Confidence rating of marine eutrophication assessments

*C. Murray, J.H. Andersen, H. Kaartokallio, P. Axe, J. Mølvær,
K. Norling & M. Krüger-Johnsen*

Confidence rating of marine eutrophication assessments

TemaNord 2011:504

© Nordic Council of Ministers, Copenhagen 2011

ISBN 978-92-893-2192-1

This publication is available as Print on Demand (PoD) and can be ordered on www.norden.org/order. Other Nordic publications are available at www.norden.org/en/publications.

Nordic Council of Ministers

Ved Stranden 18
DK-1061 København K
Phone (+45) 3396 0200
Fax (+45) 3396 0202

Nordic Council

Ved Stranden 18
DK-1061 København K
Phone (+45) 3396 0400
Fax (+45) 3311 1870

www.norden.org

Nordic co-operation

Nordic co-operation is one of the world's most extensive forms of regional collaboration, involving Denmark, Finland, Iceland, Norway, Sweden, and three autonomous areas: the Faroe Islands, Greenland, and Åland.

Nordic co-operation has firm traditions in politics, the economy, and culture. It plays an important role in European and international collaboration, and aims at creating a strong Nordic community in a strong Europe.

Nordic co-operation seeks to safeguard Nordic and regional interests and principles in the global community. Common Nordic values help the region solidify its position as one of the world's most innovative and competitive.

Content

Preface	7
Summary	9
1. Introduction	11
2. Methods	13
2.1 Data and information sources	13
2.2 Primary assessment: Eutrophication status	14
2.3 Secondary assessment: Confidence rating	15
2.4 Scoring principles	16
3. Results and discussion	19
3.1 RefCon, AcDev and AcStat	19
3.2 Quality elements	20
3.3 Areas	21
3.4 Countries	22
4. Conclusions	25
5. Acknowledgements	27
6. References	29
Appendix 1: Finnish case study	31
Appendix 2: Norwegian case study	47
Appendix 3: Swedish case study	61

Preface

The Nordic countries have a long and scientifically sound tradition in relation to the assessment of marine eutrophication, e.g. in relation to national assessments, HELCOM Periodic Assessments and OSPAR Quality Status Reports, and in the application of the HELCOM Eutrophication Assessment Tool (HEAT) and the OSPAR Comprehensive Procedure (OSPAR COMP).

Methods, indicators and assessment tools are generally very well documented. However, the final assessment result indicating whether an area is an “eutrophication problem area” (affected by eutrophication) or a “non-problem area” (not affected by eutrophication) is not currently backed by an estimation of the quality or accuracy of the assessment.

The CONFIRM project has been initiated with the aim of developing a methodology for confidence rating marine eutrophication assessments. This report presents the background, the tools and methods used, the result of case studies where the tools have been tested in Finnish, Norwegian and Swedish marine waters, as well as conclusions and recommendations for future use and improvements of the prototype tools.

We hope that this report initiates a process leading to development of other approaches and tool useful for assessing the confidence of the classification to be carried out by Nordic countries in the future not only in regard of the regional marine conventions, but also those status classifications required by the Water Framework Directive and the Marine Strategy Framework Directive.

Summary

We report the development of a methodology for assessing confidence in ecological status classifications. The method presented here can be considered a secondary assessment, supporting the primary assessment of eutrophication or ecological status. The confidence assessment is based on scoring the quality of the indicators on which the primary assessment is made. This represents a first step towards linking status classification with information regarding their accuracy and precision. Applied to an existing data set used for assessment of the eutrophication status of the Baltic Sea (including the Kattegat and Danish Straits) we demonstrate that confidence in the assessment is good or high in 149 out of 189 areas assessed (79%). Contrary to our expectations, assessments of the open parts of the Baltic Sea have a higher confidence than assessments of coastal waters. We also find that in the open waters of the Baltic Sea, some biological indicators have a higher confidence than indicators representing physical-chemical conditions. In coastal waters, phytoplankton, submerged aquatic vegetation and indicators of physical-chemical conditions have a higher confidence than indicators of the quality of benthic invertebrate communities. Our analyses also show that the perceived weaknesses of eutrophication assessments are due more to low confidence in reference conditions and acceptable deviations rather than in the monitoring data.

1. Introduction

Assessments of eutrophication status in Nordic countries have for decades been indicator based, but since 2002 Nordic eutrophication assessments have included the use of multi-metric indicator-based assessment tools.

The tools currently used in regard to assessment and classifications of eutrophication status in the Nordic countries are:

- The OSPAR Common Procedure (OSPAR COMP)
- The HELCOM Eutrophication Assessment Tool (HEAT)

OSPAR COMP is currently used by Denmark, Norway and Sweden.

HEAT, which is based based on the OSPAR COMP, but makes use of the principles in the EU Water Framework Directive for assessment of ecological status of coastal waters, is currently used by the countries bordering the Baltic Sea.

The overall objective of the CONFIRM project has been to develop a prototype tool for confidence rating of marine eutrophication assessments in the Nordic countries. Such assessments have been carried out by Nordic countries under the OSPAR Comprehensive Procedure, HELCOM EUTRO and HELCOM EUTRO-PRO, the latter encompassing the draft HELCOM Eutrophication Assessment Tool (HEAT). Detailed description of these widely accepted assessment tools can be found in OSPAR 2001, Ærtebjerg et al. 2006, HELCOM 2006 and HELCOM 2009.

Taking into account that management actions and implementation of measures in relevant sectors (e.g. agriculture, industry, waste water treatment) require major investments, there is an urgent need for tools enabling us to assess confidence in assessment results. In the Water Framework Directive (WFD) it is stated that the “level of confidence and precision” has to be “estimated” and has to be “adequate”, but how to judge adequate confidence and what further actions need to be taken if confidence is not adequate are not described.

CONFIRM attempts to fill the need for a confidence rating tool outlined above via the following activities:

This key task resulted in principles for confidence rating and a score card for confidence rating. The latter is done per indicator and focuses on the following questions:

- To what extent can we trust the reference conditions set?
- How reliable is acceptable deviation (boundary setting) determined?
- What is the quality of monitoring data and how would less than high quality data affect the outcome of the assessment?

The sequence of actions leading to confidence rating of a eutrophication assessment should include the following steps: (1) calculation of an indicator score, (2) calculation of category/quality element score, and finally an overall confidence rating directly linked to the overall result of the eutrophication assessment. Confidence score of an indicator could be based on amount of data available and the variation in data used to establish both present status and reference conditions for the indicator. This concept was discussed and agreed at the project initiation meeting, where Denmark, Finland, Germany, Holland, Norway, Sweden and the United Kingdom were represented. A key issue for CONFIRM was to arrive at jointly agreed Nordic principles and to describe these in detail.

Indicators within a group of indicators (category/quality element) are often treated as if they have the same importance for making the assessment. However, some indicators are better indicators than others due to difference in cause-effect relations, and often the data behind an indicator varies in quality and quantity. Consequently, it makes sense to include a weighting between indicators and even between categories/quality elements. While methods for incorporating a weighting procedure are numerically straightforward the principles on how to weight different indicators can be debated. One option could be a weighting determined by dose-response relationship between indicator and nutrient load (i.e. steepness of relation) or weighting could be based on the number of published examples of significant relationships. Concepts regarding weighting were also discussed at the kick-off meeting and details worked out subsequently. A weighting option has been included in the tools developed by CONFIRM. Parallel to weighting of indicators, the confidence score needs to be weighted too.

Draft tools for confidence rating were developed and tested via three case studies: one Finnish, one Norwegian, and one Swedish. The justification for carrying out these case studies was: (1) to test the draft tools, (2) to evaluate their usefulness, (3) to propose improvement if relevant, and (4) to provide practical examples to end-users. Based on the experiences and recommendation from the case studies, the principles and tools were adjusted where relevant. Further, the text describing the principles and the guidance explaining how to use the tools was updated.

The CONFIRM tools for confidence rating of eutrophication assessments are available to all interested persons and institutions in the following versions: (1) an Excel version focusing on OSPAR COMP, and (2) an Excel version focusing on HEAT. Both versions are described and downloadable free of any charge at the following website:

http://heat.dhigroup.com/confidence_rating/confirm.html

Results have been submitted for publication as a scientific paper.

2. Methods

2.1 Data and information sources

This work is based on existing data used in the HELCOM assessment of Baltic Sea eutrophication 2001–2006 (HELCOM, 2009). Most of the monitoring data representing actual status originate from the HELCOM Cooperative Monitoring in the Baltic Marine Environment (HELCOM COMBINE) Programme although some come from national monitoring and assessment activities. HELCOM COMBINE is a cooperative monitoring programme shared by the Baltic countries. Andersen et al. (submitted) describes the data origin in more detail.

An overview of the “assessment units” used is shown in Fig. 1.

Figure 1. Map of the “assessment units” used in this report.

2.2 Primary assessment: Eutrophication status

The assessment tool used to illustrate confidence ratings is the HELCOM Eutrophication Assessment Tool (HEAT). This is a multi-metric indicator based tool developed for the assessment and classification of the eutrophication status of the entire Baltic Sea. A total of 189 areas were assessed using indicators where information on reference conditions (RefCon) and acceptable deviations from reference conditions (AcDev) could be combined with national monitoring data describing the actual status (AcStat) for the period 2001–2006 (Andersen et al., submitted). Using the described RefCon, AcDev and AcStat concepts, the basic assessment principle is:

$$\text{EutroQO (indicator)} = \text{RefCon} \pm \text{AcDev} \quad (1)$$

where EutroQO is an “eutrophication quality objective” (or target), RefCon is an “anchor” for the assessment while AcDev is the “yardstick”.

For indicators which have a positive response to nutrient inputs, the classification is determined by the following:

$$\begin{aligned} &\text{If } \text{AcStat} < \text{RefCon} + \text{AcDev}, \\ &\text{then the EutroQO (or target) is met} \end{aligned} \quad (2).$$

Similarly for indicators having a negative response to nutrient inputs:

$$\begin{aligned} &\text{If } \text{AcStat} > \text{RefCon} - \text{AcDev}, \\ &\text{then the EutroQO (or target) is met} \end{aligned} \quad (3).$$

The HEAT tool integrates the elements described above and is based on: (1) Indicators representing well documented eutrophication effects with synoptic information on Reference Conditions (RefCon), Acceptable Deviations (AcDev) and Actual Status (AcStat), (2) Quality Elements *sensu* the EU Water Framework Directive (see Anon., 2000), (3) HELCOM Ecological Objectives (see HELCOM, 2009), (4) the relative weighting of indicators within Quality Elements, and (5) integration of the Quality Elements used into a final assessment based on the “One out – all out” principle *sensu* the Water Framework Directive.

The primary assessment calculated by HEAT is a classification of eutrophication status in five classes: High, Good, Moderate, Poor and Bad. The EutroQO (or target) corresponds to the boundary between good and moderate status. Assessment results are described in HELCOM (2009) and Andersen et al. (submitted).

2.3 Secondary assessment: Confidence rating

HEAT also produces an overall confidence rating for each indicator, where scores are assigned to each RefCon, AcDev and AcStat value. This scoring is based on expert judgment, where the quality of RefCon, AcDev and AcStat is assigned to one of three classes: High (score = 1), Good (score = 2), and Low (score = 3). High and Good are considered acceptable, while Low indicates a problem related to the quality of the input parameters. We acknowledge that the system has a degree of subjectivity since it relies on expert judgment. Details of the scoring principles are described in section 2.4.

A so-called Final Confidence Rating (FCR) is then calculated in three steps: After combining the RefCon, AcDev and AcStat scores into an Interim Indicator Confidence (see Table 1), then the Quality Element Interim Confidence (QE-IC) is calculated by taking the weighted arithmetic mean of the confidences of the indicators within the Quality Element (QE). The Final Confidence Rating (FCR) for a station/water body is then obtained from the arithmetic mean of Quality Element Interim Confidences (QE-IC). In calculating the FCR, the quality elements are weighted equally, though quality elements not having any indicators are ignored. For example, at a station where only two Quality Elements have indicators, the final confidence is arrived at by giving a weighting of 50% to each of these two QE's. For each station, separate Interim Confidence ratings are also calculated for RefCon, AcDev and AcStat (respectively RefCon-IC, AcDev-IC, AcStat-IC) by taking the arithmetic mean of the values for all indicators from all Quality Elements, without any weighting. All HEAT spread sheets including the secondary assessment of confidence can be found in the Electronic Supplementary Material in Andersen et al. (2010).

