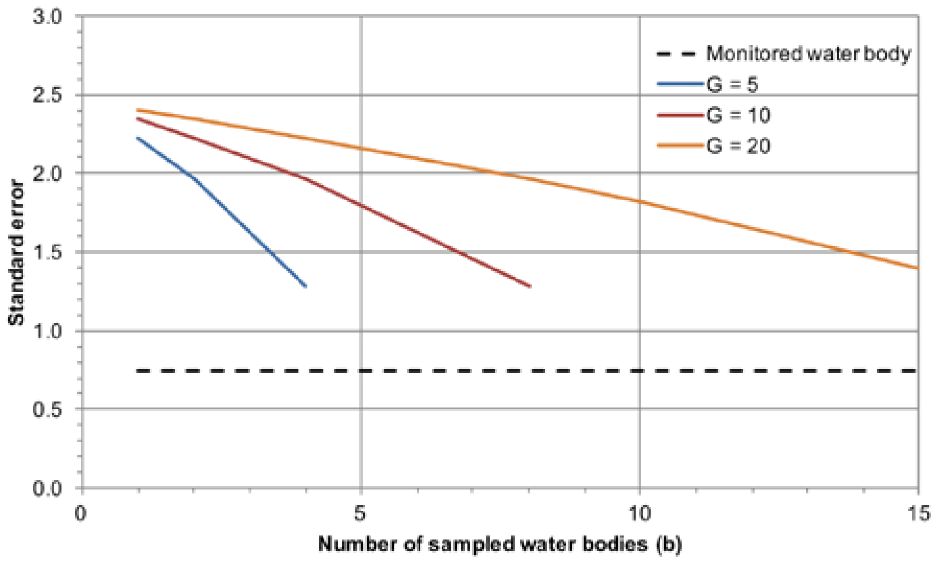


Figure 5.9: Expected uncertainty of mean BQI in a water body during a six-year period using monitoring and extrapolation from b monitored water bodies from a group of *G* water bodies. Number of years sampled, $a = 6$, number of sites per water body, $c = 5$, and number of samples per site and year, $n = 1$.

that of monitored water bodies, but that these estimates are useful in overall assessments. Assuming a mean BQI of 10, the worst-case scenario produces an error of 25% of the mean, which will gradually decrease towards 10% as the number of sampled water bodies increases.

This example suggests that the status of water bodies may be assessed with some confidence but that the most promising approaches likely vary among BQEs. Other analyses have indicated that the relationships between status and alternative indicators or other proxies can be used in a similar way to predict status, but that to assess uncertainty, it is important that formal, quantitative relationships be established. Finally, it is worth pointing out that uncertainties and errors estimated in this way are fully comparable to those estimated directly from monitoring data.

In summary, this section has described a coherent strategy for assessing uncertainty and status at the scale of water bodies within six-year assessment periods, using monitoring data and when assessments are made using expert methods. This framework can be used for all conceivable indicators and BQEs and is fully compatible with the tool for integrated assessment.

5.3.5 Towards a coherent estimation of status and uncertainty

This chapter summarises WATERS efforts to develop general methods addressing overarching issues applicable to all biological quality elements, and possibly also chemical and physical elements, to achieve a coherent and harmonised assessment process. The driver of such efforts is a desire to produce more predictable, transparent, and reliable status assessments based on sound scientific principles. Overall, the development of harmonised methods has been successful in several areas, with a common framework for uncertainty treatment and a method and tool for integrated assessment being clear products. Other areas, however, have been deemed less prone to full harmonisation. In particular, the preconditions for defining reference values and class boundaries have been considered so different among inland and coastal areas, and to some extent among BQEs, that the complete harmonisation of methods has not been desirable. This is largely because of differences in ecological processes among systems and due to differences in knowledge of historical conditions among BQEs. Thus, in this respect the explicit aim of achieving harmonised methods has been given a lower priority in favour of ecological considerations. Notwithstanding these considerations, it is also true that important aspects of harmonisation with respect to the definition of class boundaries are achievable by adopting the proposed transformation to a common EQR scale (Figure 5.4). This method allows class boundaries to be set according to the best available ecological knowledge of each indicator and BQE, while providing a harmonised value of the status that can easily be communicated to stakeholders. In summary, we believe that these results represent significant components of improved harmonisation and integration among BQEs. If they are implemented in the management process, we expect that they will contribute to a more predictable, transparent, and reliable status assessment.

6. Beyond WATERS: future challenges for research and water management

Mats Lindegarth

6.1 Objectives and achievements

The aim of WATERS has been to develop high-quality research in order to provide an improved scientific basis and practically useful assessment methods for the management of Swedish inland and coastal waters. The primary concern has been to improve assessment criteria for the biological quality elements within the Water Framework Directive (WFD); in particular, two general aims have been pursued: (1) the development of reliable and indicators that are sensitive to human disturbances and (2) the development of harmonised methods for the integrated assessment of quality elements.

Improved indicators for the BQEs in inland as well as coastal waters

Depending on the needs identified, WATERS has refined and tested existing indicators or developed new indicators for all biological quality elements (BQEs) in inland and coastal waters. Because the conditions differ strongly with respect to the national and international status of the indicators, availability of data, status of monitoring techniques, relevant human pressures, etc., the specific tasks and approaches have varied among BQEs. WATERS has used extensive databases and synoptic studies of responses to defined pressure gradients in lakes and streams around Sweden and in coastal waters in the Baltic and Skagerrak to suggest improved indicators for all BQEs. While some of these indicators are more or less operational, others are less mature. This may be because responses to pressures, and therefore class boundaries, have not been widely established or because the use of indicators requires that new techniques and sampling programmes be adopted for monitoring. Nevertheless, in agreement with the general aims, the research within WATERS has made substantial progress towards improved indicators for the BQEs in inland as well as coastal waters.

Assessment system easier to understand and communicate with improved integration

Similarly, the work on developing harmonised methods for integrated assessment has resulted in a set of generally applicable methods and routines. Despite the fact that methods for defining reference conditions and class boundaries need to be selected individually for BQEs (due to differences in the availability of historical data or data from undisturbed sites), the suggested transformation to a common EQR scale allows all indicators and BQEs to be aggregated and compared in a harmonised way. The adoption of common class boundaries for all BQEs will also help make the assessment system more comprehensible and easier to communicate. Furthermore, the development of a general framework for assessing status and uncertainty, applicable to all BQEs and at spatial and temporal scales that correspond to the WFD assessment, provides a significant advance over previous routines. Finally, the specifically developed user-friendly tool for integrated assessment involving EQR transformation, for assessing classification uncertainty for individual indicators and BQEs, and for assessing overall ecological status offers much-improved opportunities for standardised and transparent status assessment.

Thus, within the framework of the WFD, WATERS has successfully addressed the overall objectives of indicator development and methodological harmonisation. During this work, however, several general challenges partly originating from the prerequisites of the Directive have become evident and will continue to challenge future work to devise progressively more reliable and efficient assessment criteria.

6.2 Scientific lessons and challenges

Ecological systems, particularly biological assemblages of microscopic organisms, animals, and plants in aquatic environments, are inherently variable at a range of spatial and temporal scales. This has caused a number of theoretical and practical challenges throughout the programme at the same time as it holds the key to developing sensitive indicators and reliable assessment systems. Our ability to account for natural variability has influenced the work in several ways, two aspects of which deserve particular attention.

Important to account for natural variability within water body types

First, central to the WFD is the partitioning of inland and coastal waters into a system of water body types. These water body types are administrative units, but they are also defined based on ecoregions, physical properties, and chemical properties in order to reflect ecologically relevant units. For each of these types, specific reference values and class boundaries are defined for each individual indicator. This is a requirement of the WFD, but one difficulty is of course that much ecological variability is caused by processes occurring at smaller spatial scales and by processes not coupled to the factors involved in defining the typology. Biological assemblages, including those related to the BQEs, are therefore often highly variable within the defined types. This means that type-specific reference values and class boundaries are at best representative of the average state of that type and that assessments based on these benchmarks risk being very uncertain. This issue has been a general challenge for all water body types and all BQEs throughout the programme. The good news, however, is that with an appropriate analytical framework, ecological knowledge, and relevant data, it is possible to account for variability caused by small-scale processes and thus reduce the uncertainty in status assessments. There are examples of BQEs in the existing Swedish assessment criteria (e.g. fish in lakes and streams and phytoplankton in coastal waters; SwAM 2013) that use various modelling approaches to adjust for site-specific factors. The need for methods to decrease uncertainty by accounting for site-specific information is one of the most striking and general themes throughout the programme (see Section 5.1; Johnson et al. 2014). This has been manifested in the indicator development work, for example, on benthic invertebrates and macrovegetation in coastal waters (Leonardsson et al. 2016; Blomqvist et al. 2014) and on benthic invertebrates in inland waters (see Section 5.1). The importance of accounting for site-specific factors has also been important for the harmonisation of assessment methods across BQEs, so the uncertainty framework explicitly involves the possibility of accounting for predictable spatial (e.g. continuous gradients or categorical factors) and temporal (e.g.

seasonal fluctuations) variability due to fixed effects (e.g. Lindegarth et al. 2013; Carstensen and Lindegarth 2016) In summary, despite the fact that the Directive and the formal reporting require that the Member States define type-specific references and class boundaries, the experiences and methods developed within WATERS clearly indicate that accounting for site-specific factors is possible and worthwhile when assessment criteria are refined, not only for the WFD assessments in Sweden but probably also in a broader context.

Natural variability obscures responses to pressures

A second challenge posed by the natural variability of ecological systems involves the understanding of how assemblages and indicators respond to human pressures. The development of quantitative relationships between indicators and pressures is fundamental to validating the performance of indicators and to establishing class boundaries. WATERS work on indicator development, using existing data from monitoring programmes and from the synoptic gradient study of inland and coastal waters, has repeatedly demonstrated that spatial and temporal variability caused by factors other than the targeted pressure gradient often makes it difficult to detect and quantify any such relationships (responses may also be confounded by additional anthropogenic factors). This is not surprising, but nevertheless it must be stressed that addressing the complexity of natural variability is fundamental to the development of indicators and assessment criteria. Access to supporting data that can explain substantial parts of the natural variability and the adoption of flexible analytical frameworks capable of involving multiple explanatory variables (including their interactions) hold the key to such efforts.

Assessing status under climate change

The aim of the WFD is to prevent deterioration and achieve good ecological status in European surface waters by 2027. Good ecological status, according to the normative definition of the WFD, means that biological quality elements "show low levels of distortion resulting from human activity, but deviate only slightly from those normally associated with the surface water body type under undisturbed conditions" (Annex V, p. 45). In principle according to the Directive, if the status is classified as being below "good" (or if it is deteriorating from "high" to "good"), actions are to be taken to restore the status. A recent ruling in the EU court[2] has demonstrated that this principle has legislative validity even for the risk of deterioration of individual BQEs (Paloniitty 2016). Assessments of overall ecological status, and of individual BQEs, are used to support decisions about programmes of measures and to identify which measures should be taken to restore the status. This principle is reasonably straightforward when human impacts are localised and reversible. The growing evidence of global climate change, however, is manifested as large-scale changes of the physical and chemical environment accompanied by complex and potentially irreversible changes to ecological systems (e.g. Folke et al. 2004). Such impacts cannot likely be remedied by measures at national or even European scales, so the intended coupling between status assessments and program-

2) http://curia.europa.eu/juris/document/document.jsf?doclang=EN&text=&pageIndex=0&part=1&mode=DOC&docid=165446&occ=first&dir=&cid=778485

mes of measures defined by the Directive no longer holds. The managerial implications and extent of the impacts of this troublesome development remain to be seen. Nevertheless, well-functioning and scientifically sound assessment criteria based on the best available knowledge of historical conditions (i.e. reference values) and biological responses to human pressures (i.e. class boundaries) still provide invaluable knowledge of the structure of well-functioning ecological systems. Furthermore, by developing a capacity to predict what types of changes are caused by climate change, improved scientific understanding can contribute to the design of efficient remedial actions when other anthropogenic disturbances are to blame (e.g. Johnson et al. 2010).