Experiences from the HEAT classifications have lead to a principle where the final classification of eutrophication status has to be based on at least two, but preferably at least three QE's, with ideally a minimum of two indicators per QE (data not shown). This has been incorporated in two ways. Firstly, a QE with only one indicator has its QE-IC reduced by 25%. Secondly, if the assessment is based on only a single QE, its FCR is reduced by 50%. The FCR has three quality classes: High (100–75%), Good (75–50%), and Low (50–0%). This is comparable to the method used for analysis of Data Completeness and Reliability (DCR) in the ASSETS tool developed for assessment of eutrophication status of estuaries in the United States (NOAA, 2007). In HEAT, High and Good confidence ratings are considered acceptable, while Low indicates a problem related to the quality of the eutrophication classification.

Table 1 – Scoring matrix

Scores			Σ	Ind_Conf
RefCon	AcDev	AcStat		
1	1	1	3	100%
1	1	2	4	83%
1	2	2	5	67%
2	2	2	6	50%
2	2	3	7	33%
2	3	3	8	17%
3	3	3	9	0%

A score of 3 indicates the best possible classification where a "High" quality rating has been assigned to RefCon, AcDev and to AcStat. A score of 9 indicates the worst possible classification, where all scores are rated "Low". The indicator scores from 3 to 9 are converted to an indicator confidence rating ranging from 0 to 100%. The worst possible result is assigned a confidence score of 0% and the best possible score is equivalent to a confidence score of 100%. The confidence score for other combinations is arrived at by interpolating linearly.

2.4 Scoring principles

Initially, indicators used to assess open water eutrophication status were provisionally scored by a group of national experts (6 persons) from Finland (lead country), Denmark, Germany, Poland and Sweden. For coastal waters, a provisional scoring was done by 2–3 national experts from each of the Baltic Sea countries.

The type of data used for setting RefCon values, whether it be historical data, modelling or expert judgment (HELCOM, 2009), has implications for the scoring. In many cases, historical or modelled data are used directly, especially if the data are already published or if the methods are in line with the monitoring methods used for AcStat. Because RefCons are commonly based on a combination of methods, e.g. (1) historical data and statistical modelling, (2) historical data and dynamical modelling, or (3) historical data and expert judgment, the level of confidence in the RefCon has to be taken into account.

The sources of information on AcDevs, and hence the scoring, differ slightly for open and coastal waters. For open waters, we generally use +50% and -25%, but other values are used if justified (HELCOM, 2006; OSPAR, 2008; HELCOM, 2009; Andersen et al., submitted). For coastal waters, we use AcDevs originating from the WFD implementation process as far as possible (e.g. Anon., 2008a, and also Andersen et al., 2004; Krause-Jensen et al., 2005; Henriksen, 2009; Lysiak-Pastuszek et al., 2009a, b, c). If no coast-specific AcDevs are available, we use the AcDevs derived for open waters (see HELCOM, 2006 and Andersen et al., submitted). These different approaches are taken into account where scoring the indicators, since coastal AcDevs in most cases are better justified and documented compared to those for open waters. In addition, AcDevs larger than +50% and -25% are in general considered outside the range of minor or slight deviations from RefCons and are therefore given a lower confidence score.

The scoring of AcStat, where the information sources are mostly in situ monitoring (and in a few cases also modelled data) is based on the reliability

of the observations, their spatial coverage and frequency. The confidence assessment could also consider the number of years of adequate data, QA/QC procedures, and whether data have been reviewed and published. A non-exhaustive list of issues to consider when scoring the quality of an indicator is presented in Table 2. The issues of concern and the questions to be considered are closely interconnected and sometimes overlap.

Table 2 – Key issues of concern and questions to be addressed when scoring an indicator in (HEAT)

Indicator part	Concerns and questions	Yes	/	No
RefCon	Are reliable historical data used?	↑	or	↓
	Is the method used then comparable with today monitoring methods?	↑	or	↓
	Is the spatial and temporal coverage for historical data representative?	↑	or	↓
	Are RefCon data modelled and justified?	↑	or	↓
	Are RefCon data from so-called reference/undisturbed sites?	↓↓	or	↑
	Is the influence of expert judgment low?	↑↑	or	↓↓
	Are the RefCon values published in a peer reviewed paper or report series?	↑↑	or	↓
AcDev	Open waters:			
	Is the indicator a HELCOM target indicator?	↑	or	↓
	Is the AcDev value lower than + 50% (for a positive indicator response to nutrient concentrations, such as chlorophyll concentration)?	↑	or	↓
	Is the AcDev value lower than -25% (for a negative indicator response, such as Secchi depth)?	↑	or	↓
	Is a functional relation established between nutrient concentrations and indicator response?	↑	or	↓
	Is the influence of expert judgement low?	↑	or	↓
	Is the AcDev value published in a peer reviewed paper or report series?	↑↑	or	↓
	Coastal waters:			
	Does the AvDev value originate from the WFD implementation?	↑	or	↓
	Is the AcDev value lower than + 50%?	↑	or	↓↓
	Is the AcDev value lower than -25%?	↑	or	↓↓
	Is a functional relation established?	↑	or	↓
	Is the degree of expert judgment low?	↑	or	↓
	Is the AcDev value published in a peer reviewed paper or report series?	↑↑	or	↓
AcStat	Does the indicator represent eutrophication well?	↑	or	↓↓
	Is the indicator reported in a HELCOM Indicator Fact Sheet?	↑↑	or	↓
	Is the spatial coverage adequate?	↑	or	↓↓
	Is the temporal coverage (frequency) adequate and does it match seasonality?	↑	or	↓↓
	Is the uncertainty known and low?	↑↑	or	↓
	Are AcStat data published in a peer reviewed paper or report series?	↑↑	or	↓

Calculations of eutrophication status and confidence. "↑↑" indicates a large positive influence on the score; "↑" indicates a positive influence; "↓" indicate a negative influence; and "↓↓" indicates large negative influence. Note 1: Reference sites no longer exist within the Baltic Sea; 2: More information is available via www.helcom.fi.

In the second round of the confidence assessment, the scoring for open and coastal waters was tentatively checked by the convener of the assessment work. The results were then, as a third round, presented, argued and finally agreed collectively to ensure that each expert's assessment methodology was

consistent, producing a harmonized, Baltic Sea-wide assessment (see HEL-COM, 2008). The final and fourth round in the assessment process included verification by the convener and final approval by national contact points (who were not necessarily the experts who carried out the assessment).

3. Results and discussion

Confidence ratings were made for 189 areas of the Baltic Sea including the Kattegat and Danish Straits. These areas consisted of 172 coastal areas and 17 open water bodies. A summary of all confidence ratings is presented in Table 3.

Table 3 – Rating of confidence in 14 Baltic Sea basins including the Kattegat and the Danish Straits

Basin	High	Good	Low	Total
Bothnian Bay and the Quark	0 (0%)	6 (55.4%)	5 (45.5%)	11 (100%)
Bothnian Sea	0 (0%)	9 (40.9%)	13 (59.1%)	22 (100%)
The Archipelago and Åland Seas	0 (0%)	5 (83.3%)	1 (16.7%)	6 (100%)
Baltic Proper, northern parts	0 (0%)	32 (78.0%)	9 (22.0%)	41 (100%)
Gulf of Finland	0 (0%)	18 (90%)	2 (10%)	20 (100%)
Baltic Proper, Eastern Gotland Basin	1 (11.1%)	7 (77.8%)	1 (11.1%)	9 (100%)
Gulf of Riga	0 (0%)	5 (83.3%)	1 (16.7%)	6 (100%)
Western Gotland Basin	0 (0%)	17 (85.0%)	3 (15.0%)	20 (100%)
Gulf of Gdansk	1 (20.0%)	3 (60.0%)	1 (20.0%)	5 (100%)
Bornholm Basin	2 (14.3%)	9 (64.3%)	3 (21.4%)	14 (100%)
Arkona Basin	0 (0%)	4 (100%)	0 (0%)	4 (100%)
Kiel Bight and Mecklenburg Bight	0 (0%)	5 (100%)	0 (0%)	5 (100%)
Danish Straits including the Sound	6 (60.0%)	4 (40.0%)	0 (0%)	10 (100%)
Kattegat	5 (31.3%)	10 (62.5%)	1 (6.3%)	16 (100%)
Total	15 (7.9%)	134 (70.9%)	40 (21.2%)	189 (100%)

High and Good are considered acceptable, while Low indicate a problem related to the quality of the eutrophication classification. Information about the eutrophication status of these basins/areas can be found in HELCOM (2009) and Andersen et al. (submitted).

3.1 RefCon, AcDev and AcStat

For RefCons, the average confidence score was 0.75, indicating that the experts had High confidence in them. Open waters had an average RefCon confidence of 0.67, while coastal waters had an average RefCon confidence of 0.76 (Figure 2). For AcDev, the pattern was the same: The average confidence of AcDev was 0.46, while open waters and coastal waters had average AcDev confidence of 0.53 and 0.46, respectively. The AcStat confidence was somewhat higher with an average AcStat confidence of 0.86. Open waters had an average AcStat confidence of 0.96, while the AcStat confidence for coastal waters was only 0.85.

Figure 2. Confidence ratings of open waters ($n=17$) and coastal waters ($n=172$) in the Baltic Sea in regard to indicators, quality elements and final confidence ratings (FCR). For the indicators, confidence ratings of reference conditions (RefCon-IC), acceptable deviation (AcDev-IC) and actual status (AcStat-IC) are compared. For the quality element, we compare the confidence rating of phytoplankton (PHY-IC), submerged aquatic vegetation (SAV-IC), benthic invertebrate communities (BIC-IC) and physical-chemical indicators (PC-IC). The horizontal lines in regard to indicators and quality elements represent the boundary between Good and Low confidence. Please note the y axes differ in the FCR figures.

AcStat and RefCon generally had an average High confidence. We believe there are two reasons for this. Firstly, the monitoring activities carried out by the Baltic Sea countries through the HELCOM COMBINE Programme generally hold a high quality in terms of the methods used, temporal and spatial resolution as well as QA/QC procedures. Secondly, much effort has been put into establishing RefCon values, both for open (HELCOM, 2006) and coastal waters (Anon., 2008a).

The Low average confidence of AcDev, which are probably more difficult to establish than RefCon, leave room for improvements, especially for coastal waters. Despite Baltic EU Member States having spent considerable resources on implementation of the WFD, there seems to be an urgent need for improving the boundary setting, since this sets the target for having an acceptable or unacceptable ecological status.

3.2 Quality elements

The QE based on phytoplankton indicators had an average confidence of 0.71, indicating Good confidence (Table 1). The indicators used for the phytoplankton quality element were mostly, though not exclusively, based

on chlorophyll *a*. Assessments of open waters and coastal waters had an average confidence of 0.74 and 0.71, respectively. Submerged Aquatic Vegetation (SAV) is used only for assessment of eutrophication status of coastal waters because of depth limitation. The average confidence for SAV was 0.64, indicating Good confidence. For the QE on Benthic Invertebrate Communities, the average confidence was 0.33. For open waters it was 0.54, whilst it was only 0.30 for coastal waters. For the QE including physical-chemical indicators, the average confidence was 0.54. Open waters and coastal waters had an average confidence of 0.62 and 0.53, respectively.

For open waters, the ranking of Quality Elements in terms of average confidence went from Phytoplankton (0.74) > Physical-chemical conditions (0.62) > Benthic invertebrate communities (0.54), all holding a Good confidence. For coastal water the picture was different, here the rank was Phytoplankton (0.71) > Submerged Aquatic Vegetation (0.64) > Physical-chemical conditions (0.57) > Benthic invertebrates (0.30), the very last being a Low confidence QE.

To say that one or more Quality Element(s) is superior to the others is tempting. However, it should be kept in mind that all high-quality assessments of eutrophication status have to be based on a range of QEs and indicators. Causative factors, as well as primary and secondary eutrophication effects should be included in any assessment, meaning that indicators dealing with nutrient concentrations, phytoplankton and benthic communities should always be used. The benthic communities to be included in assessments of open (deep) waters will in practice be benthic invertebrates. For coastal waters, both submerged aquatic vegetation and benthic invertebrates should be included. However, based on this study, it appears that vegetation (0.64) is considered a more reliable Quality Element than invertebrate fauna (0.30) at least in terms of QE_IC. We assume there are two reasons for this: Firstly, the monitoring of SAV is mostly straightforward and based on the depth limits of the dominating SAV species. Secondly, indices currently used to assess the status of benthic invertebrate communities in coastal waters are not completely developed and perhaps not applicable in brackish or near-limnic coastal waters such as the Baltic Sea.

3.3 Areas

A total of 149 out of the 189 areas assessed were rated as having High or Good confidence. 40 areas were rated as having Low confidence (Table 3). In the Arkona Basin, Kiel Bight and Mecklenburg Bight as well as the Danish Straits including the Sound, none of the areas assessed hold Low confidence.

Discouragingly, some parts of the Baltic Sea had a high proportion of areas with Low confidence ratings (see Table 3), e.g. the Bothnian Bay (5 out of 11 or 45.5%), and the Bothnian Sea (13 out of 22 or 59.1%). The underlying reasons for this are difficult to deduce. Meaningful explanations could

be that the Gulf of Bothnia, surrounded by areas of low population density, is normally regarded as being unaffected or only slightly affected by eutrophication. Hence, despite about 40% of the runoff to the Baltic draining into the gulf, monitoring activities are relatively modest. This may result in a limited number of available indicators, low sampling frequencies, indicators with inaccurate setting of RefCons, or AcDevs that are outside the range of what is normally interpreted as minor or slight deviation from RefCons. Combinations of any or all of these are likely to result in a low FCR.

An important result is that none of Baltic Sea basins had monitoring and assessment programmes which averaged High confidence (see Table 3).

For all areas, the rank of the confidence was as follows: the Danish Straits including the Sound (0.75) > Kattegat (0.69) > Arkona Basin (0.67) > Kiel Bight and Mecklenburg Bight (0.64) > Western Gotland Basin (0.61) > Gulf of Finland (0.59) > Baltic Proper (0.58) > Bornholm Basin and Gulf of Riga (both 0.57) > Gulf of Gdansk (0.55) > Eastern Gotland Basin (0.54) > Archipelago and Åland Seas (0.52) > Bothnian Bay and the Quark (0.50) > Bothnian Sea (0.46). The monitoring activities in the majority of areas can hardly be reduced without compromising the quality of future eutrophication assessments. The monitoring of the Bothnian Sea, and perhaps also some of the other areas (e.g. coastal parts of the northern parts of the Baltic Proper, Gulf of Gdansk and the Bornholm Basin), ought to be improved in order to improve any future assessment of eutrophication status in these areas.

Of the 189 areas assessed, 13 were considered to be unaffected by eutrophication (HELCOM, 2009 and Andersen et al., submitted). However, only four of these were rated as having High or Good confidence. This does not imply that the remaining nine areas are mis-classified, but it could indicate that the quality of these status classifications can be questioned. False positive classifications (areas mis-classified as being affected by eutrophication) are perhaps a more worrying scenario, at least in budgetary terms, since investment in load reductions might not be scientifically based. Here, it would be prudent to state that the costs of acting on incomplete information or knowledge are generally believed to be significantly greater than the costs of obtaining the information.

3.4 Countries

When looking at individual countries (data not shown), it is clear that some countries (Denmark, Estonia, Germany) had better coastal monitoring programmes (providing data on AcStat) and indicators (providing data on Ref-Con and AcDev values) than others, leading to Good confidence in most places. Finland, Lithuania, Poland and Sweden had acceptable programmes, but there might be areas of concern such as in the western coastal waters of the Gulf of Bothnia (see Table 3). Poland might also have reason for concern in some lagoons. The FCR's made for Latvian and Russian coastal

waters indicate monitoring activities below par, mostly caused by insufficient monitoring activities, both in terms of spatial and temporal coverage as well as numbers of indicators.

None of the Baltic Sea countries had monitoring programmes which averaged High confidence. For coastal waters, the rank of the countries was Denmark (0.70) > Estonia (0.65) > Germany (0.64) > Sweden (0.57) > Finland (0.55) > Lithuania (0.54) > Poland (0.51) > Russia (0.47) > Latvia (0.42). The programmes in Denmark, Estonia, Finland, Lithuania, Sweden and Poland cannot be reduced without compromising the requirements of the WFD.

An interesting result was that open waters, often being monitored collectively by 2–3 countries per station, had a better average confidence (0.66) than coastal waters (0.57). There are probably several explanations. Firstly, the monitoring of open waters has been harmonized and coordinated by HELCOM for almost three decades, leading to a good understanding of the monitoring system. Secondly, HELCOM requirements have traditionally put more focus on the shared open waters rather than the national coastal waters. Thirdly, cruises, even in coastal waters, are expensive and while monitoring of offshore regions has benefitted from the cost-sharing of the cooperative HELCOM monitoring, some Baltic Sea states have had difficulty monitoring their local archipelagos. Furthermore, at least for some parameters, variability (particularly spatially) is higher in coastal than open waters.

4. Conclusions

We have shown that confidence rating of marine eutrophication (and other status assessments) can be performed using a practical, indirect and non-statistical approach. The approach should at this stage be seen as a first step only, paving the way for more sophisticated and formal methodologies. We do not intend to promote the suggested methodology, but do stress the need for confidence rating eutrophication status classifications. However, we consider the methodology reported here to be a useful tool in regard to: (1) indirect assessment of the accuracy and precision of the primary assessment of eutrophication status, (2) future revision of marine monitoring networks as well as (3) implementation of science-based nutrient management strategies, especially those pursuant to the EU Water Framework Directive and the EU Marine Strategy Framework Directive.

The reported method has several uses, the following five being the most obvious: (1) it can be used to implement or strengthen nutrient management strategies in the upstream catchment where the primary assessment indicates that the “downstream” area in question is “affected by eutrophication” and where the FCR is High or Good, (2) it can be used to improve scientific knowledge of the basis on which load reductions are being decided or it can prevent mistaken investments in load reductions where the primary assessment indicates that the area in question is an “area affected by eutrophication” but with a low FCR, (3) it can be used to improve setting of RefCons and AcDevs, e.g. by encouraging applied research in regard to boundary/target setting, (4) it can be used to improve monitoring activities in areas where the current (2001–2006) confidence is low, and (5) it can be used to warn against unsupported reductions in monitoring activities, particularly where a reduction would change the FCR from Good to Low.

The approach could also be useful in regard to assessments according to the Water Framework Directive (Anon., 2000) and the Marine Strategy Framework Directive (Anon., 2008b), since EU Member States are already required not only to assess ecological/environmental status of transitional and coastal marine waters, but also to provide information about levels of confidence and precision in assessments (Irvine, 2004).

The guiding principle for the use of the scoring results and confidence estimates in managing and developing marine environmental assessments should be to strengthen and/or maintain that which already has an appropriate quality and to improve what does not have a good quality. By doing so, we are likely, in the longer term, to end up with better monitoring programmes, better indicators, better environmental targets, better assessments, and ultimately more effective science-based nutrient management strategies.

5. Acknowledgements

This study is funded by Nordic Council of Ministers (CONFIRM). JHA has been partly funded by HELCOM (via the EUTRO-PRO project). We would like to thank the following colleagues in the HELCOM EUTRO-PRO project: Mats Blomkvist, Ulrich Claussen, Vivi Fleming-Lehtinen, Pirkko Kauppila, Aiste Kubiliute, Elisabeth Lysiak-Pastuszek, Georg Martin, Günther Nausch, Alf Norkko and Anna Villnäs for providing the data on which this study is based. Special thanks to the DHI-NTU Water & Environment Research Centre in Singapore and to Uwe Brockmann, David Connor, Maria Laamanen and Flemming Møhlenberg for stimulating discussions.

6. References

- Andersen, J.H., Conley, D.J., Hedal, S., 2004. Palaeo-ecology, reference conditions and classification of ecological status: The EU Water Framework Directive in practice. *Marine Pollution Bulletin*. 49, 282–290.
- Andersen, J.H., Murray, C., Kaartokallio, H., Axe, P., Molvær, J., 2010. A simple method for confidence rating of eutrophication status classifications. *Marine Pollution Bulletin*. 60, 919–924.
- Andersen, J.H., and 18 co-authors, 2010. Getting the measure of eutrophication in the Baltic Sea: towards better assessment principles and methods. Accepted by *Biogeochemistry*.
- Anon., 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Communities* L 327/1.
- Anon., 2008a. Commission decision of 30 October 2008 establishing, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise. *Official Journal of the European Union* L 332/20.
- Anon., 2008b. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). *Official Journal of the European Communities* L 164/19.
- Claussen, U., Zevenboom, W., Brockmann, U., Topcu, D., Bot, P., 2009. Assessment of the eutrophication status of transitional, coastal and marine waters within OSPAR. *Hydrobiologia* 629(1), 49–58.
- Borja, A., Bald, J., Franco, J., Larreta, J., Muxika, I., Revilla, M., Rodríguez, J.G., Solaun, O., Uriarte, A., Valencia, V., 2009. Using multiple ecosystem components, in assessing ecological status in Spanish (Basque Country) Atlantic marine waters. *Marine Pollution Bulletin* 59, 54–64.
- EEA, 2001. Eutrophication in Europe's coastal waters. Topic report 7/2001. Available at http://www.eea.europa.eu/publications/topic_report_2001_7/Topic_Report_7_2001.pdf.
- Krause-Jensen, D., Greve, T.M., Nielsen, K., 2005. Eelgrass as a Bioindicator under the Water Framework Directive. *Water Resources Management* 19, 63–75.
- HELCOM, 2008. Minutes of the eight meeting of HELCOM project to elaborate the HELCOM Baltic Sea-wide integrated thematic assessment on eutrophication (HELCOM EUTRO-PRO). 14 pp.
- HELCOM, 2009. Eutrophication in the Baltic Sea – An integrated thematic assessment of eutrophication in the Baltic Sea region. *Baltic Sea Environmental Proceedings* No. 115B. Helsinki Commission. 148 pp. Available via <http://www.helcom.fi>.
- Henriksen, P., 2009. Reference conditions for phytoplankton at Danish Water Framework Directive intercalibration sites. *Hydrobiologia* 629(1), 255–262.
- Irvine, K., 2004. Classifying ecological status under the European Water Framework Directive: the need for monitoring to account for natural variability. *Aquatic Conservation: Marine and Freshwater Ecosystems* 14, 107–112.
- Lysiak-Pastuszek, E., Krzyminski, W., Lewandowski, L., 2009. Development of tools for ecological quality assessment in the Polish marine areas according to the Water Framework Directive. Part I – Nutrients. *Oceanological and Hydrobiological Studies*, 38(3): 87–99.
- Lysiak-Pastuszek, E., Krzyminski, W., Lewandowski, L., 2009. Development of tools for ecological quality assessment in the Polish marine areas according to the Water Framework Directive. Part II – Chlorophyll-a. *Oceanological and Hydrobiological Studies*, 38(3): 101–112.
- Lysiak-Pastuszek, E., Krzyminski, W., Lewandowski, L., 2009. Development of tools for ecological quality assessment in the Polish marine areas according to the

- Water Framework Directive. Part III – Secchi depth. *Oceanological and Hydrobiological Studies*, 38(3): 113–124.
- NOAA, 2007. Effects of Nutrient Enrichment in the Nation's Estuaries: A Decade of Change, National Estuarine Eutrophication Assessment Update. NOAA Coastal Ocean Program Decision Analysis Series No. 26. National Centers for Coastal Ocean Science, Silver Spring, MD. 322 pp. Available via <http://www.noaa.gov>.
- OSPAR, 2008. Second integrated report on the eutrophication status of the OSPAR maritime area. Eutrophication Series. OSPAR Commission. 107 pp. Available via <http://www.ospar.org>.
- Ærtebjerg, G., Andersen, J.H., Hansen, O.S., (Eds.), 2003. Nutrients and Eutrophication in Danish Marine Waters. A Challenge for Science and Management. National Environmental Research Institute. 126 pp.