6.3 Impacts on water management

The research developed within WATERS can arguably contribute to improving the Swedish biological assessment criteria for the WFD in inland and coastal waters. This will improve the scientific basis for decision making, having positive effects on water management and ultimately on the structure and functioning of Swedish and European waters. There are, however, a number of challenges to be addressed and steps to be taken to maximise the impact of these results in a managerial context. These steps provide a logical extension of the research within WATERS and must involve both managers and researchers.

Improving status assessments

The first and most obvious area of application of WATERS results is in developing a harmonised and transparent assessment system, with state-of-the-art indicators and user-friendly tools, that can be used by the responsible agencies and managers. The WATERS tool for integrated assessment (see Section 5.2) provides a promising, functional framework fulfilling the

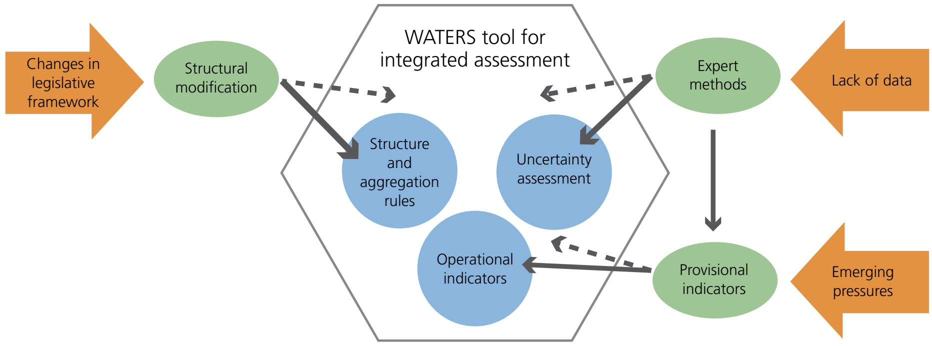


Figure 6.1: Schematic of how the WATERS assessment tool can form the core of an adaptive system in which new demands and advances in assessment methods can be evaluated within the framework and ultimately incorporated into the system.

requirements stipulated by the Directive, meeting demands for user-friendliness, and providing transparent assessments at hierarchical levels. Therefore, we propose that it be adopted as a common tool for status assessment in Swedish surface waters (Figure 6.1). This tool can then form the core of the assessment system, incorporating routines and information on uncertainties (see Section 5.3) and, due to its modular structure, progressively more sensitive and appropriate indicators as these become fully operational (Chapters 3 and 4).

Progressive implementation in a robust framework

Apart from the implementation of this overarching framework, additional steps remain to be taken before scientific advances within WATERS are fully operational for practical status assessment. One main objective will be to finalise and test indicators. The extent and content of this work varies among BQEs, but for a number of new indicators, these tasks include defining reference values and class boundaries, collecting data using novel sampling methods and monitoring programmes, and determining uncertainties using the uncertainty framework. Nevertheless, the modular system provided by the framework for integrated assessment allows the progressive incorporation of new indicators following analyses of the consequences for overall status assessments (Figure 6.1). Because the tool for integrated assessment is transparent at different hierarchical levels and fully integrated with the uncertainty framework, it also provides estimates of confidence in classification for individual indicators and BQEs. As discussed above, this is potentially very useful in the light of recent rulings in the EU court.

Additional guidance needed

Finally, it is worth stressing that although the WATERS tool for integrated assessment provides a framework for automated integrated assessment using many indicators and BQEs, it is not intended to be used as a "black box". All assessments need to be manually and critically assessed in terms of their realism and general agreement with what is expected by informed people. This is not to say that evidence based on data can easily be discarded, but it is important that any status assessment, good or bad, can be understood and carefully scrutinised.

Develop tools, routines, and data flows for estimating status and uncertainty

In addition, the proposed general framework for uncertainty assessment will make significant contributions to improving status assessments. Note that not only does it provide a method for uncertainty assessment but, importantly, it also defines a methodology for correctly estimating the state of the indicators (mean ±SE) when data are sampled over several years, seasons, sites, cores, etc., and when supporting variables are used to reduce uncertainty. This is a substantial advance over existing Swedish guidance documents and assessment practices. Otherwise, important features of the uncertainty framework directly addressing water management needs are: (1) methods to determine uncertainty at the scale of six-year assessment periods (or any other spatial and temporal resolution desired), (2) methods to determine confidence in classification, and (3) methods to estimate the uncertainty of assessments using expert methods, such as those based on incomplete data or spatial extrapolation. It is our conviction that all these properties and potentials can help improve the scientific basis of status assessments and of the decision support for programmes of measures, which are all central to the water management cycle. This is of course dependent on successful implementation of the estimation routines. Using available data from national data hosts, WATERS has developed an "uncertainty library" and, in combination with the proposed methodology, this provides a useful start (Bergström and Lindegarth 2016). Nevertheless, if these methods are to be used routinely, they need to be implemented in a system that has a functioning data flow and appropriate user-friendly tools. Development of such routines was not within the scope and mandate of WATERS and requires collaboration with all agencies involved and with the existing data infrastructure. Nevertheless, it is crucial for the successful implementation of these methods.

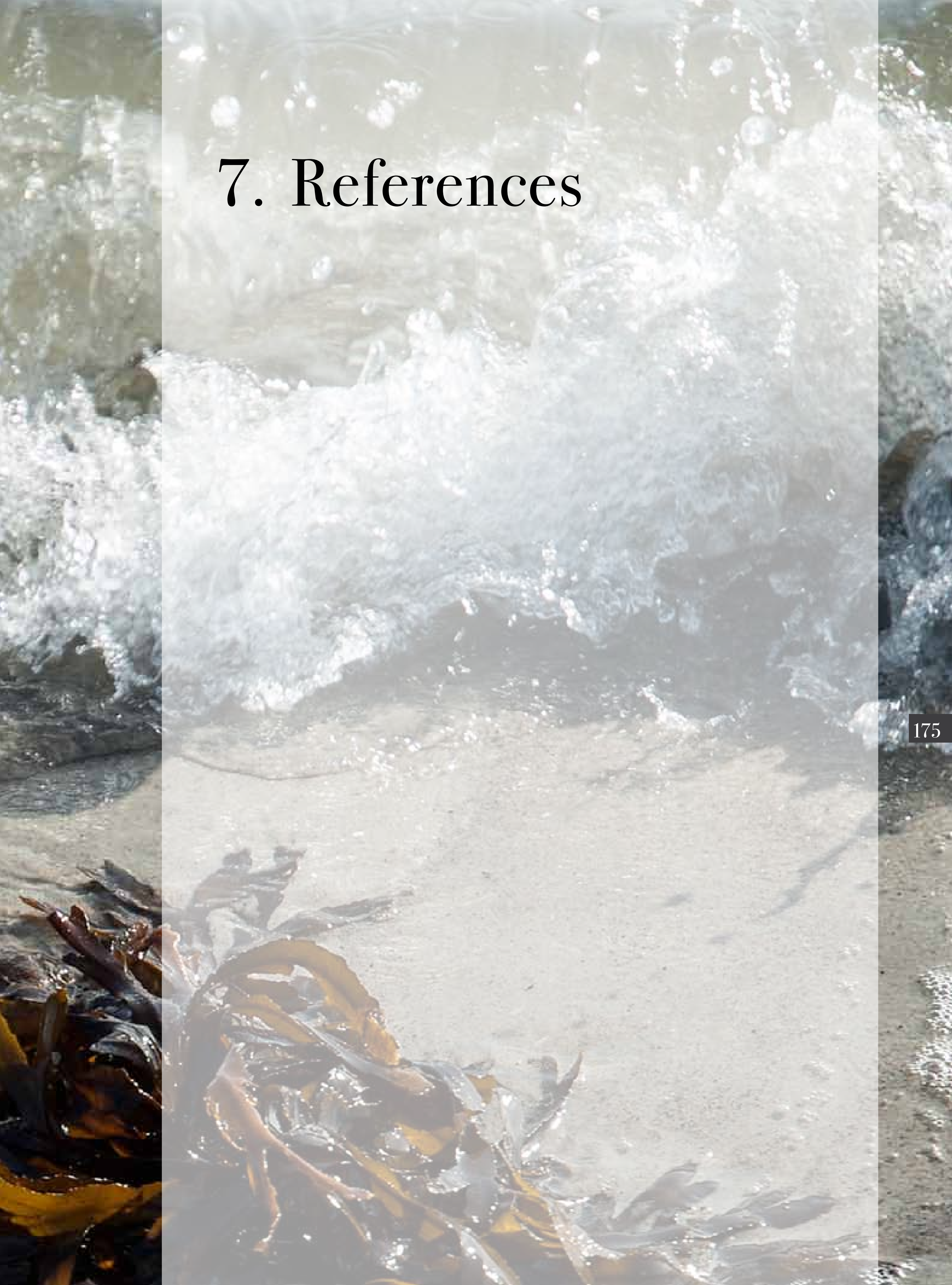
Use the framework to design efficient monitoring programmes

A second area where the WATERS results can make important contributions to water management is in sampling techniques and monitoring programmes, where the uncertainty framework can be used to make improvements. Because the methods of the framework can be used to estimate the precision of estimated means at any relevant spatial and temporal scale, the uncertainty can always be estimated *a posteriori* and predicted *a priori*. Thus informed decisions about precision are always possible, which will help provide a more solid empirical basis for status assessments and related contexts in water management. Not only do such tools have the potential to identify situations in which the data may be imprecise, but they also permit the identification of situations in which little additional precision will be gained by adding more samples, locations, or sampling occasions. The method can also be used to reallocate sampling efforts among scales, areas, or years to achieve the most cost-effective and comprehensive monitoring programmes.

Harmonise beyond the BQEs

Third, it is also quite likely that the WATERS results can contribute greatly to assessment routines in other contexts, for example, other quality elements within the WFD and other Directives such as the Marine Strategy Framework Directive (MSFD) and the Habitats Directive. For example, many indicators developed for coastal waters can also be used to assess status according to the MSFD. Another example is the uncertainty framework, which has successfully been used to assess monitoring designs for the MSFD. Perhaps the most obvious extension is to use the framework for assessing and monitoring designs for the chemical and physical quality elements within the WFD. The opportunities are abundant, but should be seized one at a time.

7. References



Allan JD, 2004. Landscapes and riverscapes: The Influence of Land Use on Stream Ecosystems. Annual Review of Ecology. Evolution and Systematics 35: 257-84

Andersen JH, Aroviita J, Friberg N et al, 2016. Approaches for integrated assessment of ecological and eutrophication status in Nordic countries. AMBIO. http://link.springer.com/article/10.1007/s13280-016-0767-8

Andersson A, Hajdu S, Haecky P et al, 1996. Succession and growth limitation of phytoplankton in the Gulf of Bothnia. Marine Biology. 126: 791-801.

Andersson A, Falk S, Samuelsson G and Hagström Å, 1989. Nutritional characteristics of a mixotrophic nanoflagellate, Ochromonas sp. Microb. Ecol. 17: 251-262.

Andersson A, Höglander H, Karlsson C and Huseby S, 2015. Key role of phosphorus and nitrogen in regulating cyanobacterial community composition in the northern Baltic Sea. Estuarine, Coastal and Shelf Science 164:161-171.

Andersson J, 2015. Undersökningstyp: Provfiske med kustöversiktsnät, nätlänkar och ryssjor på kustnära grunt vatten. Havs- och vattenmyndigheten. Programområden Kust och Hav Version 1:1 2015-07-08.

Andrén C and Jarlman A, 2008. Benthic diatoms as indicators of acidity in streams. undamental and Applied Limnology 173(3): 237-253.

Angeler DG, 2007. Resurrection ecology and global climate change research in freshwater ecosystems. Journal of the North American Benthological Society, 26: 12-22.

Angeler DG, Viedma O, Sánchez-Carrillo S and Alvarez-Cobelas M, 2008. Conservation issues of temporary wetland Branchiopoda (Anostraca, Notostraca: Crustacea) in a semiarid agricultural landscape: What spatial scales are relevant?. Biological Conservation, 141, 1224-1234.

Argillier C, Caussé S, Gevrey M et al, 2013. Development of a fish-based index to assess the eutrophication status of European lakes. Hydrobiologia, 174:193-211.

Armitage PD, Moss D, Wright JF and Furse MT,1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. Water research, 17: 333-347.

Aroviita J, Mykrä H, Muotka T and Hämäläinen H, 2009. Influence of geographical extent on typology- and model-based assessments of taxonomic completeness of river macroinvertebrates. Freshwater Biology, 54: 1774-1787.

Arranz I, Mehner T, Benejam L et al, 2015. Density-dependent effects as key drivers of intraspecific size structure of six abundant fish species in lakes across Europe. Canadian Journal of Fisheries and Aquatic Sciences, 72:1-16, dx.doi.org/10.1139/cjfas-2014-0508.

Arwidsson T, 1926. Studier över sjöarnas vegetation i Lilla Lule älvs vattenområde. Arkiv för botanik 20A: 1-31.

Baattrup-Pedersen AB, Andersson BTK, Karttunen T et al 2001. Macrophytes. Biological Monitoring in Nordic rivers and lakes. Nordic Council of Ministers, TemaNord 2001: 513, ed Skriver J.

Baattrup-Pedersen A and Riis T, 1999. Macrophyte diversity and composition in relation to substratum characteristics in regulated and unregulated Danish streams. Freshwater Biology, 42: 375-385.

Bain MB, Finn JT and Booke HE, 1988. Streamflow regulation and fish community structure. Ecology, 69(2): 382-392.

Baker ME and King RS, 2010. A new method for detecting and interpreting biodiversity and ecological community thresholds. Methods in Ecology and Evolution, 1: 25-37.

Balsby JS and Holmgren K. Sources of uncertainty in 6-year cycles of monitoring lake fish abundance. Manuscript.

Battarbee RW, Flower RJ, Juggins S et al, 1997. The relationship between diatoms and surface water quality in the Høylandet area of Nord-Trøndelag, Norway. Hydrobiologia, 348: 69-80.

Beier U, Degerman E, Sers B et al, 2007. Bedömningsgrunder för fiskfaunans status i rinnande vatten – utveckling och tillämpning av VIX. Environmental quality criteria to determine the status of fish in running waters – development and application of VIX Fiskeriverket Informerar, 2007:5, 59 p (In Swedish with English summary)

Bergström L, Carstensen J, Bergström U and Olsson J. Coastal fish indicator responses to natural and anthropogenic drivers – variability at temporal and different spatial scales. Submitted manuscript.

Bergström L and Karlsson M, 2016a. Undersökningstyp: Djupstratifierat provfiske med småryssjor. Havs- och vattenmyndigheten. Programområden Kust och Hav Version 1: 0, 2016-02-10.

Bergström L, Karlsson M, Bergström U et al. 2016b Distribution of mesopredatory fish determined by habitat variables in a predator-depleted coastal system. Marine Biology, 163:201.

Bergström L, Karlsson M and Pihl L, 2013. Comparison of gill nets and fyke nets for the status assessment of coastal fish communities. WATERS Report no. 2013:7. Havsmiljöinstitutet, Sweden.

Bergström L and Olsson J, 2015. Coastal fish community indicators in Sweden - variation along environmental gradients. WATERS Report no. 2015: 1.

Bergström P and Lindegarth M, 2016. Developing practical tools for assessing uncertainty of Swedish WFD indicators: A library of variance components and its use for estimating uncertainty of current biological indicators. WATERS Report no. 2016:2. Havsmiljöinstitutet, Sweden.

Birk S, Bonne W, Borja A et al, 2012. Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators, 18: 31–41

Birk S and Ecke F 2014. The potential of remote sensing in ecological status assessment of coloured lakes using aquatic plants. Ecological Indicators 46:398-406.

Blomqvist M, Börjesson D and Olsson P 2012b. Djuputbredning av ålgräs. Metoderfarenheter från fältstudier kring Orust och Tjörn 2012. http://waters.gu.se/digitalAssets/1543/1543120_waters-gradientstudie-metodik-a--lgra--s-va--stkusten-2012.pdf

Blomqvist M, Wikström SA, Carstensen J et al, 2014. Response of coastal macrophytes to pressures. WATERS Report no. 2014:2 Havsmiljöinstitutet, Sweden.

Blomqvist M, Cederwall H, Leonardsson K and Rosenberg R, 2007. Bedömningsgrunder för kust och hav – Bentiska evertebrater. Commissioned report to Swedish Environmental Protection Agency.

Blomqvist M, Krause-Jensen D, Olsson P et al, 2012a. Potential eutrophication indicators based on Swedish coastal macrophytes. WATERS Report no. 2012:2. Havsmiljöinstitutet, Sweden.

Blomqvist M and Leonardsson K, 2016. A probability based index for assessment of benthic invertebrates in the Baltic Sea. WATERS Report no. 2016:3. Havsmiljöinstitutet, Sweden.

Boesch D, Heckey R, O'Meila C et al, 2006. Eutrophication of Swedish seas. Swedish Environmental Protection Agency Report 5509. www.naturvardsverket.se.

Bolker BM, Brooks ME, Clark CJ et al, 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends. Ecol. Evol. 24(3):127-135

Bonada N, Prat N, Resh VH and Statzner B, 2006. Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. Annual Review of Entomology 51, 495–523.

Borja A, Franco J and Pérez V, 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Mar. Pollut. Bull. 40, 1100–1114.

Borja A, Prins T, Simboura N et al, 2014. Tales from a thousand and one ways to integrate marine biodiversity components when assessing the environmental status. Frontiers in Marine Science. http://journal.frontiersin.org/Journal/10.3389/fmars.2014.00072/abstract.

Brattberg G, 1986. Decreased phosphorus loading changes phytoplankton composition and biomass in the Stockholm archipelago. Vatten, 42:141-153.

Brauns M, Garcia XF, Walz N and Pusch MT, 2007. Effects of human shoreline development on littoral macroinvertebrates in lowland lakes. Journal of Applied Ecology, 44: 1138–1144.

Breiman L, 2001. Random forests. Machine learning, 45, 5-32.

Brinkhurst RO, 1974. The benthos of lakes. – Blackburn Press, Caldwell, NJ, USA.

Brucet S, Pédron S, LauridsenTL et al, 2013. Fish diversity in European lakes: geographic predictors dominate over anthropogenic pressures. Freshwater Biology, 58:1779-1793.

Brucet S, Poikane S, Lyche-Solheim and Birk S, 2013. Biological assessment of European lakes: ecological rationale and human impacts. Freshwater Biology, 58: 1106–1115.

Brusven MA, Meehan WR and Biggam RC, 1990. The role of aquatic moss on community composition and drift of fish-food organisms. Hydrobiologia 196:39-50.

Brönmark C and Hansson LA, 2005. The Biology of Lakes and Ponds. Second Edition. Oxford University Press Inc, New York. 285 pp.

Buffagni A, Erba S, Cazzola M and Kemp JL, 2004. The AQEM multimetric system for the southern Italian Apennines: assessing the impact of water quality and habitat degradation on pool macroinvertebrates in Mediterranean rivers. In Integrated Assessment of Running Waters in Europe (pp. 313-329). Springer Netherlands.

Burrell TK, O'Brien JM, Graham SE et al, 2014. Riparian shading mitigates stream eutrophication in agricultural catchments. Freshwater Science, 33: 73-84.

Carlson P, Johnson RK and McKie BG, 2013. Optimizing stream bioassessment: habitat, season, and the impacts of land use on benthic macroinvertebrates. Hydrobiologia, 704: 363-373.

Carlson PE, McKie BG, Sandin L and Johnson RK, 2016. Strong land-use effects on the dispersal patterns of adult stream insects: implications for transfers of aquatic subsidies to terrestrial consumers. Freshwater Biology, doi: 10.1111/fwb.12745.

Carstensen J and Lindegarth M, 2016. Confidence in ecological indicators: A framework for quantifying uncertainty components from monitoring data. Ecol. Ind. 67: 306-317

Carstensen J and Heiskanen AS, 2007. Phytoplankton responses to nutrient status: application of a screening method to the northern Baltic Sea. Marine Ecology Progress Series 336: 29-42.

Carstensen J, 2015. Baltic GIG justification revision. Report to European commission.

Carvalho L, Poikane S, Lyche Solheim A et al, 2012. Strength and uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes. Hydrobiologia 704 (1): 127-140.

Caussé S, Gevrey M, Pédron S et al, 2011. WISER Deliverable 3.4-4: Fish indicators for ecological status assessment of lakes affected by eutrophication and hydromorphological pressures. http://www.wiser.eu/download/D3.4-4.pdf.

CEMAGREF 1982, Etude des méthodes biologiques d´appréciation quantitative de la qualité des eaux,, Rapport Division Qualité des Eaux Lyon-Agence Financière de Bassin Rhône-Méditerranée-Corse: 218 p.

CEN 2003. Water quality: guidance standard for the routine sampling and pretreatment of benthic diatoms from rivers. EN 13946:2003. Comité Européen de Normalisation, Geneva, Switzerland.

CEN 2003. Water quality – sampling of fish with electricity. European standard. European Committee for Standardization. Ref. No. EN 14011:2003.

CEN 2004. Water quality: guidance standard for the identification, enumeration and interpretation of benthic diatom samples from running waters. EN 14407:2004. Comité Européen de Normalisation, Geneva,Switzerland.

CEN 2004. Water quality - Guidance standard for the surveying of macrophytes in lakes. CEN/TC 230/WG 2/TG 3/N80, European Committee for Standardisation.

CEN 2006. Water quality – guidance in the scope and selection of fish sampling methods. European standard. European Committee for Standardization. EN 14962:2006.

CEN 2014a. Water quality. Guidance for the routine sampling and preparation of benthic diatoms from rivers and lakes. EN 13946:2014. Comité Européen de Normalisation, Geneva, Switzerland.

CEN 2014b. Water quality. Guidance for the identification and enumeration of benthic diatom samples from rivers and lakes. EN 14407:2014. Comité Européen de Normalisation, Geneva, Switzerland.

CEN 2014. Water quality – guidance on the estimation of fish abundance with mobile hydroacoustic methods. European standard. European Committee for Standardization. EN 15910:2014.

CEN 2015. Water quality – sampling of fish with multi-mesh gillnets. European standard. European Committee for Standardization. EN 14757:2015.

Cermeño P, Rodriguez-Ramos T, Dornelas M et al, 2013. Species richness in marine phytoplankton communities is not correlated to ecosystem productivity. Marine Ecology Progress Series, Vol. 488:1-9.