Appendix 1: Finnish case study

*Hermanni Kaartokallio, Vivi Fleming-Lehtinen, Anna Villnäs & Alf Norkko,
SYKE Marine Research Centre, Finland*

A1.1 Introduction/background

This case study presents an eutrophication assessment confidence rating work carried out in conjunction with the first HELCOM thematic eutrophication assessment of the Baltic Sea (HELCOM 2009a) at the Finnish Institute of Marine Research (presently Finnish Environment Institute – Marine Research Centre). The study includes a HEAT assessment with three quality elements for four Northern sub-basins as well as with two quality elements (QE) for three eastern-southern sub-basins. For northern sub-basins some quality elements include multiple indicators, otherwise there is only one indicator per quality element, reflecting the current availability of open-sea eutrophication indicators in the Baltic Sea. Indicators for category I, II and III eutrophication effects were included in the study. For the QE “plankton”, the summer surface chlorophyll *a* indicator reflects the biomass of planktonic primary producers. For the QE “invertebrate benthic fauna”, a gamma-diversity indicator based on the average number of taxa was employed. The QE “physico-chemical features” included Secchi depth and winter dissolved nutrients ($\text{NO}_2 + \text{NO}_3$, PO_4) in the surface waters. Our aim was to employ concrete, robust and transparent criteria in the confidence assessment of the indicators selected for the current study.

A1.2 General principles in confidence rating and weighting

Reference conditions (RefCon), boundary setting between acceptable and unacceptable status (acceptable deviation, AcDev), actual status determination (for the years 2001–2006, AcStat) and weighting between indicators in each quality element for selected open Baltic Sea sub-basins were confidence rated in the study. In the confidence rating of particular indicators, the factors indicated in the following were taken into account: For RefCon and AcStat confidence ratings, spatial and temporal representativeness of historical data were the main decisive factors as reference setting was essentially based on historical data (Secchi depth based only on historical data, chlorophyll *a* and nutrients based on historical data and using some extrapolation) or on the entire monitoring dataset held by FIMR (benthic invertebrates monitoring data, starting from the late 1960s).). Confidence rating of

boundary setting was based primarily on the availability of detailed scientific information on functional relations between the eutrophication pressures and direct and indirect responses of the particular indicator. As the quality of indicators included in this study was rather homogeneous, the principle in indicator weighting was based either on well-known ecosystem traits of the particular sub-basin (physico-chemical indicators in the Bothnian Bay) or areal coverage of particular indicators within the sub-basin (invertebrate benthic fauna in the Eastern Gotland Basin).

A1.3 Confidence rating of indicators used in the study

In the following chapter, indicators included in the study are described and principles of confidence rating of each indicator element (RefCon, AcDev, AcStat) are presented in detail. For the sake of clarity indicators are classified according to HEAT quality elements, although indicators are grouped in a slightly different manner into OSPAR COMP quality elements.

A1.3.1 Quality element: Plankton

A1.3.1.1. Indicator: Chlorophyll *a*

Phytoplankton respond rapidly to changes in nutrient levels and therefore the biomass and the species composition of phytoplankton can be used as indicators of eutrophication. The close connection between phytoplankton growth and nutrient concentrations in the water make phytoplankton a useful tool in the assessment of eutrophication status. The sole phytoplankton indicator employed in this study is the chlorophyll *a* concentration. This is a phytoplankton biomass indicator that has relatively coherent reference conditions for open Baltic Sea sub-basins. For this indicator, the combined HELCOM COMBINE (kept at ICES) and FIMR datasets (HELCOM 2006, Olsonen 2008) with June-September surface (≤ 2 m depth) observations were used (Figure 1). The indicator was constructed for four open Northern Baltic Sea sub-basins with sufficient data coverage for reference setting and actual status determination (Bothnian Bay, Bothnian Sea, Gulf of Finland, Northern Baltic Proper). Compared to other indicators, e.g. Secchi depth and benthic invertebrate diversity, the amount of chlorophyll *a* data available was limited, especially in the Gulf of Bothnia.

Reference conditions

Reference setting for chlorophyll *a* in the Gulf of Bothnia (including the Bothnian Bay and the Bothnian Sea) and Gulf of Finland was done according to principles set out during HELCOM indicator development work at FIMR HELCOM (2006,) and Fleming-Lehtinen (2007) using respective sub-basin median values of earliest historical observations typically originating from the mid 1970s to early 1980s for chlorophyll *a* (Fleming-Lehtinen

et al. 2008). The trends from the early observations were extrapolated to the period before the 1950s, which was considered an appropriate period for reference conditions. The results of the extrapolation were compared to and adjusted with those derived by model simulations (Schernewski & Neumann 2005). In Northern Baltic Proper a reference condition derived from BSEP 104 (HELCOM 2006) was used.

Confidence rating of the reference conditions was done on the basis of the availability and spatio-temporal and seasonal representativeness of the historical data (Figure 1). All four sub-basins were rated as confidence class 2 because of the lack of historical data from the reference period, the long time difference (25–30 years) between the oldest data and the reference period and subsequent uncertainties arising from the use of expert judgment/modelling results in filling the time gap.

Acceptable deviation

A normative maximum acceptable deviation of 50% was used for all sub-basins. This 50% maximum acceptable deviation was agreed in the HELCOM thematic eutrophication assessment based on principles adopted in HELCOM (2006) and is compliant with WFD requirements.

The confidence rating on acceptable deviation was done on the basis of the underlying research data on functional relations. As a normative 50% maximum acceptable deviation (HELCOM 2006) was used for all four sub-basins, they were rated as confidence class 2.

Actual Status

Actual status was defined as a mean of summer (June–September, years 2001–2006) surface chlorophyll *a* observations for each open sub-basin included.

The confidence rating of actual status for the years 2001–2006 was done on the basis of the availability and spatio-temporal and seasonal representativeness of the assessment data. As the assessment data was deemed spatially sufficient and having the same seasonal and spatial coverage as the reference data (Figure 1) for open parts of the basins, all sub-basins were rated having confidence class 1. Overall limited availability of open-sea chlorophyll *a* data were taken into account in classification.

Weighting

No weighting was done because chlorophyll *a* was the only indicator representing the plankton quality element in this study.

Figure 1. Spatial coverage of the oldest historical and 2001–2006 summer (June–September) chlorophyll *a* data used in reference condition and actual status determinations respectively. Red dots show the locations of oldest data, green dots show the data used in status assessment. Red dots have been increased in size to make them visible. Basin abbreviations: BoB= Bothnian Bay, BoS= Bothnian Sea, GoF= Gulf of Finland, NBP= Northern Baltic Proper, EGB= Eastern Gotland Basin, BS= Bornholm Sea, AS= Arkona Basin.

A1.3.2 Quality element: Invertebrate Benthic Fauna

A1.3.2.1 Indicator: Average number of taxa

Benthic invertebrate communities are good indicators of environmental status. Because of their relative longevity (years to decades), the composition of benthic communities integrates environmental conditions over longer periods of time. Hence, variations in environmental characteristics, such as salinity, oxygen, food supply, biotic interactions, and different types of dis-

turbances (both natural and anthropogenic) are reflected in the composition of communities in time and space. While physiological tolerances to salinity and oxygen may play the most important role in defining the distribution of species, macrozoobenthic communities are generally food limited (Pearson & Rosenberg 1978) and the abundance and biomass of benthic invertebrates correlates to some extent with the deposition of pelagic organic material (Josefson & Conley 1997). Benthic environments are at the receiving end for the accumulation and burial of organic material, and healthy benthic communities play an important role in benthic-pelagic coupling and in mineralization of organic matter settling on the seafloor. Benthic communities generally respond to organic enrichment in a predictable manner. Initial stimulatory effects on benthos are gradually replaced by degradation of communities as eutrophication advances. Increasing organic enrichment and bottom-water hypoxia and anoxia alters benthic community composition by reducing sensitive species and increasing the proliferation of tolerant species. At advanced stages of eutrophication, significant reductions in diversity and ecosystem function accompany these compositional changes.

The number of species was used as a parameter when defining eutrophication status of the open sea areas in the Baltic. Assessment was done for six open basins (Bothnian Bay, Bothnian Sea, Gulf of Finland, the Northern Baltic Proper, Eastern Gotland Basin, Bornholm Sea and the Arkona Basin). The Eastern Gotland basin was assessed in two parts (southern and northern/central parts). The number of species is defined as gamma diversity, describing diversity over a larger area (Gray 2000). The measure is based on point (alpha) diversity, i.e. the total number of species at a station per sampling occasion, including at least three replicates. Gamma diversity then describes the average diversity of stations in a sub-area per year (Figure 2). Reference values and acceptable deviations defined here for benthic communities in the open Baltic Sea are based on data derived from disturbed communities. The lack of reference data from unaffected benthic communities, constrains this evaluation but species diversity provides a reasonably good picture of the zoobenthic status in the Baltic Sea.

Figure 2. Example of gamma diversity $\pm$ standard deviation based on ca 15 stations in the Bothnian Bay during 1964–2006.

Reference conditions

When defining reference values and acceptable deviations for open Baltic Sea areas, data from ca 3000 sampling occasions at over 100 stations during 1964–2006 were included (FIMR HELCOM monitoring data (Olsonen 2008) with additional data provided from Sweden for the Bornholm and Arkona basins, Figure 3). No corrections have been made for seasonal variability within locations or areas. Generally only stations with a depth over 40 m were included. Anoxic and/or hypoxic periods were excluded from the data, as were occasions when there were zero or single occurrence of species. The reference value for each sub-basin was identified as the average of the 10 % of the maximum diversity values during the monitoring period (Figure 4). The maximum value was used only when the temporal data coverage was not sufficient. Confidence rating of class 1 was given for reference conditions in sub-basins where the data had a high reliability (single source, uniform methods) and a sufficient spatial and temporal coverage. In Eastern Gotland basin and the Bornholm Sea, occurrence of recurrent anoxia and lack of benthic fauna lowered the confidence of reference condition obtained to class 2. In the Arkona Basin insufficient reference data also justified a class 2 confidence rating.

Acceptable deviation

It was considered that the acceptable deviation should relate to natural fluctuations in the number of species in an area. Taking advantage of the long-term data, acceptable deviation was defined as relative standard deviation in an area per year. An average acceptable deviation for each sub-area was based on data from several years. The highest acceptable deviation allowed was limited to 40 %, which was exceeded in the southern part of the Eastern Gotland Basin and the Bornholm Basin; areas severely influenced by hypoxia. Confidence rating class 2 was given for all sub-basins assessed as the direct effect of eutrophication cannot be fully distinguished from other pres-

tures and natural variation with the method employed here. Also lack of data contributed in the confidence rating in the Northern Baltic Proper, southern parts of the Eastern Gotland Basin and the Arkona Basin.

Figure 3. Spatial coverage of benthic invertebrate data used in reference condition setting and actual status determination as described in the text. Red dots show the reference data and green dots the data used in actual status determination. For basin abbreviations see Figure 1.

Figure 4. Reference values and the border between good and moderate ecological status (showing Reference conditions-Acceptable deviation). Benthic invertebrate status is described as an average for the assessment period 2001–2006 (Villnäs & Norkko, in prep.).

Actual status

Actual status was defined for the period of 2001–2006 using the same dataset as for reference conditions and acceptable deviation (figures 3). Confidence rating for the actual status assessment was class 1 for all sub-basins except for the southern parts of the Eastern Gotland Basin, Bornholm Sea and Arkona Basin, where a rating of class 2 was implied due to paucity of data.

Weighting

Within the quality element, weighting was only done in the Eastern Gotland basin, where North and Central parts were given a weight 67% and southern parts 33% according to areal contribution of these parts to the basin area.

A3.3 Quality element: Physico-chemical features

A3.3.1 Indicator: Summer Secchi depth

Water transparency is a measure of the clarity of water. It indicates attenuation of light penetrating into water and is governed by its absorption and scattering properties. Scattering and absorption of light is dependent on the amount of particulate matter and dissolved substances in the water. The material in the water is typically living or dead organic particles (e.g. phytoplankton), inorganic small particles, or a combination of both and dissolved coloured substances e.g. humic substances.

Secchi depth, a measurement of water transparency, can be used as an indicator of eutrophication as increasing phytoplankton biomass is likely to have a high influence on water transparency in summer (defined here as from June to September). Changes in water transparency can thus be used to give an indication of changes in summer phytoplankton biomass. In this study, all available summer Secchi observations were pooled together to enable reliable assessment on a sub-basin scale. Data used in the open sea-assessment con-

tains all available data from FIMR (Olsonen 2008) and HELCOM COMBINE (kept at ICES) databases as well as additional data from several HELCOM contracting states (Latvia, Lithuania, Poland, and Sweden).

Reference conditions

Reference conditions were defined for each sub-basin. The reference condition used was the median value of a five-year period with the oldest available data from the beginning of 20th century. Reference conditions were derived from this historical data for all sub-basins included in this study. For details on the data see Fleming-Lehtinen et al. (2007).