Clarke RT, 2012. Assessing uncertainty of indicators of Water Framework Directive ecological status class. WISER deliverable 6.1-2. http://www.wiser.eu/download/D6.1-2.pdf

Clarke RT, Davy-Bowker J, Sandin L et al, 2006. Estimates and comparisons of the effects of sampling variation using 'national' macroinvertebrate sampling protocols on the precision of metrics used to assess ecological status. Hydrobiologia 566:477-503

Clarke RT and Hering D, 2006. Errors and uncertainty in bioassessment methods - major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia, 566:433-440.

Clarke RT, Furse MT, Gunn RJ et al, 2002. Sampling variation in macroinvertebrate data and implications for river quality indices. Freshwater Biology, 47: 1735-1751.

Clarke RT, Lorenz A, Sandin L et al, 2006. Effects of sampling and sub-sampling variation using the STAR-AQEM sampling protocol on the precision of macroinvertebrate metrics. Hydrobiologia, 566, 441-460.

Cochran W G 1977. Sampling techniques (Third ed.). Wiley. ISBN 0-471-16240-X.

Cronberg G and Annadotter H, 2006. Manual on aquatic cyanobacteria. A photo guide and synopsis of their toxicology. Intergovernmental Oceanographic Commission of UNESCO. International Society for the Study of Harmful Algae (ISSHA): 106 pp.

Dahl J, Johnson RK and Sandin L, 2004. Detection of organic pollution of streams in southern Sweden using benthic macroinvertebrates. Hydrobiologia, 516: 161-172.

Dale VH and Beyeler SC, 2001. Challenges in the development and use of ecological indicators. Ecological indicators, 1: 3-10.

Davy-Bowker J, Clarke RT, Johnson et al, 2006. A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. Hydrobiologia, 566, 91-105.

Donohue I, Jackson AL, Pusch MT and Irvine K, 2009. Nutrient enrichment homogenizes lake benthic assemblages at local and regional scales. Ecology, 90, 3470–3477.

Drakare S, 2014. Översyn av typology för sjöar och vattendrag (Review of the typology for lakes and rivers). Sveriges lantbruksuniversitet, Institutionen för vatten och miljö, Rapport 2014:2, 40 pp (in Swedish).

Duarte CM, 1991. Seagrass depth limits. Aquatic Botany 40: 363–377.

Duarte CM, Conley DJ, Carstensen J and Sánchez-Camacho M, 2009. Return to Neverland: shifting baselines affect eutrophication restoration targets. Estuar. Coasts 32, 29–36. http://dx.doi.org/10.1007/s12237-008-9111-2

Dudgeon D, Arthington AH, Gessner MO et al, 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological reviews, 81:163-182.

Dudley B, Dunbar M, Penning E et al, 2012. Measurements of uncertainty in macrophyte metrics used to assess European lake water quality. Hydrobiologia:1-13.

EC 2000. Directive 200/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities L 327/1.

EC 2003. WFD CIS Guidance Document No. 7. Guidance on the Intercalibration Monitoring under the Water Framework Directive. Produced by Working Group 2.7 – Monitoring. Available from http://circa.europa.eu/

EC 2010. WFD CIS Guidance Document No14. Guidance on the Intercalibration Process 2008-2011.

EC 2011. Guidance document on the intercalibration process 2008-2011. Guidance document No. 14. Common Implementation strategy for the Water Framework Directive (2000/60/EC), doi: 10.2779/99432, 102 p.

EC 2013. Commission decision of 20 September 2013 establishing, pursuant to Directive 2000/60/EC of the European Parliament and the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise and repealing Decision 2008/915/EC. Official Journal of the European Communities, L266:1-47.

Ecke F, 2007. Bedömningsgrunder för makrofyter i sjöar - Bakgrundsrapport (Water Quality Element of macrophytes in lakes - background report, in Swedish). Research Report 2007:17, Luleå University of Technology, Department of Chemical Engineering and Geosciences, Research Report 2007:17.

Ecke F, 2009. Drainage ditching at the catchment scale affects water quality and macrophyte occurrence in Swedish lakes. Freshwater Biology 54:119-126.

Edwards RW and Brown MW, 1960. An Aerial Photographic Method for Studying the Distribution of Aquatic Macrophytes in Shallow Waters. Journal of Ecology 48:161-163.

Elith J, Graham CH, Anderson RP et al, 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29, 129–151.

Ellis N, Smith SJ and Pitcher CR, 2012. Gradient forests: calculating importance gradients on physical predictors. Ecology, 93, 156-168.

Emmrich M, Pedron S, Brucet S et al, 2014. Geographical patterns in the body-size structure of European lake fish communities along abiotic and biotic gradients. Journal of Biogeography, 41: 2221-2233.

Erba S, Furse MT, Balestrini R et al, 2009. The validation of common European class boundaries for river benthic macroinvertebrates to facilitate the intercalibration process of the Water Framework Directive. Hydrobiologia, 633, 17-31.

Eriksson BK, Ljunggren L, Sandström A et al, 2009. Declines in predatory fish promote bloom-forming macroalgae. Ecological Applications 19: 1975-1988.

Eriksson BK, Sieben K, Eklöf J et al, 2011. Effects of altered offshore food webs on coastal ecosystems emphasize the need for cross-ecosystem management. AMBIO 40: 786-797.

Fielding AH and Bell JF, 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation, 24, 38-49.

Fjellheim A and Raddum GG, 1990. Acid precipitation: biological monitoring of streams and lakes. Science of the total environment, 96: 57-66.

Folke C, Carpenter S, Walker B et al, 2004. Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review of Ecology Evolution and Systematics. 35: 557-581 DOI: 10.1146/annurev.ecolsys.35.021103.105711

Fredriksson R, 2014. Kartläggning och sammanställning av indikatorbaserad statusklassning av kustfisk-samhällen i Östersjön samt analys av representativitet av provfisken i förhållande till kustvattentyper, kustvattenförekomster och naturtyper. Report to SLU Aqua. Contact SLU.

Free G, Little R, Tierney D et al, 2006. A reference based typology and ecological assessment system for Irish lakes. Preliminary investigations. ERTDI Report 57, Irish Environmental Protection Agency, Wexford.

Friberg N, Sandin L, Furse M et al, 2006. Comparison of macroinvertebrate sampling methods in Europe. Hydrobiologia, 566: 365–378.

Fürhaupter K and Meyer T 2009. Handlungsanweisung zum Monitoring in den äußeren Küstengewässern der Ostsee nach den Vorgaben der EU-Wasserrahmenrichtlinie, Qualitätskomponente Makrophyten, BALCOSIS-Verfahren; report for LLUR-SH and LUNG-MV. [In German]

Fölster J, Andrén C, Bishop K et al, 2007. A novel environmental quality criterion for acidification in Swedish lakes – an application of studies on the relationship between biota and water chemistry. Water, Air, and Soil pollution. 7: 331–338. doi: 10.1007/s11267-006-9075-9.

Fölster J, Johnson RK, Futter M and Wilander A, 2014. The Swedish monitoring of surface waters: 50 year of adaptive monitoring. Ambio, 43, 3–18.

Gerstmeier R, 1989. Lake typology and indicator organisms in application to the profundal chironomid fauna of Starnberger See (Diptera, Chironomidae). Archiv für Hydrobiologie, 116: 227– 234.

Giller P, Hillebrand H, Berninger UG et al, 2004. Biodiversity effects on ecosystem functioning: emerging issues and their experimental test in aquatic environments. Oikos, 104: 423-436.

Gorman OT and Karr JR 1978. Habitat structure and stream fish communities. Ecology, 59(3): 507-515.

Gowen RJ, Tett P, Bresnan E et al, 2012. Anthropogenic nutrient enrichment and blooms of harmful phyto-plankton. Oceanography and Marine Biology: An Annual Review: 50: 65-126.

Granéli E, Wallström K, Larsson U et al, 1990. Nutrient limitation of primary production in the Baltic Sea area. Ambio. 19 (3):142-151.

Grémare A, Labrune C, Vanden Berghe E et al, 2009. Comparison of the performances of two biotic indices based on the MacroBen database. Mar. Ecol. Prog. Ser. 382: 297–311.

Hallstan S, Johnson RK, Willén E and Grandin U, 2012. Comparison of classification-the-modelling and species-by-species modeling for predicting lake phytoplankton assemblages. Ecological Modelling, 231: 11-19.

Hallstan S and Johnson RK. Establishing reference conditions for lakes and streams: comparison of model- and typology-based approaches. Manuscript.

Hansen JP 2012. Benthic vegetation in shallow inlets of the Baltic Sea: Analysis of human influences and pro-posal of a method for assessment of ecological status. Plants & Ecology. Department of Botany, Stockholm University. Report 2012/2. 47 pages. URN: urn:nbn:se:su:diva-126535, OAI: oai:DiVA.org:su-126535, DiVA: diva2:900928. http://www.diva-portal.org

Harvey T, Kratzer S, Walve J et al. The effect of optical properties on Secchi depth and implications for eutrophication management. Manuscript.

Hawkins CP and Vinson MR 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. Journal of the North American Benthological Society, 19: 501-517.

Hawkins CP, Norris RH, Gerritsen J et al, 2000. Evaluation of landscape classifications for biological assessment of freshwater ecosystems: synthesis and recommendations. Journal of the North American Benthological Society, 19: 541–556.

Hawkins CP, Cao Y and Roper B, 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshwater Biology, 55: 1066-1085.

Heino J, Paavola R, Virtanen R and Muotka T, 2005. Searching for biodiversity indicators in running waters: do bryophytes, macroinvertebrates, and fish show congruent diversity patterns? Biodiversity & Conservation, 14: 415-428.

HELCOM 2006. Assessment of coastal fish in the Baltic Sea. Baltic Sea Environment Proceedings No. 103A, 1-26.

HELCOM 2007. The Baltic Sea Action Plan. HELCOM Ministerial Meeting Krakow, Poland, 15 November 2007.

HELCOM 2012. Indicator-based assessment of coastal fish community status in the Baltic Sea 2005-2009. Baltic Sea Environment Proceedings No. 131 Available at: www.helcom.fi/publications.

HELCOM 2013. HELCOM core indicators: Final report of the HELCOM CORESET project. Balt. Sea Environment Proceedings No. 136.

HELCOM 2015a. Guidelines for coastal fish monitoring. Sampling methods of HELCOM http://helcom.fi/Lists/Publications/Guidelines%20for%20Coastal%20fish%20Monitoring%20of%20HELCOM.pdf

HELCOM 2015b. Abundance of coastal fish key functional groups. http://helcom.fi/baltic-sea-trends/indicators/abundance-of-coastal-fish-key-functional-groups/

HELCOM 2015c. Abundance of key coastal fish species. http://helcom.fi/baltic-sea-trends/indicators/abundance-of-key-coastal-fish-species

Hellawell JM, 1986. Biological Indicators of Freshwater Pollution and Environmental Management. London: Elsevier. 546 pp.

Henrikson L and Medin M, 1986. Biologisk bedömning av försurningspåverkan på Lelångens tillflöden och grundområden 1986. Aquaekologerna, Rapport till länsstyrelsen i Älvsborgs län.

Hering D, Borja A, Carstensen J et al, 2010. The European Water Framework Directive at the age of 10: A critical review of the achievements with recommendations for the future. Science of the Total Environment, 408, 4007-4019.

Hering D, Moog O, Sandin L and Verdonschot PF, 2004. Overview and application of the AQEM assessment system. Hydrobiologia, 516: 1–20.

Hillebrand H, Dürselen C-D, Kirschtel D et al, 1999. Biovolume calculation for pelagic and benthic microalgae. Journal of Phycology 35: 403-424.

Holmgren K, 2013. Importance of fish age in assessment of fish fauna status. Aqua Reports 2013: 5, 66 pp (in Swedish with English summary)

Holmgren K, 2016a. Monitoring and ecological status assessment of fish assemblages in inland waters. WATERS Report no 2016:6, Havsmiljöinstitutet, Sweden.