Secchi depth reference conditions were rated as confidence class 1 for all sub-basins except the Arkona Basin on the basis of sufficient spatial distribution of historical data (Figure 5). The Arkona Basin received a confidence rating 2 due to the uneven spatial distribution of reference data into one south-north transect in the mid-basin (Figure 5).

Acceptable deviation

For Secchi depth, the acceptable deviation was set to the upper normative limit of -25 % of the reference conditions (HELCOM 2006). The main reason for selecting an acceptable deviation maximum limit lower than -50 % was that the relationship between Secchi depth and chlorophyll *a* concentration is not linear (Sandén and Håkansson 1996). When moving from high to lower Secchi depths, the corresponding change in chlorophyll *a* concentrations is steeper in the lower end of the Secchi depths. Therefore, it is reasonable to consider that the change in Secchi depths should not exceed 25 % decrease from the reference conditions.

Confidence rating on acceptable deviation was defined separately for each sub-basin. For the four northern sub-basins an analysis of functional relations between chlorophyll *a* and Secchi depth has been presented in HELCOM (2009b) and they accordingly received a confidence rating of class 1. Other sub-basins, where functional relations had not been analysed yet with sufficient detail were rated to confidence class 2.

Figure 5. Spatial coverage of oldest historical and 2001–2006 summer (June–September) Secchi depth data used in reference condition and actual status determinations, respectively. Red dots show the reference data and green dots data used in the actual status determination. For basin abbreviations see figure 1.

Actual status

The mean value for summer (June–September, years 2001–2006) observations was used in defining actual status for each sub-basin. The data used in actual status included data from FIMR and ICES databases as well as complementary data from HELCOM contracting states as mentioned above (HELCOM 2009a). Confidence rating of actual status was done on the basis of spatio-temporal and seasonal representativeness and sufficiency of the assessment data (Figure 5). As the assessment data was sufficient and spatially representative, all sub-basins included in this study were rated having confidence class 1.

Weighting

In the Bothnian Bay, the Secchi depth indicator was weighted 40%, whereas nutrient indicators were weighted 20% for NO_2+NO_3 and 40% for PO_4 due to generally prevailing hard P-limitation of phytoplankton primary production in the basin. The same weighting was used for HEAT (QE Physico-chemical elements) and OSPAR COMP (QE causative factors), as the Secchi depth indicator in the Bothnian Bay was considered to reflect more direct organic loading in the basin, hence belonging to QE “causative factors” rather than “Other possible effects” as in other sub-basins. In the other four sub-basins with Secchi and nutrient indicators, weighting was done on an equal basis, with all three indicators within the quality element weighted at 33%.

A3.3.2. Indicators: Winter surface nutrients (NO_2+NO_3 and PO_4)

Nutrient enrichment is a proximate cause of eutrophication, leading to the increased production of organic matter as phytoplankton need nutrients for growth. Accordingly, low concentrations of available nutrients in water will limit primary production. Increased production affects water transparency, phytoplankton communities, benthic fauna and vegetation as well as oxygen conditions. Seasonal variations in nutrient supply, removal and transformation processes create seasonal patterns in seawater nutrient concentrations. These are most pronounced for dissolved inorganic nitrogen and phosphorus. These dissolved nutrients are typically depleted in surface waters during summer and reach their maximum surface concentrations during winter, when nutrient regeneration exceeds nutrient usage in primary production. Winter dissolved nutrients are often used as indicators of eutrophication and were also employed in this study.

The two nutrient indicators included were the winter (December-February) sum of the concentration of dissolved nitrate+nitrite (NO_2+NO_3) and concentration of dissolved phosphate (PO_4) in the surface waters. For nutrient indicators, a combined HELCOM COMBINE/FIMR dataset (HELCOM 2006, Olsonen 2008) with December-February surface (≤ 2 m depth) observations was used. The indicator was constructed for five open Baltic Sea sub-basins with sufficient data coverage (Figures. 6 and 7) for reference setting and actual status determination (Bothnian Bay, Bothnian Sea, Gulf of Finland, Northern Baltic Proper, Eastern Gotland Basin).

Reference conditions

Reference setting for dissolved nutrient indicators (in all five sub-basins except the Northern Baltic Proper) followed principles set out during indicator development work in (HELCOM 2006, Fleming-Lehtinen 2007) using respective sub-basin median values of earliest historical observations typically originating from the mid- and late 1960s for nutrients (Fleming-Lehtinen et al. 2008). Trends from the early observations were extrapolated to the period before the 1950s, which was considered an appropriate period

for reference conditions. The results of the extrapolation were compared to and adjusted with those derived by model simulations (Schernewski & Neumann 2005). Northern Baltic Proper reference conditions, presented in BSEP104 (HELCOM 2006), were defined from historical data.

Confidence rating of reference conditions was based on the availability and spatio-temporal and seasonal representativeness of the historical data. All five sub-basins included were rated having confidence class 1 because of sufficient historical data. Although the oldest data is not directly from the reference period, the time gap between reference period and oldest data was remarkably shorter than for chlorophyll *a*, which can be seen as a factor decreasing uncertainty of reference setting by extrapolation and expert judgment. Northern Baltic Proper reference conditions were rated as having a confidence rating of 2 for both indicators, as we had no direct access to the underlying data.

Acceptable deviation

A normative maximum acceptable deviation of 50% was used for all sub-basins. 50% maximum acceptable deviation was agreed in the HELCOM thematic eutrophication assessment based on principles adopted in (HELCOM 2006) and is compliant with WFD requirements. Confidence rating of acceptable deviation was done on the basis of the underlying research data. As normative 50% maximum acceptable deviation (HELCOM 2006) was used for all four sub-basins, they were rated to confidence class 2.

Actual Status

Actual status was defined as a mean of winter (December-February, years 2001–2006) surface nutrient observations for each open sub-basin. Confidence rating of actual status for the years 2001–2006 was done on the basis of availability and spatio-temporal and seasonal representativeness of the assessment data. As the assessment data were sufficient and met the above mentioned criteria, all sub-basins were rated having confidence class 1.

Weighting

In the Bothnian Bay nutrient indicators were weighted 20% for $\text{NO}_2 + \text{NO}_3$ and 40% for PO_4 due to generally prevailing hard P-limitation of phytoplankton primary production in the basin. In the other four sub-basins including nutrient indicators weighting was done on an equal basis, with all three indicators within the quality element receiving weights of 33%.

Figure 6. Spatial coverage of oldest historical and 2001–2006 winter (December–February) surface $\Sigma\text{NO}_2 + \text{NO}_3$ concentration data used in reference condition and actual status determinations, respectively. Red dots show the reference data and green dots data used in the actual status determination. For basin abbreviations see figure 1.

Figure 7. Spatial coverage of oldest historical and 2001–2006 winter (December–February) surface PO_4 concentration data used in reference condition and actual status determinations, respectively. Red dots show the reference data and green dots data used in the actual status determination. For basin abbreviations see figure 1.

A1.4 Summary of classification results

A summary of classification results with HEAT and OSPAR COMP tools are shown in tables 1 and 2, respectively. Both eutrophication/ecological status classifications and confidence ratings are presented. Overall, the final results obtained with both tools agree very well, the only difference being in status classification in the Bothnian Bay. Here, the HEAT tool gives an overall classification of “good”, whereas OSPAR COMP status is “moderate” due to the difference in classification of quality element physico-

chemical features/Causative factors (see tables 1 and 2). Status classification was acceptable (above the good-moderate border) in the Bothnian Bay and Arkona Basin (using HEAT tool) or only in the Arkona Basin (Using OSPAR COMP tool). The Bothnian Sea was classified as “poor” and all the other sub-basins as “bad” using both tools. The confidence rating was identical using both tools, giving a final confidence class II to all sub-basins except the Arkona Basin, which received confidence class III. Lowered confidence of individual quality elements in the southern sub-basins (see tables) was mostly due to the low number (typically only 1 indicator/quality element) of indicators used in these areas.

Table 1 – Ecological status classification and confidence rating using HEAT tool in the open Baltic Sea sub-basins for the three quality elements included.

HEAT	Bothnian Bay	Bothnian Sea	Gulf of Finland	N Baltic Proper	E Gotland Basin	Bornholm Sea	Arkona Basin
Ecological status classification							
Plankton	Good	Poor	Bad	Bad			
Inv. benthic fauna	High	High	Bad	Bad	Bad	Bad	Good
Phys.-chem. features	Good	Good	Poor	Poor	Moderate	Good	High
Final status	Good	Poor	Bad	Bad	Bad	Bad	Good
Confidence rating (%)							
Plankton	50	50	50	50			
Inv. benthic fauna	63	63	63	63	50	38	38
Phys.-chem. features	90	89	89	78	83	63	50
Final confidence class	II	II	II	II	II	II	III

Table 2 – Ecological status classification and confidence rating using OSPAR COMP tool in the open Baltic Sea sub-basins for the three quality elements included.

OSPAR COMP	Bothnian Bay	Bothnian Sea	Gulf of Finland	N Baltic Proper	E Gotland Basin	Bornholm Sea	Arkona Basin
Eutrophication status classification							
Causative Factors	Moderate	Good	Poor	Poor	Poor		
Direct effects	Good	Poor	Bad	Bad			
Indirect Effects	High	High	Bad	Bad	Bad	Bad	Good
Other possible effects		Poor	Bad	Poor	Good	Good	High
Final Status	Moderate	Poor	Bad	Bad	Bad	Bad	Good
Confidence rating (%)							
Causative Factors	87	83	83	67	83		
Direct Effects	50	50	50	50			
Indirect Effects	63	63	63	63	50	38	38
Other possible effects		75	75	75	63	63	50
Final confidence class	II	II	II	II	II	II	III

A1.5 Conclusions

Confidence rating of eutrophication indicators presented in this study shows an example of how various indicator features can be used in confidence rating. The indicator confidence rating is a simple and effective tool for eutrophication assessments, that in addition to making the assessment more reliable and transparent also pinpoints the possible development challenges of particular indicators. The study shows that concrete, relatively simple and robust indicator features can be used in confidence rating in a coherent manner.

A1.6 References

- HELCOM (2006): Development of tools for assessment of eutrophication in the Baltic Sea. *Balt. Sea Environ. Proc.* No. 104.
- Olsonen, R. (ed.) (2008): FIMR monitoring of the Baltic Sea environment – Annual report 2007. – Meri – Report Series of the Finnish Institute of Marine Research 62 (pp. 81–88) Helsinki, Finland.
- Fleming-Lehtinen, V. (ed.) (2007): HELCOM EUTRO: Development of tools for a thematic eutrophication assessment for two Baltic Sea sub-regions, The Gulf of Finland and The Bothnian Bay. MERI No. 61, Report Series of the Finnish Institute of Marine Research, Helsinki, Finland.
- Fleming-Lehtinen, V., M. Laamanen, H. Kuosa, H. Haahti & R. Olsonen (2008): Long-term development of inorganic nutrients and chlorophyll a in the open Northern Baltic Sea. *Ambio* 37:86–92.
- Schernewski, G., & T. Neumann (2005): The trophic state of the Baltic Sea a century ago: a model simulation study. *Journal of Marine Systems* 53:109–124.
- Pearson, T.H., & R. Rosenberg (1978): Macroben- thic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology: an Annual Review* 16: 229–311.
- Josefson, A., & D. J. Conley (1997): Benthic responses to a pelagic front. *Marine Ecology Progress Series* 147:49–62.
- Gray, J.S. (2000): The measurement of marine species diversity, with an application to the benthic fauna of the Norwegian continental shelf. *Journal of Experimental Marine Biology and Ecology* 250:23–49.
- Sandén, P., & B. Håkansson (1996): Long term trends in Secchi depth in the Baltic Sea. *Limnology and Oceanography* 41(2):346– 351.
- HELCOM (2009a): Eutrophication in the Baltic Sea – An integrated thematic assessment of the effects of nutrient enrichment in the Baltic Sea region. *Balt. Sea Environ. Proc.* No. 115B. 148 pp.
- HELCOM (2009b): Eutrophication in the Baltic Sea – An integrated thematic assessment of the effects of nutrient enrichment in the Baltic Sea region. Background Report. Available online. *Balt. Sea Environ. Proc.* No 115C.