Holmgren K 2016b. Progress report on work 2015 according to ”Tillägg Uppdrag gällande interkalibrering av ekologisk status avseende fisk i sjöar (HaV:s dnr 2570-14)”. Report 2016-03-02, 16 p.

Holmgren K, Kinnerbäck A, Pakkasmaa S et al, 2007. Bedömningsgrunder för fiskfaunans status i sjöar. Fiskeriverket Informerar, 2007:3, 54 p (In Swedish with English summary)

Howland WG, 1980. Multispectral aerial photography for wetland vegetation mapping. Photogrammetric Engineering and Remote Sensing 46:87-99.

Hurlbert SH, 1971. Nonconcept of species diversity: critique and alternative parameters. Ecology 52 (4): 577–586.

Husson E, Lindgren F and Ecke F, 2014b. Assessing Biomass and Metal Contents in Riparian Vegetation Along a Pollution Gradient Using an Unmanned Aircraft System. Water, Air, & Soil Pollution 225: 1-14.

Husson E, Hagner O and Ecke F, 2014a. Unmanned aircraft systems help to map aquatic vegetation. Applied Vegetation Science 17: 567-577.

Hynes HB, 1970. The ecology of running waters (Vol. 555). Liverpool: Liverpool University Press.

Höglander H, Karlson B, Johansen M et al, 2013. Overview of coastal phyto-plankton indicators and their potential use in Swedish waters. WATERS Report no. 2013:5, Havsmiljöinstitutet, Sweden. 86 pp.

Jarlman A, Kahlert M, Sundberg I and Herlitz E, 2016. Påväxt i rinnande vatten – kiselalgsanalys. In: Handledning för miljöövervakning. Metoder. Programområde: Sötvatten. Version 3:2. In: Miljöövervakningens metoder och undersökningstyper inom programområde Sötvatten. Havs- och vattenmyndighetens on-line document, 24p.

Johnson RK and Hering D, 2009. Response of taxonomic groups in streams to gradients in resource and habitat characteristics. Journal of Applied Ecology, 46: 175-186.

Johnson RK and Wiederholm T, 1989. Classification and ordination of profundal macroinvertebrate communities in nutrient poor, oligo-mesohumic lakes in relation to environmental data. Freshwater Biology, 21: 375-386.

Johnson RK and Angeler DG, 2010. Tracing recovery under changing climate: response of phytoplankton and invertebrate assemblages to decreased acidification. Journal of the North American Benthological Society, 29: 1472-1490.

Johnson RK, Lindegarth M and Carstensen J, 2014. Reference conditions and class boundaries: initial set of guidelines for reference conditions and class boundaries. WATERS Report no. 2014:5. Havsmiljöinstitutet, Sweden.

Johnson RK and Goedkoop W, 2007. Bedömningsgrunder för bottenfauna i sjöar och vattendrag – Användarmanual och bakgrundsdokument. Rapport 2007:4. Institutionen för miljöanalys, SLU.

Johnson RK, Hallstan S and Xuemin Z. Lake littoral invertebrate assemblages respond to multiple pressures. Submitted manuscript.

Johnson RK and Hering D, 2010. Spatial congruency of benthic diatom, invertebrate, macrophyte, and fish assemblages in European streams. Ecological Applications, 20: 978-992

Johnson RK, McKie BG, Angeler DG et al. Decomposing multiple stressor effects on invertebrate assemblages of boreal streams. Submitted manuscript.

Johnson RK, 1998. Spatio-temporal variability of temperate lake macroinvertebrate communities: detection of impact. Ecological Applications 8: 61-70.

Johnson RK, Angeler DG, Moe J and Hering D, 2014. Cross-taxon responses to elevated nutrients in European streams and lakes. Aquatic Sciences, 76: 51-60.

Johnson RK, Bennion H, Hering D et al, 2010. Climate change: Defining Reference Conditions and Restoring Freshwater Ecosystems, In: "Climate Change Impacts on Freshwater Ecosystems" (eds M Kernan, B Moss and RW Battarbee). Blackwell & Wiley, 203-235.

Johnson RK, Goedkoop W and Sandin L, 2004. Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. Freshwater Biology, 49: 1179–1194.

Johnson RK, Goedkoop W, Fölster J and Wilander A, 2007. Relationships between macroinvertebrate assemblages of stony littoral habitats and water chemistry variables indicative of acid-stress. Water Air and Soil Pollution, 7: 323–330.

Johnson RK, Hering D, Furse MT and Clarke RT, 2006a. Detection of ecological change using multiple organism groups: metrics and uncertainty. Hydrobiologia, 566: 115-137.

Johnson RK, Hering D, Furse MT and Verdonschot PF, 2006c. Indicators of ecological change: comparison of the early response of four organism groups to stress gradients. Hydrobiologia, 566: 139-152.

Johnson RK, Hering D, Furse MT and Clarke RT, 2006. Detection of ecological change using multiple organism groups: metrics and uncertainty. Hydrobiologia, 566: 115-137.

Johnson RK, Lindegarth M and Carstensen J, 2013. Establishing reference conditions and setting class boundaries. WATERS Report no. 2013:2, 66 p.

Johnson RK, Wiederholm T and Rosenberg DM, 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. Freshwater Biomonitoring and Benthic Invertebrates. Chapman and Hall, pp. 40–158.

Jonsson P and Persson J, 1996. Bottenförändringar i S:t Anna skärgård. Institutionen för geovetenskap, Uppsala universitet

Juanes JA, Guinda X, Puente A and Revilla JA, 2008. Macroalgae, a suitable indicator of the ecological status of coastal rocky communities in the NE Atlantic. Ecological Indicators 8: 351-359

Jyväsjärvi J, Aroviita J and Hämäläinen H, 2014. An extended Benthic Quality Index for assessment of lake profundal macroinvertebrates: addition of indicator taxa by multivariate ordination and weighted averaging. Freshwater Science, 33: 995-1007.

Jyväsjärvi J, Aroviita J and Hämäläinen H, 2012. Performance of profundal macroinvertebrate assessment in boreal lakes depends on lake depth. Fundamental and Applied Limnology, 180: 91–100.

Jyväsjärvi J, Nyblom J and Hämäläinen H, 2010. Palaeolimnological validation of estimated reference values for a lake profundal macroinvertebrate metric (Benthic Quality Index). J. Paleolimnol, 44: 253–264.

Kahlert M, 2011. Jämförande test av kiselalgers och bottenfaunas lämplighet som indikatorer för närsaltshalt och surhet inom miljömålsuppföljningen. Rapport Länsstyrelsen Blekinge 2011:7.

Kahlert M and Jarlman A, 2009. Påväxt i rinnande vatten – kiselalgsanalys. In: Handledning för miljöövervakning. Metoder. Programområde: Sötvatten. Version 3:1. In: Miljöövervakningens metoder och undersökningstyper inom programområde Sötvatten. Havs- och vattenmyndighetens on-line document, 19p.

Kahlert M and McKie B, 2014. Comparing new and conventional methods to estimate benthic algal biomass and composition in freshwaters (2014). Environ. Sci.: Processes Impacts 16(11): 2627-34.

Kahlert M and Trigal C. Diatom community turnover along environmental gradients in streams. Manuscript.

Kahlert M, Andrén C and Jarlman A, 2008. Kiselalger i vattendrag. In: Naturvårdsverkets författningssamling. Naturvårdsverkets föreskrifter och allmänna råd om klassificering och miljökvalitetsnormer avseende ytvatten. 2008:1, 22-24 ISSN 1403-8234.

Kahlert M, 2012. Utveckling av en miljögiftsindikator – kiselalger i rinnande vatten. Länsstyrelsen Blekinge län, Karlskrona, Report 2012:12, 40 pp.

Kahlert M and Gottschalk S, 2014. Differences in benthic diatom assemblages between streams and lakes in Sweden and implications for ecological assessment. Freshwater Science 33(2): 655-669.

Kahlert M, Andrén C and Jarlman A, 2007. Bakgrundsrapport för revideringen 2007 av bedömningsgrunder för Påväxt – kiselalger i vattendrag, Rapport / Sveriges lantbruksuniversitet, Institutionen för vatten och miljö, 2007:23.

Kansanen PH, Paasivirta L and Väyrynen T, 1990. Ordination analysis and bioindices based on zoobenthos communities used to assess pollution of a lake in southern Finland. Hydrobiologia, 202: 153-170.

Karlson B, 2014 Monitoring methods of phytoplankton in the Baltic Sea and Kattegat-Skagerrak, Swedish Agency for Marine and Water Management, Göterborg, report 2014:26, ISBN 978-91-87025-89-1, 28 pp.

Karlsson M, Pihl L and Bergström L. Assemblage structure and functional traits of littoral fish in Swedish coastal waters. Manuscript.

Karlsson M, 2015. Undersökningstyp: Provfiske i Östersjöns kustområden - Djupstratifierat provfiske med Nordiska kustöversiksnät. Havs- och vattenmyndigheten. Programområden Kust och Hav Version 1:3 2015-07-07.

Karlsson M, Ragnarsson Stabo H et al, 2014. Nationell plan för kunskapsförsörjning om fritidsfiske inom fisk havs och vattenförvaltningen (A national data collection framework for recreational fisheries). Aqua reports 2014.

Kautsky H, 1992. Methods for monitoring of phytobenthic plant and animal communities in the Baltic Sea. In: Plinski, M ed), The ecology of Baltic terrestrial, coastal and offshore areas: protection and management. Sopot, pp. 21–59.

Kautsky L, Wibjörn C and Kautsky H, 2006. Förslag till och vidareutveckling av bedömningsgrunder för kust och hav enligt krav i ramdirektivet vatten – Makroalger och några gömfröiga vattenväxter [In Swedish].

Kelly MG and Whitton BA, 1995. The Trophic Diatom Index: a new index for monitoring eutrophication in rivers. J. Appl. Phycol. 7: 433-444.

Kelly M, Bennett C, Coste M et al, 2009. A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia 621(1): 169 – 182.

Kelly M, Urbanic G, Acs E et al, 2014. Comparing aspirations: intercalibration of ecological status concepts across European lakes for littoral diatoms. Hydrobiologia. 734(1):125-141.

Kelly MG, 1998. Use of the trophic diatom index to monitor eutrophication in rivers, Water Research 32: 236-242.

Kelly M, 2013. Data rich, information poor? Phytobenthos assessment and the Water Framework Directive, European Journal of Phycology, 48(4): 437-450.

Kelly MG, Birk S, Willby NJ et al, 2016. Redundancy in the ecological assessment of lakes: Are phytoplankton, macrophytes and phytobenthos all necessary?. Science of The Total Environment. doi.org/10.1016/j.scitotenv.2016.02.024

Kolada A, Willby N, Dudley B et al, 2014. The applicability of macrophyte compositional metrics for assessing eutrophication in European lakes. Ecological Indicators 45: 407-415.

Korpinen S, Prins T, Green N et al, 2015. Review of thematic multi-metric indicator-based assessment tools. EEA/NSV/13/002 – ETC/ICM, 51 pp.

Kowalik RA, Cooper DM, Evans CD and Ormerod SJ, 2007. Acidic episodes retard the biological recovery of upland British streams from chronic acidification. Global Change Biology, 13: 2439-2452.

Kratz TK, Webster KE, Bowser C et al, 1997. The influence of landscape position on lakes in northern Wisconsin. Freshwater Biology 37: 209–217

Krause-Jensen D, Middelboe AL, Carstensen J and Dahl K, 2007. Spatial patterns of macroalgal abundance in relation to eutrophication. Marine Biology 152: 25-36

Krause-Jensen D, Sagert S, Schubert H and Boström C, 2008. Empirical relationships linking distribution and abundance of marine vegetation to eutrophication. Ecological Indicators, 8: 515–529.

Kremp A, Lindholm T, Dressler N et al, 2009. Bloom forming Alexandrium ostenfeldii (Dinophyceae) in shallow waters of the Åland Archipelago, Northern Baltic Sea. Harmful Algae 8: 318-328.

Köhler HR and Triebskorn R, 2013. Wildlife ecotoxicology of pesticides: can we track effects to the population level and beyond? Science, 341(6147): 759-765.

Labrune C, Amouroux J, Sarda R et al, 2006. Characterization of the ecological quality of the coastal Gulf of Lions (NW Mediterranean). A comparative approach based on three biotic indices. Mar. Pollut. Bull. 52: 34–47.

Lagrue C, Kominoski JS, Danger M et al, 2011. Experimental shading alters leaf litter breakdown in streams of contrasting riparian canopy cover. Freshwater Biology, 56: 2059-2069.

Lange K, Townsend CR, Gabrielsson R et al, 2014. Responses of stream fish populations to farming intensity and water abstraction in an agricultural catchment. Freshwater Biology, 59: 86-299.

Larsson U, Hajdu S, Walve J et al, 2006. Bedömningsgrunder för kust och hav – Växtplankton, näringsämnen, klorofyll och siktdjup. Stockholms universitet.

Lauringson V, Kotta J, Kersen P et al, 2012. Use case of biomass-based benthic invertebrate index for brackish waters in connection to climate and eutrophication. Ecol. Ind. 12:123-132. Doi:10.1016/j.ecolind.2011.04.009.

Leka J and Kanninen A 2003. Field surveys of aquatic macrophytes as a tool for monitoring and assessing the ecological status of the boreal lakes. Pages 127-130 (in M Ruoppa, P Heinonen, A Pilke, S Rekolainen, H Toivonen and H. Vuorista eds) How to assess and monitor ecological quality in freshwaters. TemaNord.

Leonardsson K, Blomqvist M, Magnusson M et al, 2015. Calculation of species sensitivity values and their precision in marine benthic faunal quality indices. Mar. Pollut. Bull. 93: 94-102.

Leonardsson K, Blomqvist M and Rosenberg R, 2009. Theoretical and practical aspects on benthic quality assessment according to the EU-Water Framework Directive – examples from Swedish waters. Mar. Pollut. Bull. 59: 1286-1296.

Leonardsson K, Blomqvist M and Rosenberg R, 2016. Reducing spatial variation in environmental assessment of marine benthic fauna. Mar. Pollut. Bull. 104: 129-138.

Leonardsson K, Ericson Y, Olsson J and Bergström L, 2016. Optimerad övervakning av fisk i kustvatten. On commission by the Swedish agency for marine and water management.

Lindegarth M, Bergström P, Carstensen J and Johnson RK, 2016. Assessing status of biological quality elements in water bodies without adequate monitoring: current approaches and recommended strategies for “expert assessment”. WATERS Report no. 2016:7. Havsmiljöinstitutet, Sweden.

Lindegarth M, Blomqvist M, Magnusson M and Rosenberg R, 2014. Monitoring of benthic fauna for the MSFD on the Swedish west coast: Modelling precision and uncertainty of current and future programs using WATERS uncertainty framework, WATERS Report no. 2014:3. Havsmiljöinstitutet, Sweden.

Lindegarth M, Carstensen J and Johnson RK, 2013a 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, Göteborg, Sweden.

Lindegarth M, Carstensen J and Johnson RK, 2013b. Monitoring biological indicators for the WFD in Swedish water bodies: Current designs and practical solutions for quantifying overall uncertainty and its components. WATERS Report no. 2013:6. Havsmiljöinstitutet, Göteborg, Sweden.

Lohammar G, 1938. Wasserchemie und höhere Vegetation Schwedischer Seen. Doctoral thesis. Uppsala universitet, Uppsala.

Lohammar G, 1949. Über die Veränderungen der Naturverhältnisse gesenkter Seen. Verh. Intern. Ver. Limnol 10: 266-274.

Lohammar G, 1965. The vegetation of Swedish lakes. Acta Phytogeographica Suecica, 50, Svenska Växtgeografiska Sällskapet, Uppsala.

Lyche Solheim A, Phillips G, Drakare S et al, 2014. Water Framework Directive Intercalibration Technical Report – Northern Lake Phytoplankton ecological assessment methods. JRC Technical Reports. European Commission Report, Ed. Poikane S. EUR 26503 EN. 254 pp. doi: 10.2788/70684.

Magnusson M, Blomqvist M and Rosenberg R, 2016. Response of the benthic quality index (BQI) to eutrophication induced impact in two Swedish coastal areas. WATERS Report no. 2016:9. Havsmiljöinstitutet, Sweden.

Maguire KC, Nieto-Lugilde D, Blois JL et al, 2016. Controlled comparison of species- and community-level models across novel climates and communities. Proc. R. Soc. B 283: 20152817. http://dx.doi.org/10.1098/rspb.2015.2817

Marbà N, Krause-Jensen D, Alcoverro T et al, 2013. Diversity of European seagrass indicators: patterns within and across regions. Hydrobiologia 704: 265–278.

Marmion M, Parviainen M, Luoto M et al, 2009. Evaluation of consensus methods in predictive species distribution modelling. Diversity and distributions, 15(1): 59-69.

Matthaei CD, Piggott JJ and Townsend CR, 2010. Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment and water abstraction. Journal of Applied Ecology, 47: 639-649.

McFarland B, Carse F and Sandin L, 2010. Littoral macroinvertebrates as indicators of lake acidification within the UK. Aquatic Conservation, 20: 105–116.

McKie BG, Truchy A, Johnson RK et al. WATERS inland gradient study: species by site dataset of biological response variables in selected lakes and streams. WATERS Report no. 2016:5. Havsmiljöinstitutet, Sweden.

Mehner T, Keeling C, Emmrich M et al, 2016. Effects of fish predation on density and size spectra of prey fish communities in lakes. Canadian Journal of Fisheries and Aquatic Sciences, 73: 1-13, doi: 10.1139/cjfas-2015-0034.

Menden-Deuer S and Lessard EJ, 2000. Carbon to volume relationships for dinoflagellates, diatoms and other protist plankton. Limnology and Oceanography, 45: 569-579.

Metcalfe JL, 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. Environmental pollution, 60: 101-139.

Miler O, Ostendorp W, Brauns M et al, 2015. Ecological assessment of morphological shore degradation at whole lake level aided by aerial photo analysis. Fundamental and Applied Limnology, 4: 353–369.

Mjelde M, 2013. Vannplanter. I: Direktoratsgruppa vanndirektivet 2013. Klassifisering av miljøtilstand i vann. Økologisk og kjemisk klassifiseringssystem for kystvann, grunnvann, innsjøer og elver. Veileder 02:2013.

Mjelde M, Hellsten S and Ecke F, 2013. A water level drawdown index for aquatic macrophytes in Nordic lakes. Hydrobiologia 704: 141-151.

Moksnes PO, Albertsson J, Elfwing T et al, 2013. Sammanvägd bedömning av miljötillståndet i havet. Havsmiljöinstitutet (Report), 45 pp.

Moldan F, Kronnäs V, Wilander A et al, 2004. Modelling acidification and recovery of Swedish lakes. Water, Air, and Soil Pollution, Focus. 4: 139–160. doi: 10.1023/B:WAFO.0000028351.14511.66.

Monteith DT, Hildrew AG, Flower RJ et al, 2005. Biological responses to the chemical recovery of acidified fresh waters in the UK. Environmental Pollution, 137: 83-101.

Moss D, Furse MT, Wright JF and Armitage PD, 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology 17: 41–52.

Naiman RJ, 2010. River Ecosystem Ecology: A Global Perspective, 182.

Nielsen R, Kristiansen A, Mathiesen L and Mathiesen H, 1995. Distributional index of the benthic macroalgae of the Baltic Sea area. Acta Botanica Fennica, 155:1–51

Noges P, van de Bund W, Cardoso AC et al, 2009 Assessment of the ecological status of European surface waters: a work in progress. Hydrobiologia, 633: 197–211.

Norwegian Environment Agency. 2013. Klassificering av miljotilstand i vann. Veileder 02:2013. http://www.vannportalen.no/globalassets/nasjonalt/dokumenter/veiledere-direktoratsgruppa/klassifisering-av-miljotilstand-i-vann-_02-2013_revidert_2015.pdf

Nyström Sandman A, Qvarfordt S, Wikström SA and Blomqvist M, 2016. Evaluation of methodological aspects on data collection for assessing ecological status of vegetation according to the WFD. WATERS Report no. 2016:4. Havsmiljöinstitutet, Sweden.

Olin M, Holmgren K, Rask M et al, 2014. Water Framework Directive Technical Report: Northern Lake Fish fauna ecological assessment. EUR – Scientific and Technical Research series, Report EUR 26515 EN, 55 p. Doi: 10.2788/76197.

Olin M, Rask M, Ruuhijärvi J and Tammi J 2013. Development and evaluation of the Finnish-based lake classification method. Hydrobiologia, 713: 149-166.

Orfanidis S, Panayotidis P and Stamatis N 2003. An insight to the ecological evaluation index (EEI). Ecological Indicators 3: 27–33.

Ormerod SJ and Durance I, 2009. Restoration and recovery from acidification in upland Welsh streams over 25 years. Journal of Applied Ecology, 46: 164-174.

OSPAR 2005. Ecological Quality Objectives for the Greater North Sea with regard to nutrients and eutrophication effects. OSPAR Background Document on Eutrophication, OSPAR Commission. www.ospar.org

Pall K and Moser V, 2009. Austrian Index Macrophytes (AIM-Module 1) for lakes: a Water Framework Directive compliant assessment system for lakes using aquatic macrophytes. Hydrobiologia 633: 83-104.

Paloniitty T, 2016. The Weser Case: Case C-461/13 BUND V GERMANY J Environmental Law 28(1): 151-158 doi:10.1093/jel/eqv032

Pardo I, Gómez-Rodríguez C, Wasson J-G et al, 2012. The European reference condition concept: A scientific and technical approach to identify minimally-impacted river ecosystems. Science of the Total Environment, 420, 33-42

Parker JD, Burkepile DE, Collins DO et al, 2007. Stream mosses as chemically-defended refugia for freshwater macroinvertebrates. Oikos 116: 302-312.

Pearson TH and Rosenberg R, 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanogr. Mar. Biol. 16: 229–311.

Pearson TH and Rosenberg R, 1987. Feast and Famine: Structuring Factors in Marine Benthic Communities. In: JH Gee, Giller PS eds), 27th Symposium of the British Ecological Society Aberystwyth 1986. Blackwell Scientific Publications, Oxford, pp. 373–395.

Penning W, Dudley B, Mjelde M et al, 2008a. Using aquatic macrophyte community indices to define the ecological status of European lakes. Aquatic Ecology 42: 253-264.

Penning W, Mjelde M, Dudley B et al, 2008b. Classifying aquatic macrophytes as indicators of eutrophication in European lakes. Aquatic Ecology 42: 237-251.