Appendix 2:

Norwegian case study

Jarle Molvær & Karl Norling, NIVA, Norway

A2.1 Background

The objectives of CONFIRM are to develop a methodology for confidence rating of marine eutrophication assessments, including guidance and an assessment tool in the form of a downloadable open-source program. In addition to the methodology and programs for confidence rating, the tools include possibilities of weighting between indicators (per category/quality element) and between categories and quality elements.

As agreed at the September 2008 meeting at DHI, we have applied CONFIRM in the Oslofjord, along a gradient from inner Oslofjord to the border between outer Oslofjord and Skagerrak waters. The data are described below and the results are shown in the attached HEAT-file.

A2.2 Data

The data is mainly based on four projects:

- The monitoring program for inner Oslofjord (latest report: Magnusson et al., 2008)
- The monitoring program for outer Oslofjord during 2001–2005 (a series of reports, but especially the summary report from Dragsund et al., 2006).
- The monitoring program for outer Oslofjord during 2007 (Walday et al., 2008).
- The monitoring program for inner Oslofjord (Olsgard 1993)

These reports include information about anthropogenic nutrient loads and effects in the recipients.

The classification was performed for 8 stations along a transect from the southern part of outer Oslofjord (station OF1) to the Bunnefjord in inner Oslofjord (station EP1, Figure 1). A description of available data is found in Tables 1–5 and “A2.5 Supplementary information: monitoring stations”.

Figure 1. The Oslofjord showing the 8 stations (areas) included in the classification.

Table 1 – Description of stations and data

Station	Depth	Plankton	NEAGIG	NQI1	H'	HB	BHQ	totN	totP	SD
EP1	153	X		X	X	X	X	X	X	X
CQ1	88	X		X	X	X	X	X	X	X
DK1	100	X		X	X	X	X	X	X	X
IM2	203	X		X	X	X	X	X	X	X
OF5	198	X	X	X	X	X	X	X	X	X
OF4	305	X		X	X	X	X	X	X	X
OF2	356	X	X	X	X	X	X	X	X	X
OF1	360–460	X	X	X	X	X	X	X	X	X

(see also Figure 1 and Figures 2,3,4,5 in A2.5 Supplementary information monitoring stations). Stations in inner Oslofjord are shaded grey.

A2.3 Parameters

A2.3.1 Phytoplankton

90th percentiles of chlorophyll a are calculated for every station. According to agreements under the WFD in the NEA-GIG area, the 90th percentile values for chlorophyll a are used for classification of water quality. To fulfill the requirements for classification of water quality by using chlorophyll a, data must be collected throughout the whole phytoplankton growing season.

Data from the whole phytoplankton growing season (February–October) from 2002–2007 (six years) is used in the calculations for stations EP1, DK1 and IM2 and classification according to Table 2.

For the stations OF1, OF2, OF4 and OF5, data for the period 2001–2005 (five years) were used for the calculations. However, data from these stations were sampled in June to September. Therefore, the data basis for classifications for these stations in outer Oslofjord is not according to the requirements.

Table 2 – Classification table for phytoplankton measured as chlorophyll *a*. Units are µg/l

Watertype	Salinity	Reference	Ecological status classification according to parameter				
		condition	High	Good	Moderate	Poor	Bad
Exposed	≥20	2,3	<3,5	3,5–<7	7–<11	11–<20	>20
Moderate exposed	≥30	2,0	<3	3–<6	6–<9	9–<18	>18
Sheltered	≥30	2,0	<3	3–<6	6–<9	9–<18	>18
Freshwater influence	18–<30	2,5	<3,3	3,3–<5	5–<7,5	7,5–14	>14

Submerged aquatic vegetation

The hardbottom flora has been classified for the outer Oslofjord, using the lower depth for growth in the water types Sk1, Sk 2 and Sk 3 (NeaGIG 8, 9, 10). Data were collected as part of the Norwegian Coastal Monitoring Programme and the Monitoring Programme for Outer Oslofjord during 2007. Note that while station OF2 (Figure 1) lies in open waters, the associated station for macroinvertebrates south of Tønsberg is situated in more enclosed waters.

There is no recent data from the inner Oslofjord.

Invertebrate benthic fauna

Macroinvertebrate data from inner and outer Oslofjord have been sampled in two major monitoring programmes, see Figure 1 and Figure 5.

Table 3 – Reference conditions used for classification of macroinvertebrate samples.

Indicative parameter	Reference condition	Ecological status classification according to parameter				
		High	Good	Moderate	Poor	Bad
NQI1	0,78	>0,72	0,63–0,72	0,49–0,63	0,31–0,49	<0,31
NQI2	0,73	>0,65	0,54–0,65	0,38–0,54	0,20–0,38	<0,20
H'	4,4	>3,8	3,0–3,8	1,9–3,0	0,9–1,9	<0,9
ES100	32	>25	17–25	10–17	5–10	<5
ISI	9	>8,4	7,5–8,4	6,1–7,5	4,2–6,1	<4,2
AMBI		<1,2	1,2–3,3	3,3–4,3	4,3–5,5	>5,5

The classification used in CONFIRM-Oslofjorden is primarily NQI1 which is an inter-calibrated index within Northeast Atlantic Geographical Inter-calibration Group (NEAGIG) benthic expert group, Table 3 shows the classification of NQI1 in comparison to several other indices used in Norway and the EU.

Hyperbenthos is averaged data from 2001–2006 because of single hauls each year. Reference conditions are due to documented prawn fisheries in Bunnfjorden more than 100 years ago (Wollebæk, 1906). Already in 1914, Petersens investigations in Bunnefjorden showed very low numbers of macroinvertebrates (Petersen, 1915).

Sediment Profile Imaging (SPI) is a recent technique that takes an image 10–25 cm down in surface sediments. Profile images were taken at the fjord stations for benthic macroinvertebrate studies in the different areas classified according to CONFIRM (Figures. 1, 3, 5

The BHQ index was averaged from analysis of four images at each station. This index parameterizes sediment surface structures, sub-surface structures, and the apparent depth of the aRPD¹. The BHQ index ranges from 0 (severely disturbed with no macrofauna) to 15 (undisturbed with mature benthic community) and is related to the successional stages of the benthic fauna according to the Pearson and Rosenberg model, see Figure 6.

Physio-chemical features

New hydrochemical classification tables are being developed, but presently the classification system from 1997 (SFT 1997) is used (Table 4).

¹ Apparent Redox Potential Discontinuity

Table 4 – Classification of nutrients, chlorophyll a, and Secchi depth for salinity >20.

		Classes				
Parameters		I Very Good	II Good	III Fair	IV Bad	V Very bad
Surface layer	Total phosphorus (µg P/l)	<12	12–16	16–29	29–60	>60
Summer	Phosphate (µg P/l)	<4	4–7	7–16	16–50	>50
(June-August)	Total nitrogen (µg N/l)	<250	250–330	330–500	500–800	>800
	Nitrate (µg N/l)	<12	12–23	23–65	65–250	>250
	Ammonium (µg N/l)	<19	19–50	50–200	200–325	>325
	Chlorophyll a (µg/l)	<2	2–3.5	3.5–7	7–20	>20
	Secchi depth (m)	>7.5	7.5–6	6–4.5	4.5–2.5	<2.5
Surface layer	Total phosphorus (µg P/l)	<21	21–25	25–42	42–60	>60
Winter	Phosphate (µg P/l)	<16	16–21	21–34	34–50	>50
(December-February)	Total nitrogen (µg N/l)	<295	295–380	380–560	560–1300	>1300
	Nitrate (µg N/l)	<90	90–125	125–225	225–350	>350
	Ammonium (µg N/l)	<33	33–75	75–155	155–325	>325

The classification for inner and outer Oslofjord from monitoring programmes is shown in Table 5 and Table 6, respectively.

According to OSPAR the classification for nutrients is based on winter data and as Tables 5–6 are based on data from the summer, winter-data were therefore calculated separately. From Tables 5–6 only Chlorophyll a and Secchi depth were used.

As Reference condition the values in Class I were used, and with the transition from Class II to Class III as “Acceptable Deviation”.

Table 5– Classification of water quality in inner Oslofjord during June – August 2007 (from Magnusson et al., 2008)

Station	Salinity	Tot-P	Tot-N	Chl-a	Secchi Depth
	(PSU)	(µg/l)	(µg/l)	(µg/l)	(meter)
Ap1					3.70
Ap2	19.00	12.80	245.00	4.60	3.53
Aq1					3.14
Aq2	18.30				2.82
Aq3**	16.10	14.00	260.00	4.80	2.10
Bk1**					2.43
Bk2**					3.03
Bk4**	14.40	14.20	317.00	4.55	2.95
Bn1	18.70	11.30	246.00	4.10	4.00
Br1	20.00				3.87
Cq1	19.80	11.40	240.00	3.74	3.78
Cq1-dep*	19.70				3.97
Dk1	19.60	12.40	240.00	3.53	4.29
Ej1					4.08
Ep1	19.90	11.20	232.00	3.04	4.42
Gp1**	17.00				3.20

Sample taken ca. 50m from the barge at the disposal site in Bekkelags basin

** The station is classified by salinity 0–20

Table 6 – Classification of water quality in outer Oslofjord during 2007 (from Walday et al. 2008 and Dragsund et al. 2006).

Station	Nitrate	Phosphate	Tot-P	Tot-N	Chl-a	Oxygen	Secchi Depth	Season
	(µg/l)	(µg/l)	(µg/l)	(µg/l)	(µg/l)	(ml/l)	(meter)*	
OF-1	12	1	12	215	3.11	5.02	4	Summer
	149	18	-	-	-	-	-	Winter
OF-2	11	1	12	235	3.36	4.31	3	Summer
	190	16	-	-	-	-	-	Winter
OF-4	24	2	11	255	3.04	5.33	5	Summer
	155	15	-	-	-	-	-	Winter
OF-5	11	3.7	11	222	2.43	3.67	3	Summer
	180	17	-	-	-	-	-	Winter
OF-7	23	3.3	12	254	4.40	5.22	5	Summer
	188	17	-	-	-	-	-	Winter
BC-1	195	3.7	11	385	4.16	0.21	3	Summer
	269	7	-	-	-	-	-	Winter
FG-1	67	3.6	12	388	4.16	4.27	3	Summer
	180	14	-	-	-	-	-	Winter
GI-1	16	2.7	13	307	2.40	3.30	4	Summer
	146	17	-	-	-	-	-	Winter

The use of Secchi depth will be very dependent on the lighting conditions that day, including time of day for sampling.

Salinities below 20 are considered

The colour indicates environmental class: I – Very good, II – Good, III – Fair, IV – Bad, and V – Very Bad

A2.4 Results from the classification according to HEAT and OSPAR

The classification of the eight stations is shown in Table 7 and in the attached HEAT- and OSPAR-files. The results from this test may be summarized as follows:

Overall conclusion:

- CONFIRM: very promising and useful. It improves the Classification, makes it more transparent and the Confidence clearly shows whether one may rely on it, thereby reducing the more obscure expert judgements.
- The classifications fits well with the overall conclusions from the monitoring studies
- HEAT and OSPAR give very similar results. The differences are commented on below.

However, not all problems are solved and we mention some:

- Reference conditions are often difficult to establish – with implications for the EQR.
- Scores are based on individual estimates, and there may be a tendency to choose the “safe” no 2.
- Weightings are also based on individual estimates where not defined by intercalibration.

- Results may partly depend on the number of indicators, and in the “one-out-all-out” context this may determine the classification. A category where only one indicator is used may have a strong influence on the final classification. In Table 6 the difference in classification for OF4 may be explained by where the Chlorophyll and the Secchi depth were placed in the HEAT- and the OSPAR classification respectively.

More written documentation/guidelines and more experience may largely eliminate these problems.

Table 7 – Comparison of HEAT and OSPAR eutrophication status and confidence rating

Station	Eutrophication status		Confidence rating	
	HEAT	OSPAR	HEAT	OSPAR
OF1	Poor	Moderate	III	III
OF2	Moderate	Moderate	III	III
OF4	Poor	Moderate	III	III
OF5	Poor	Poor	III	III
IM2	Moderate	Moderate	II	III
DK1	Moderate	Moderate	II	III
CQ1	Poor	Poor	III	III
EP1	Bad	Bad	II	III

A2.5 Supplementary information monitoring stations

Figure 2. Hydrographic stations in the inner Oslofjord (from Magnusson et al., 2008)

Figure .5 Bottom fauna stations in the outer Oslofjord (from Walday et al., 2008).

Figure 6. Sediment profile images (colours enhanced) are shown on the top where brownish colour indicate oxidised conditions and black reduced conditions, and the benthic habitat quality (BHQ) indices (Nilsson and Rosenberg, 1997) are presented for depths >20 m. The benthic quality indices (BQIs) for the different environmental status according to the Water Framework Directive are presented for depths >20 m and 620 m at the bottom of the figure

A2.6 Supplementary information on nutrient loads

Figure 7. Calculated annual loads of phosphorous and nitrogen to Inner Oslo Fjord 1910–2002 (Bergstol et al, 1981, Balsrud et al, 1986, Holtan, 1990, Nedland, 1997, Wivestad, 1999 and Fagrådet's Annual Report 2001. The figure also includes load via natural runoff. Calculation of loads for 2004 (Bjerkeng, personal communication)

Figure 8. Calculated anthropogenic loads of phosphorous and nitrogen 1990–2006 compared with base year 1985 using TEOTIL (Selvik et al, 2007)

Figure 9. Nitrogen loads by source type for the catchment areas around the Outer Oslo Fjord. The legend shows the height of a column corresponding to 2900 tons nitrogen spread over 4 sources (green: aquaculture, beige: agriculture, brown: population; white: industry)

Figure 10. Phosphorous loads by source type for the catchment areas around the Outer Oslo Fjord. The legend shows the height of a column corresponding to 51 tons phosphorous spread over 4 sources (green: aquaculture, beige: agriculture, brown: population; white: industry)

A2.7 Literature

- Dragsund, E., Aspholm, O., Tangen, K., Bakke, S.M., Heier, L., Jensen, T., 2006. Overvåking av eutrofitilstanden i ytre Oslofjord. Femårsrapport 2001–2005. DNV rapportnr. 2006-0831. 127 sider.
- Magnusson, J., Andersen, T., Amundsen, R., Berge, J.A., Beylich, B., Bjerkeng, B., Bjerknes, E., Gjøseter, J., Grung, M., Holth, T.F., Hylland, K., Johnsen, T., Lømsland, E., Paulsen, Ø., Rønning, I., Sørensen, K., Schøyen, M., Walday, M., 2008. Overvåking av forurensningssituasjonen i indre Oslofjord 2007. NIVA-rapport nr. 5637-2008. 116 sider.
- Nilsson, H.C., Rosenberg, R., 1997. Benthic habitat quality assessment of an oxygen stressed fjord by surface and sediment profile images. *Journal of Marine Systems* 11, 249–264.
- Petersen CGJ 1915. On the animal communities of the sea bottom in the Skagerrak, the Christiania Fjord and the Danish waters. Rep Danish Biological Station 23:1–28.
- Olsgard, F. 1993. Overvåking av forurensningssituasjonen i indre Oslofjord. Undersøkelser av bløtbunnsfauna 1993. Overvåkningsrapport 622/95) TA nr. 1258/1995. 106 sider.
- SFT, 1997. Klassifisering av miljøkvalitet i fjorder og kystfarvann. Statens forurensningstilsyn. Veiledning 97:03. 36 sider. (In Norwegian).
- Walday, M., Gitmark, J., Naustvoll, L., Nilsson, H.C., Pedersen, A., Selvik, J.R., 2008. Overvåking av ytre Oslofjord 2007. Årsrapport. NIVA-rapport nr. 5640-2008. 63 sider.
- Wollebæk, 1906. Nogle statistiske oplysninger om fisket efter dybvandsræger (*Pandalus borealis*) i Kristianiafjorden indenfor Drøbak. www.vannportalen.no

Appendix 3: Swedish case study

Philip Axe, SMHI, Sweden

A3.1 Introduction

Monitoring data covering nutrient concentrations, Secchi depth and phytoplankton characteristics have been used in several ecosystem assessments on the Swedish West Coast. In the Swedish national eutrophication report to OSPAR (Håkansson et al, 2007) the offshore Skagerrak was classified as a Non-Problem area, while the coastal Skagerrak (within 1 nautical mile of the baseline) was classified as a Problem Area. This appendix discusses this result and compares it to those obtained using the methodology adopted in Sweden for evaluation of water bodies under the Water Framework Directive, and also the methodology associated with the HELCOM Eutrophication Assessment Tool (HEAT) used recently in the classification of the eutrophication status of the Baltic (HELCOM, 2009).

A3.2 Study area: Gullmarfjord

Gullmarfjord, on the Swedish west coast, is Sweden's only true fjord, in the sense that it is a deep inlet formed through glaciation, where the deep inner waters are separated from the outer sea by a shallow sill. Depth at the sill is about 60 metres, while depths in the inner parts vary between 115 metres at Alsbäck, to 70 metres at Björkholmen. The fjord receives fresh water from a stream and a river (indicated by the blue arrow in Figure 1). Monthly measurements of nutrient concentrations in the river, combined with daily measurements of discharge collated by the Environmental Monitoring and Assessment Group of the Department of Aquatic Science and Assessment at the Swedish Agricultural University (Miljöanalys) indicate that approximately 700 tonnes of total nitrogen and 35 tonnes of phosphorus are delivered to the fjord each year. These figures were previously higher, before the introduction of a water treatment works on Örekilsälven. While approximately 8000 people live in the town of Lysekil, at the mouth of the fjord, discharges from the town's water treatment works are released from the north side of the town, and are not considered to affect the fjord.

Earlier industrial use of the fjord resulted in areas being affected by heavy metal and organic pollutants, while eutrophication became a particular problem in the 1970s and 1980s. The sill at the fjord's mouth restricts the renewal of deep water and this, combined with eutrophication pressures, has

resulted in autumn hypoxia between 70 and 120 metres since about 1980. (Figure 2). High nutrient concentrations still occur in the surface waters, although these are generally in conjunction with heavy rainfall.

Offshore of the fjord is the meeting point for the waters of the Baltic Outflow and the Jutland Coastal Current, as they merge to form the Baltic Current. The Baltic outflow consists of relatively fresh water which flows out from the Baltic through the Sound and Great Belts before continuing northwards along the Swedish west coast on the surface of the Kattegat. This water generally has lower nitrogen and phosphorus concentrations than the saltier Kattegat water below it, although river discharges, particularly from the Göta Älv increase nutrient loading.

Figure 1. Study area, showing the Gullmar fjord, sampling stations and large scale hydrographic features

The Jutland Coastal current flows up the Danish west coast before separating from it at Skagen (the northern tip of Denmark). This water is saline, originating in the southern North Sea. It often has high nitrite concentrations, due to the outflows from the Rhine and Elbe. Water typical of the Jutland Current is frequently observed at Släggö, where it is thought to have an important influence on primary production (L. Andersson, pers. comm.).

Figure 2. Time series of oxygen concentration with depth in central Gullmarfjord, showing periods of hypoxia (from Håkansson, 2007)

As the Jutland Current and Baltic Outflow combine to form the Baltic Current, they form a dynamic front, parallel to the Swedish coast. The position of this front, and the depth of the Baltic Current, are monitored monthly as part of the Swedish national pelagic monitoring programme. A transect of five stations are placed across the front. At three stations (Å13, Å15 and Å17) Secchi depth, nutrient and chlorophyll *a* concentrations and plankton are measured throughout the year. High resolution CTD profiles (measuring temperature, salinity, oxygen and chlorophyll *a* fluorescence) are made at these stations, at Å14, and Å16 (marked in grey on Figure 1) and at Släggö. The offshore station (Å17) is considered typical of the central Skagerrak water, the coastal stations (Å13, and also Släggö) are considered typical of the coastal waters, while the stations Å14 – Å16 allow the position of the front to be inferred.

Because of the fjord's unique status in Sweden, and the presence of several marine research stations along its coastline, it has been well monitored. Observations from Släggö extend at least as far back as 1902. Present day monitoring is commissioned by the Bohus Coast Water Quality Association (monitoring at Släggö and Björkholmen) and by the Gullmarfjord Water Quality Association (bottom water sampling at Alsback). Monitoring is similar to that at in the national monitoring programme. Additional monitoring of benthic fauna occurs at other locations in the same water bodies. The combined monitoring programmes mean that the area is covered by monthly sampling at all stations, while Släggö is visited twice per month.

A3.3 Classification methodologies

OSPAR Common Procedure

The OSPAR Common Procedure structures the classification of the eutrophication status of a water body using four categories (Table 1). The first considers the nutrient loading and concentrations. The second category examines the direct effects of nutrient enrichment, such as increased phytoplankton growth and reduced Secchi depth. The third category concerns indirect effects, such as hypoxic events or reductions in benthic macrofauna, while the fourth is based on events considered to be related to eutrophication, such as harmful algal blooms and the occurrence of algal toxins. The evaluation usually covers a five year period, and is applied across a whole region (such as the offshore Skagerrak) and not station by station.

For each region where the Common Procedure is applied, OSPAR have defined background and elevated nutrient concentrations, and suitable ranges for nutrient concentration ratios, to be used as the metrics in assessing whether a region suffers eutrophication. To eliminate the influence of fresh water, mixing diagrams are used to correct observed concentrations to a nominal salinity of 30 psu, before evaluation. It is not precisely defined however whether the threshold concentrations are to be compared to a mean, median or maximum observed concentration. In the Swedish report (Håkansson, 2007) the annual median and 90th percentile nutrient concentrations were used.

Table 1 – Parameters used to describe the four categories of eutrophication indicators in the Swedish national eutrophication report to OSPAR (Håkansson et al, 2007)

Category I	Category II	Category III	Category IV
Nutrient enrichment	Direct effects	Indirect effects	Other possible effects of eutrophication / Voluntary parameters
Nutrient budgets: discharges from land, atmospheric deposition and transboundary inputs	Maximum and minimum chlorophyll a concentrations (February to November)	Oxygen deficiency: Autumn 5 th percentile and median of bottom oxygen concentrations and saturations, as trends and levels	Algal toxins
Winter nutrient concentrations (means) corrected to 30 psu	Secchi depth	Benthic macrofauna from 15 stations, 4 depth strata, using a BQI approach based on tolerance, abundance and diversity	Primary production in Gullmar fjord
Winter nutrient ratios	Phytoplankton indicator species		

Chlorophyll a was compared to the OSPAR Problem (elevated) level of 2 µg/l. No salinity correction was applied. It was also unclear as to which metric should be used, and so once again the annual median and 90th percentiles were used. Secchi depth was not assessed on this occasion. Of the third assessment category variables, oxygen concentrations and saturations were reported, as were trends. Macrobenthos quality was presented and discussed. Of the Category Four variables, the frequency and occurrence of harmful

algal blooms was reported, but as with the Category Three variables, it was not possible to relate the reported signals directly to eutrophication.

Water Framework Directive

Annex V of the Water Framework Directive (WFD) defines Quality Elements for the status assessment of transitional and coastal waters primarily in terms of biological elements (phytoplankton, aquatic flora, benthic invertebrate fauna and fish) supported by hydromorphological elements and physico-chemical elements (see Table 2). Status classification is made according to five classes, ranging from High to Bad. High status is achieved when where there is at most only very minor anthropogenic influence on the hydromorphological and physico-chemical characteristics of the water body and the biological elements are almost indistinguishable from those expected in pristine conditions. Bad status is characterised by severe alterations to biological quality elements. High and Good status are considered acceptable. Moderate, Poor or Bad status are not. In order to make these judgements, Reference conditions, describing for example nutrient concentrations or phytoplankton biomass under pristine conditions, have been defined, based on expert assessment, historical data and numerical models. Acceptable deviations from reference conditions have also been estimated, to define the position of the Good/Moderate status boundary as well as the boundaries between the other status classes.

Table 2 – Biological elements and their supporting elements for Transitional and Coastal waters, from Annex V of the Water Framework Directive.

Biological elements Composition, abundance and biomass of phytoplankton	Hydromorphological elements supporting the biological elements		Chemical and physico-chemical elements supporting the biological elements	
	Morphological conditions	Tidal regime	General	Specific pollutants
Composition and abundance of other aquatic flora	Depth variation	Freshwater flow*	Transparency Thermal conditions	Pollution by all priority substances identified as being discharged into the body of water
Composition and abundance of benthic invertebrate fauna	Structure and substrate of the transitional/coastal bed	Direction of dominant currents**	Oxygenation conditions	Pollution by other substances identified as being discharged in significant quantities into the body of water
Composition and abundance of fish fauna*	Structure of the intertidal zone	Wave exposure	Salinity Nutrient conditions	

Elements relevant to transitional waters only are indicated by *. Those applicable to coastal waters only are indicated by **

Reference conditions were determined by examining the relation between parameter concentrations and total nitrogen. A clear relation between Secchi depth and total nitrogen concentration exists in Swedish waters (Larsson et

al, 2006), and historical values of Secchi depth are available from the 1930s (collated by Aarup, 2002). Based on this relationship, the historical (1930) total nitrogen concentration was estimated to be $14.6 \mu\text{mol/l}$, which was in reasonable agreement with the salinity-affected range of $14\text{--}17 \mu\text{mol/l}$ proposed independently by Hansson & Håkansson (2005).