Persson M, Karlson B, Hellmér M et al, 2014. Årsrapport 2011-2013. Kontrollprogrammet för tvåskaliga blötdjur, Livsmedelsverket, rapport 11 2014, 37 pp. ISSN 1104-7089

Phillips G, Lyche-Solheim A, Skjelbred B et al, 2012. A phytoplankton trophic index to assess the status of lakes for the Water Framework Directive. Hydrobiologia 704(1): 75-95

Piggott JJ, Lange K, Townsend CR and Matthaei CD, 2012. Multiple stressors in agricultural streams: a mesocosm study of interactions among raised water temperature, sediment addition and nutrient enrichment. PLoS One 7, e49873. http://dx.doi.org/

Pihl L, Karlsson M and Bergström L. Fish community indicators along eutrophication gradients in coastal waters, Sweden. Manuscript.

Pocock MJ and Jennings N, 2008. Testing biotic indicator taxa: the sensitivity of insectivorous mammals and their prey to the intensification of lowland agriculture. Journal of Applied Ecology, 45, 151-160.

Poikane S, Birk S, Carvalho L et al, 2015. A hitchhiker's guide to European lake ecological assessment and intercalibration. Ecological Indicators, 52: 533–544.

Poikane S (ed), 2009. Water Framework Directive intercalibration report. Part 2: Lakes. European Commission. Joint Research Centre.

Poikane S, Johnson RK, Aroviita J et al, 2016. Benthic macroinvertebrates in lake ecological assessment: A review of methods, intercalibration and practical recommendations. Science of the Total Environment, 543: 123-134.

Pont D, 2012. WFD Intercalibration Technical Report: Cross-GIG Rivers, fish fauna. Final Draft, March 2012, 104 p.

Potapova M and Charles D, 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators 7: 48–70.

Raddum GG, Fjellheim A and Hesthagen T, 1988. Monitoring of acidification by the use of aquatic organisms. Internationale Vereinigung fur Theoretische und Angewandte Limnologie, Verhandlungen IVTLAP, 23(4).

Rimet F, 2012. Recent views on river pollution and diatoms. Hydrobiologia, 683: 1–24.

Rolff C, 2009. Hur har de marina bedömningsgrunderna utfallit i statusbedömningen?- En beskrivande nationell överblick. SWECO Environment 2009-02-12.

Rosenberg DM and Resh VH (eds), 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. New York: Chapman and Hall. 488 pp.

Rosenberg R, Blomqvist M, Nilsson CH et al, 2004. Marine quality assessment by use of benthic species-abundance distributions; a proposed new protocol within the European Union Water Framework Directive. Mar. Pollut. Bull. 49: 728-739.

Rørslett B, 1991. Principal determinants of aquatic macrophyte richness in northern European lakes Aquatic Botany 39: 173-194.

Sairanen S, Rask M, Stridsman S and Holmgren K, 2008. Part II. A. Fish communities. Fish communities of 15 lakes in River Torne basin: aspects of lake typology and ecological status. Pp. 65-88 in E Luokkanen, P Olofsson, H Ville & B Sundström (eds). TRIWA II Management of an international river basin district – Torne River. The Finnish Environment 10/2008.

Sandin L and Johnson RK, 2004. Local, landscape and regional factors structuring benthic macroinvertebrate assemblages in Swedish streams. Landscape Ecology, 19: 501-514.

Sandström A, Philipsson P, Asp A et al, 2016. Assessing the potential of remote sensing derived water quality data to explain variations in fish assemblages and to support status assessments in large lakes. Hydrobiologia (in press)

Sandström A, Ragnarsson-Stabo H, Axenrot T and Bergstrand E, 2014. Has climate variability driven the trends and dynamics in recruitment of pelagic fish species in Swedish Lakes Vänern and Vättern in recent decades? Aquatic Ecosystem Health and Management, 17: 349-356.

Schartau AK, Lagergren R and Hesthagen T, 2012. INTERREG prosjektet Enningdalselven. Uttesting av overvåkingsmetodikk og systemer for klassifisering av økologisk tilstand (Bedömningsgrunder) (jf. Vanndirektivet. NINA Rapport 875, 72 p In Norwegian with Swedish and English summaries)

Schartau AK, Moe SJ, Sandin L et al, 2008. Macroinvertebrate indicators of lake acidification: testing on data from UK, Norway and Sweden. Aquatic Ecology, 42: 293–305.

Schaumburg J, Schranz C, Hofmann G et al, 2004b. Macrophytes and phytobenthos as indicators of ecological status in German lakes – a contribution to the implementation of the Water Framework Directive. Limnologica 34:302–314.

Schaumburg J, Schranz C, Förster J et al, 2004a. Ecological classification of macrophytes and phytobenthos for rivers in Germany according to the Water Framework Directive. Limnologica 34: 283-301.

Schindler DW and Vallentyne JR, 2008. The Algal Bowl: Overfertilization of the World's Freshwaters and Estuaries. University of Alberta Press.

Schmedtje U, Birk S, Poikane S et al, 2009. Guidance on the Intercalibration Process 2008.2011 (version 5.0). 62 p.

Schneider S and Melzer A, 2003. The trophic index of macrophytes (TIM) - a new tool for indicating the trophic state of running waters. International Review of Hydrobiology 88: 49-67.

SEPA 2003. Handbok för miljöövervakning. Undersökningstyp: Makrofyter i vattendrag. Version 1:2 2003-12-04.

SEPA 2003. Undersökningstyp: biotopkartering vattendrag. Version 1: 2003-06-17

SEPA 2007. 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 p.

SEPA 2008. Naturvårdsverkets föreskrifter och allmänna råd om klassificering och miljökvalitetsnormer avseende ytvatten. NFS 2008:1

SEPA 2010. En sammanställning av enkätundersökningen angående bedömningsgrunder för sjöar, vattendrag, kustvatten och hydromorfologi. PM 2010-03-17 Dnr 739-8201-08Me. 210p.

SEPA 2015. Handledning för miljöövervakning – Undersökningstyp: Makrofyter i sjöar. Version 3:0, 2015-06-26.

Solimini AG, Free G, Donohue I et al, 2006. Using benthic macroinvertebrates to assess ecological status of lakes current knowledge and way forward to support WFD Implementation. Office for Official Publications of the European Communities, Luxembourg.

Spears BM, Gunn ID, Carvalho L et al 2009. An evaluation of methods for sampling macrophyte maximum colonisation depth in Loch Leven, Scotland. Aquatic Botany 91: 75-81.

Spjut D and Degerman E 2015a. Fisk som indikator på hydromorfologisk påverkan. Fem delstudier med olika påverkansbedömning. SLU, Institutionen för akvatiska resurser, Dnr: SLU.aqua.2015.5.5-122, 2015-03-31, 59 p.

Spjut D and Degerman E 2015b. Bedömning av morfologisk påverkan i vattendrag med elfiskedata. Aqua Reports, 2015:17, 43 p (In Swedish with English summary)

Staniszewski R, Szoszkiewicz K, Zbierska J et al, 2006. Assessment of sources of uncertainty in macrophyte surveys and the consequences for river classification. Hydrobiologia 566: 235-246.

Stendera S and Johnson RK, 2008. Tracking recovery trends of boreal lakes: use of multiple indicators and habitats. Journal of North American Benthological Society, 27: 529-540.

Stevenson RJ, Pan Y and van Dam H, 2007. Assessing environmental conditions in rivers and streams with diatoms. In: The Diatoms. Applications for the Environmental and Earth Sciences. Cambridge University Press, pp. 57-85.

Stoddard JL, Larsen P, Hawkins CP et al, 2006. Setting expectations for the ecological condition of running waters: the concept of reference condition. Ecological Applications, 16: 1267-1276.

Stoddard JL, Jeffries DS, Lükewille A et al, 1999. Regional trends in aquatic recovery from acidification in North America and Europe. Nature, 401: 575-578.

Strayer DL and Dudgeon D, 2010. Freshwater biodiversity conservation: recent progress and future challenges. Journal of the North American Benthological Society 29: 344-358.

Sundblad G and Bergström U, 2014. Shoreline development and degradation of coastal fish reproduction habitats. AMBIO 43: 1020-1028

Sundblad G, Bergström U, Sandström A and Eklöv P, 2014. Nursery habitat availability limits adult stock sizes of predatory coastal fish. Ices Journal of Marine Science 71: 672-680

Svenonius H, 1925. Luleåtraktens flora. Svensk Botanisk Tidskrift 19: 431-484.

SwAM 2012a. God havsmiljö 2020. Marin strategi för Nordsjön och Östersjön. Del 1: Inledande bedömning av miljötillstånd och socioekonomisk analys (Initial assessment for the marine strategy framework directive). Swedish Agency for Marine and Water Management Report 2012:19

SwAM 2012b. God havsmiljö 2020. Marin strategi för Nordsjön och Östersjön. Del 2: God miljöstatus och miljökvalitetsnormer, Havs- och vattenmyndighetens rapport, Gothenburg, p. 159. Swedish Agency for Marine and Water Management Report 2012:20

SwAM 2012c. Havs- och vattenmyndighetens föreskrifter om vad som kännetecknar god miljöstatus samt miljökvalitetsnormer med indikatorer för Nordsjön och Östersjön, HVMFS 2012: 18

SwAM 2013. Havs- och vattenmyndighetens föreskrifter om klassificering och miljökvalitetsnormer avseende ytvatten. HVMFS 2013:19.

Söderberg K & Mattsson P 2011. Mätkampanj för kustfisk 2011 – design och dimensionering. Report from the Swedish Board of Fisheries, Institute of Coastal Research (Fiskeriverkets Kustlaboratorium).

Taylor JR, 1997. An introduction to error analysis: The study of uncertainties in physical measurements. Sausalito, CA: University Science Books.

Thienemann A, 1918. Untersuchungen über die Beziehungen zwischen dem Sauerstoffgehalt des Wassers und der Zusammensetzung der Fauna in norddeutschen Seen. Archiv für Hydrobiologie, 12: 1-65.

Theinemann A, 1922. Hydrobiologissch Untersuchungen an Quellen. Archiv Für Hydrobiologie, 14: 151-190.

Toivonen H and Huttunen P, 1995. Aquatic macrophytes and ecological gradients in 57 small lakes in southern Finland. Aquatic Botany 51: 197-221.

Tonkin JD and Death RG, 2012. Consistent effects of productivity and disturbance on diversity between landscapes. Ecosphere, 3: 1-19.

Tremp H and Kohler A, 1995. The usefulness of macrophyte monitoring-systems, exemplified on eutrophication and acidification of running waters. Acta Botanica Gallica, 142: 541-550.

Valta-Hulkkonen K, Pellikka P, Tanskanen H et al, 2003. Digital false colour aerial photographs for discrimination of aquatic macrophyte species. Aquatic Botany 75:71-88.

van de Bund W, 2009. Water Framework Directive intercalibration technical report, Part 1: Rivers. European Commission, Joint Research Centre, Institute for Environment and Sustainability, Ispra, Italy.140 p. DOI 10.2788/23384

Van Sickle J, Hawkins CP, Larsen DP and Herlihy AT, 2005. A null model for the expected macroinvertebrate assemblage in streams. Journal of the North American Benthological Society, 24, 178-191.

Wallin M, Wiederholm T and Johnson RK, 2003. Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters. Final version 7.0, 2003-03-05, 93 p.

Wallsten M, 1981. Changes of lakes in Uppland central Sweden during 40 years. Uppsala University, Uppsala.

Walve J, Höglander H, Andersson A et al, 2016. Exploration and evaluation of potential phytoplankton indicators for Swedish coastal areas. WATERS Report no. 2016:8. Havsmiljöinstitutet, Sweden. (Includes three appendices: A) Gulf of Bothnia, B) Northern Baltic Proper and C) Skagerrak and Kattegat).

Walve J, Johansen M, Karlson B et al, 2014. Phytoplankton in coastal waters – evaluation of the length of the seasonal assessment period. Deliverable 3.3-2, WATERS Report no. 2014: 1. Havsmiljöinstitutet, Sweden.

WFD-UKTAG (Water Framework Directive – United Kingdom Technical Advisory Group). 2014. UKTAG Lake Assessment Method Phytoplankton – Phytoplankton Lake Assessment Tool with Uncertainty Module (PLUTO). ISBN: 978-1-906934-47-7.

Wiederholm T, 1980. Use of benthos in lake monitoring. Journal Water Pollution Control Federation, 537-547.

Wikström SA, Blomqvist M, Qvarfordt S and Nyström Sandman A, 2016. Response of macrophyte indicators to natural and anthropogenic gradients in two coastal areas of Sweden. Deliverable 3.2-3, WATERS Report no. 2016:1. Havsmiljöinstitutet, Sweden.

Willén E, 2007. Växtplankton i sjöar – Bedömningsgrunder. Institutionen för Miljöanalys, SLU, Rapport 2007:6, 41pp.

Wright JF, Blackburn JH, Gunn RJ et al, 1996. Macroinvertebrate frequency data for the RIVPACS III sites in Great Britain and their use in conservation evaluation. Aquatic Conservation: Marine and Freshwater Ecosystems, 6: 141-167.

Young J, Watt A, Nowicki P et al, 2005. Towards sustainable land use: identifying and managing the conflicts between human activities and biodiversity conservation in Europe. Conservation Biology, 14: 1641-1661.

Zelinka M and Marvan P, 1961. Zur präzisierung der biologischen klassifikation der reinheit fließender gewässer. Archiv für Hydrobiologie, 57: 389-407.

Ådjers K, Appelberg M, Eschbaum R et al, 2006. Trends in coastal fish stocks of the Baltic Sea. Boreal Environment Research 11: 13-25.

Økland J and Økland KA 1986. The effects of acid deposition on benthic animals in lakes and streams. Experientia, 42: 471-486.

Östman Ö, Eklöf J, Eriksson BK et al, 2016. Top-down control as important as nutrient enrichment for eutrophication effects in North Atlantic coastal ecosystems. Journal of Applied Ecology 53: 1138-1147

Acknowledgements

For five years WATERS has done research with the aim to improve the Swedish assessment criteria for the WFD. The programme has been funded by the Swedish Environmental Protection Agency (SEPA; through contracts 10/179 and 13/33) and the Swedish Agency for Marine and Water Management (SwAM), with additional funding was also provided by SEPA and the County Administration Board of Västra Götaland in connection with gradient studies in coastal waters.

Many people besides those directly involved in the programme have contributed during the five years. The steering group with representatives from SEPA (Cecilia Lindblad and Sif Johansson) and SwAM (Mats Svensson and Lennart Sorby), and the reference group with additional representation from SwAM (Kristina Samuelsson, Jonas Svensson, Ann-Karin Thorén, Ulrika Stensdotter Blomberg), Water District Authorities (Robert Dobak, Johan Erlandsson, Fredrik Gunnarsson, Mikael Gyllström, Ann-Louise Haglund, Lennart Johansson, Ragnar Lagergren, Christer Larsson, Lisa Lundstedt, Ingemar Perä, Carina Pålsson, Juha Salonsaari, Susanna Vesterberg) Seppo Hellsten from SYKE (Finland) and Steinar Sandøy from Miljødirektoratet (Norway) have regularly provided us with valuable input as well as local, national and international perspectives on our work over the years.

The work within WATERS has also benefited immensely from scientific and practical input from a large number of fellow researchers and colleagues. In particular we would like to acknowledge the contributions from (in alphabetical order): Anders Asp, Thomas Axenrot, Thorsten Balsby, Ulrika Beier, Johanna Bergkvist, Björn Bergquist, Magnus Dahlberg, Erik Degerman, Lena Gipperth, Susanna Hajdu, Siv Huseby, Tina Johansen Lilja, Malin Karlsson, Anders Kinnerbäck, Ulf Larsson, Svante Nyberg, Leonard Sandin, Alfred Sandström, Christina Schmidt, David Spjut and Ulla Li Zweifel.

Finally, the many people involved in supplying data are gratefully acknowledged for making the analyses of these data sets possible.

WATERS partners

AARHUS UNIVERSITY

Stockholms universitet

AquaBiota WATER RESEARCH

Marine Monitoring RESEARCH & CONSULTING

SVERIGES VATTENEKOLOGER

UMEÅ UNIVERSITET

UNIVERSITY OF GOTHENBURG

Hafok AB

SLU

SMHI

TOXICON AB

List of WATERS deliverables and additional reports. Deliverables, summary reports and other outputs can be found on the USB card accompanying this report and at **www.waters.gu.se/english/Publications.**

	COORDINATION	
1.1-1	Programme agreements	Internal
1.1-2	Mid-term report	WATERS report 2013:3
1.1-3	Status-update report	WATERS report 2014:4
1.1-4	WATERS final report	WATERS report 2016:10
1.2-1	Internal work-platform	Internal
1.2-2	External web page	www.waters.gu.se
1.2-3	Leaflet and power point presentation for external presentations	www.waters.gu.se
	INTEGRATED ASSESSMENT AND HARMONISATION	
2.1-1	Database on published literature on reference conditions and class boundaries used in the review task	Internal
2.1-2	Existing literature on reference conditions and classes	WATERS report 2013:2
2.1-3	Initial set of guidelines for reference conditions and class boundaries	WATERS report 2014:5
2.1-4	Section in final report on the reference condition and class boundary setting framework	WATERS report 2016:10 (Ch. 2 and 5)
2.2-1	Uncertainty assessment in ecological data	WATERS report 2013:1
2.2-2	Guidelines for sampling designs to assess different uncertainty components	WATERS report 2013:6
2.2-3	Section in final report on WFD uncertainty assessment	WATERS report 2016:10 (Ch. 2 and 5)
2.3-1	Review of existing assessment systems	Andersen et al. 2016
2.3-2	Conceptual Understanding of WFD Assessment Principles	www.waters.gu.se
2.3-3	Assessing status of biological quality elements in water bodies without adequate monitoring: current approaches and recommended strategies for "expert assessment"	WATERS report 2016:7
2.3-4	Section in final report on whole system assessment	WATERS report 2016:10 (Ch. 2 and 5)
2.4-1	Summary report for statistical workshop year 1	WATERS report 2012:1
2.4-2	Summary report for statistical workshop year 2	WATERS report 2013:4
2.4-3	Summary report for statistical workshop year 3	WATERS report 2013:8
Additional report	Monitoring of benthic fauna for the MSFD on the Swedish west-coast: Modelling precision and uncertainty of current and future programs using WATERS uncertainty framework	WATERS report 2014:3
Additional report	Developing practical tools for assessing uncertainty of Swedish WFD indicators: a library of variance components, and its use for estimating uncertainty of current biological indicators.	WATERS report 2016:2
	COASTAL WATERS	
3.1-1	Calculation of species sensitivity values and their precision in marine benthic faunal quality indices	Leonardsson et al. 2015
3.1-2	Reducing spatial variation in environmental assessment of marine benthic fauna	Leonardsson et al. 2016
3.1-3	Response of the benthic quality index (BQI) to eutrophication induced impact in two Swedish coastal areas	WATERS report 2016:9
3.1-4	A probability based index for assessment of benthic invertebrates in the Baltic sea	WATERS report 2016:3
3.1-6	Summary on new methods for setting class-boundaries in assessment of marine benthic invertebrates	Summary report
3.1-7	Section in final report on the use of BQI in quality assessments	WATERS report 2016:10 (Ch. 4)

3.2-1	Potential eutrophication indicators based on Swedish coastal macrophytes	WATERS report 2012:2
3.2-2	Response of coastal macrophytes to pressures	WATERS report 2014:2
3.2-3	Response of macrophyte indicators to natural and anthropogenic gradients in two coastal areas of Sweden	WATERS report 2016:1
3.2-4	Evaluation of methodological aspects on data collection for assessing ecological status of vegetation according to the WFD	WATERS report 2016:4
3.2-5	Section in final report on quality assessment for macrophytes in Sweden	WATERS report 2016:10 (Ch. 4)
3.3-1	Overview of coastal phytoplankton indicators and their potential use in Swedish water	WATERS report 2013:5
3.3-2	Phytoplankton in coastal waters – evaluation of the length of the seasonal assessment period	WATERS report 2014:1
3.3-3	Exploration of potential phytoplankton indicators for Swedish coastal areas	WATERS report 2016:8
3.3-4	Section in final report on recommendations for using phytoplankton in environmental assessment	WATERS report 2016:10 (Ch. 4)
3.4-1	Assemblage structure and functional traits of littoral fish in Swedish coastal waters	Submitted manuscript
3.4-2	Comparison of gill nets and fyke nets for the status assessment of coastal fish communities	WATERS report 2013:7
3.4-3	Coastal fish community indicators in Sweden – variation along environmental gradients	WATERS report 2015:1
3.4-4	Fish community indicators along eutrophication gradients in coastal waters, Sweden	Submitted manuscript
3.4-5	Section in final report on quality assessment of littoral fish in Sweden	WATERS report 2016:10 (Ch. 4)
3.5-1	Insamlade biologiska data i WATERS kustgradientstudie	WATERS report 2015:2
	INLAND WATERS	
4.1-1	Unmanned aircraft systems help to map aquatic vegetation	Husson et al, 2014
4.1-2	Response of macrophyte assemblages and metrics to detect human induced change	Submitted manuscript
4.1-3	Comparison of manual mapping and automated object-based image analysis of non-submerged aquatic vegetation from very-high-resolution UAS images	Submitted manuscript
4.1-4	Section in final report on establishing reference conditions and class boundaries for using macrophyte assemblages in environmental assessment	WATERS report 2016:10 (Ch. 3)
4.2-1	Response of phytoplankton assemblages and metrics to detect human induced change	Submitted manuscript
4.2-2	Accounting for natural variation when assessing ecological status of lakes using phytoplankton: comparison of typology and model-based approaches	Submitted manuscript
4.2-3	Section in final report on establishing reference conditions and class boundaries in phytoplankton	WATERS report 2016:10 (Ch. 3)
4.3-1	The use of pelagic phytoplankton and littoral benthic diatom assemblages for detecting changes in resource levels and acidification	Submitted manuscript
4.3-2	Differences in benthic diatom assemblages between streams and lakes in Sweden and implications for ecological assessment	Kahlert and Gottschalk, 2014
4.3-3	Similar small-scale variation of diatom assemblages on different substrates in a mesotrophic stream	Kahlert & Savatijevic Rasic, 2015
4.3-4	Section in final report on establishing reference conditions and class boundaries of benthic diatoms	WATERS report 2016:10 (Ch. 3)
4.4-1	Lake littoral invertebrate assemblages respond to multiple pressures	Submitted manuscript
4.4-2	Establishing reference conditions for lake and stream invertebrates: comparison of model- and typology-based approaches	Submitted manuscript
4.4-3	Section in final report on establishing reference conditions and class boundaries in benthic invertebrates	WATERS report 2016:10 (Ch. 3)
4.5-1	Monitoring and ecological status assessment of fish assemblages in inland waters	WATERS report 2016:6
4.5-2	Summary report of uncertainty associated with using freshwater fish assemblages in environmental assessment	Summary report
4.5-3	Section in final report on establishing reference conditions and class boundaries of fish	WATERS report 2016:10 (Ch. 3)
4.6-1	WATERS inland gradient study: species by site dataset of biological repsonse variables in selected lakes and streams	WATERS report 2016:5

Havsmiljöinstitutet
Swedish Institute for the Marine Environment

Swedish Agency for Marine and Water Management

NATUR VÅRDS VERKET

SWEDISH ENVIRONMENTAL PROTECTION AGENCY