Reference conditions for chlorophyll a were determined by studying the relationship between chlorophyll a and Secchi depth, and between chlorophyll a and total nitrogen. This indicated that reference conditions should be between 0.9 and $1.5 \mu\text{g/l}$, dependent upon which relationship was used. The final accepted reference condition was $1.10 \mu\text{g/l}$ offshore, increasing to $1.90 \mu\text{g/l}$ inshore at Björkholmen. The position of the other class boundaries was determined on the basis of the spread of data, and tested within the NEA GIG project (European Commission, 2008).

The Swedish implementation of the Water Framework Directive has some differences to methods used in neighbouring countries: Because coastal and transitional waters are strongly influenced by fresh water run off, Sweden has implemented a system whereby nutrient reference conditions (and class boundaries) are functions of salinity. At low salinities, the reference conditions are very similar to those adopted for fresh water. At an upper salinity limit, reference conditions are set to an “offshore” level. This permits classification under natural variability (if the water body has been affected by extreme runoff, for example). Figure 3 shows how the status classes for total nitrogen change with salinity on the Swedish west coast, as well as indicating the status classes of observed data (from 2001). A simpler version of this approach is used for the phytoplankton parameters biovolume, chlorophyll a concentration and community structure in some parts of Sweden’s Baltic coast (Naturvårdsverket, 2007).

Instead of using seasonal or annual mean biovolume, nutrient or chlorophyll a concentration, and calculating a mean ecological quality ratio (EQR) based on that value and the reference condition, the Swedish implementation of the WFD calculates the EQR for each individual observation (using the salinity specific reference condition and not allowing an EQR value of greater than 1). The EQR for that parameter is then the mean of all the individual EQR values. This mean EQR is then used to classify the status for that parameter, using a numerical scale of 0 to 5, where 0 is the bottom of the “Bad” status, and 5 implies that all the EQR values equalled 1. By combining parameters within a quality element using their status scores, rather than the EQR, it is possible to include parameters such as oxygen where an EQR approach has not yet been developed.

Due to problems in identifying the maximum winter nutrient concentrations on the Swedish West Coast, where the spring bloom can occur in February, or April, a method was adopted which uses the individual nutrient observations that go to make up the highest mean surface water concentration (Hansson and Håkansson, 2005). This means that the data used to evaluate DIN may have come from a different sampling occasion than that used for the DIP assessment,

although as the method uses the “worst case” for each, it should give a more conservative assessment than using a simple winter mean.

When there is more than one indicator within a quality element – for example both biovolume and chlorophyll a describing phytoplankton status – then final status for that quality element is the mean of all the individual indicator statuses. This approach allows the uncertainty of a quality element’s status to be described in terms of class boundaries, and avoids problems of averaging class boundaries when the borderline EQR values may vary. The final status is determined as specified in the WFD, using the “one out, all out principle”.

In this study, the following parameters were used:

Biological quality

- Phytoplankton biovolume
- Chlorophyll a concentration

Chemical & physico-chemical elements supporting the biological quality elements

- Secchi depth
- Total nitrogen and phosphorus concentrations during winter
- Total nitrogen and phosphorus concentrations during summer
- Winter concentrations of dissolved inorganic nitrogen and phosphorus (winter DIN and DIP).

Figure 3. Variation of reference conditions and class boundaries for total nitrogen with salinity, for “Type 2” water bodies (Swedish West Coast fjords). From Naturvårdsverket, 2007. Class names are given in Swedish, and refer to the English classes “High”, “Good”, “Moderate”, “Poor” and “Bad”

HEAT methodology

In using HEAT, the same biological and physico-chemical parameters were used as for the example using the WFD methodology. The difference is that mean concentrations and biovolume were used in HEAT, rather than individual values used in the Swedish WFD implementation. This was to allow the indicator weighting system in HEAT to be used. Where appropriate, the salinity correction to the reference conditions and acceptable deviations was made using the mean salinity.

HEAT also differs from the Swedish implementation of the WFD as the class boundaries are not explicitly defined: the user gives only the acceptable deviation, and then boundaries are calculated by some function. The position of the boundaries and the width of classes may differ from those in the Swedish WFD implementation.

HEAT allows the user to input some scientific judgement to the assessment process by applying confidence scores to the reference conditions, acceptable deviations and monitoring data, as well as by weighting the various indicators within a quality element. The overall EQR of a quality element is then calculated using this weighting. This differs from the Swedish WFD implementation, where each parameter EQR generates a status score, and the overall status is the mean of these statuses.

HEAT uses the confidence ratings to make an overall confidence score for each parameter, and then combines them to produce the overall quality element confidence score. The final confidence score is based on the combined quality element confidence scores, and is classified as Class I, II or III, where I is the highest confidence rating, and III is the least safe.

Applying HEAT to the Swedish West Coast, nutrient reference conditions and acceptable deviations were calculated based on the mean salinity at each station during the 2001 – 2006 assessment period. On the Swedish West Coast the other parameters are not considered to have a strong salinity dependency within their water type. Mean values of each parameter were then tested against their respective reference conditions.

Because the WFD reference conditions were derived on the basis of historical data, and had been validated through the NEA GIG process they were considered to be of good quality, and given a confidence rating of 1, with the exception of biovolume, which was rated as 2, as less data exists for validation. Acceptable deviations varied, from 94% in biovolume at Släggö, to 26% for summer total nitrogen. Where acceptable deviations were small, they were considered to be well defined, and given higher quality ratings than those that permitted large variation in the tested parameter (such as biovolume). Because the area is sampled so regularly, the quality of the measurements was considered to be very good, and scored “1”, with the exception of Secchi depth and biovolume at the most offshore station. Because this station is usually visited after Släggö, Å13 and Å15, it is often dark before the research vessel is on station, so Secchi observations are less

common. Biovolume analyses from this station have only started recently, so were also given a confidence rating of “2”.

A3.4 Results of the assessments

OSPAR

Median winter DIN concentrations were close to (inshore) or below (offshore) the OSPAR background concentration of 10 µMol/l, and with the exception of data from 2001, the 90th percentile concentrations below the OSPAR elevated levels. DIP levels were also below the background (median concentrations) and elevated levels (90th percentile). Nutrient loading to the coastal zone remains high however.

Chlorophyll a concentrations were high in the coastal areas, and some fjords with poor water exchange exhibited oxygen deficiency, which impacted on the benthic fauna. A study of offshore shallow water environments showed in addition that offshore benthic communities were in much better condition than coastal ones.

The result of the assessment (Table 3). indicated that the inshore Skagerrak was a eutrophication problem area, while the offshore Skagerrak was a non-problem area.

Table 3 – Results of the 2007 Swedish Implementation of the OSPAR Common Procedure for the Skagerrak.

		Inshore Skagerrak	Offshore Skagerrak
Category I	Nutrient inputs	+	-
Degree of	Winter DIN and DIP	-	-
nutrient enrich-	Winter N:P ratio	-	-
ment			
Category II:	Chlorophyll a concentration	+	-
Direct effects of	Phytoplankton indicator species	+	-
eutrophication	Macrophytes	?	
Category III:	Oxygen deficiency	+	-
Indirect effects	Changes/kills in zoobenthos or fish	+	-
of eutrophication	Organic carbon matter	+	
Category IV:	Algal toxins	?	
Other possible	Primary production	-	
effects of eutro-			
phication			
Overall		Problem Area	Non-problem area
classification			

Red indicates problem; green non-problem and yellow uncertain.

WFD

The Swedish regional water authorities make the results of their assessments of environmental status available on the “Vattenkartan” web site (<http://gis.lst.se/vattenkartan/>). Results of the assessments for Secchi depth,

phytoplankton, bottom fauna and oxygen are shown in Figure 4, indicating that the overall status in the study area is governed by bottom fauna.

Table 4 shows the EQR values calculated for phytoplankton and physical-chemical supporting parameter data from the six individual stations used in the HEAT calculations. As in the OSPAR classification, status improves offshore, with high status being attained in the central Skagerrak. Table 5 shows the results summarized by Quality Element, and an overall classification based on the “one out, all out principle”. The table also shows an estimate of the uncertainty in the assessment for each quality element: the figure in brackets is the standard deviation of the status scores for all the parameters within that QE. Where the standard deviation is more than 0.5, it suggests that the actual classification class could be uncertain. In this case, at Björkholmen, the standard deviation was greater than 1, so the “Good” classification should be treated as uncertain. At Å17, the variability in the status scores is 0.42, and so this “High” status classification can be accepted with confidence.

i) Secchi depth classification

ii) Phytoplankton classification

iii) Benthic invertebrate fauna classification iv) Oxygen classification

Figure 4. Classification of Gullmar fjord (and adjacent waters) for i) Secchi depth; ii) Phytoplankton; iii) Benthic Invertebrate Fauna and iv) Oxygen. Blue indicates High status, green is Good, yellow is Moderate and orange is Poor. No water body is classified here as Bad.

Table 4 – Results of the station classification, using the methodology from the Swedish implementation of the Water Framework Directive.

	Björkholmen	Alsback	Släggö	Å13	Å15	Å17
Phytoplankton	Biovolume		0.833 4.49			0.986 4.96
	chlorophyll a concentration		0.461 2.51			0.927 4.65
Physico-chemical supporting parameters	Secchi depth	0.691 3.34	0.747 3.65	0.682 2.96	0.819 3.93	0.708 3.24
	Winter total nitrogen concentration	0.853 3.70		0.814 3.27	0.869 3.88	0.985 4.88
	Winter total phosphorus concentration	0.848 3.98		0.817 3.70	0.896 4.31	0.939 4.59
	Winter DIN concentration	0.191 0.68		0.564 2.56	0.578 2.63	0.841 4.21
	Winter DIP concentration	0.805 4.02		0.802 4.01	0.809 4.04	0.921 4.61
	Summer TN concentration	0.757 2.82		0.776 2.93	0.795 3.06	0.870 3.89
	Summer TP concentration	0.759 3.41		0.875 4.26	0.919 4.52	0.949 4.70
	Explanation of status scores	0 – 1: Bad	1 – 2: Poor	2 – 3: Moderate	3 – 4: Good	4 – 5: High
						(insufficient data)

Results are presented as status colours, as EQR values (upper values in each cell) and as status values (lower italicized values in each cell)

Table 5 – Summary of the results in Table 4, showing the mean status score for each Quality Element, as well as the standard deviation of that score, in terms of class boundaries

	Björkholmen	Alsback	Släggö	Å13	Å15	Å17
Status score & standard deviation for phytoplankton			3.50 (1.41)			4.80 (0.22)
Status score & standard deviation for physico-chemical parameters	3.14 (1.16)	3.65 (N.A.)	3.38 (0.63)	3.77 (0.68)	4.30 (0.57)	4.55 (0.42)
Overall classification	Good	Good	Good	Good	High	High
Explanation of status scores	0 – 1: Bad	1 – 2: Poor	1 – 2: Poor	2 – 3: Moderate	3 – 4: Good	4 – 5: High

HEAT assessment

Results from the HEAT assessment are presented in Table 6, presenting the EQR for each component, as well as the confidence assessment), and summarised in Table 7.

Table 6 – Results of the station classification, using HEAT, with reference conditions and Acceptable deviation from the Swedish implementation of the Water Framework Directive, albeit calculated for mean conditions.

	Björkholmen	Alsbäck	Släggö	Å13	Å15	Å17
Phytoplankton	Biovolume		0.36			1.00
			50%			33%
Physico-chemical supporting parameters	chlorophyll a concentration	0.60	0.57	0.60	1.00	1.00
		67%	67%	67%	67%	67%
	Secchi depth		0.69	0.76	0.73	0.62
			83%	83%	83%	67%
	Winter total nitrogen concentration	0.75	0.81	0.87	1.00	1.00
		100%	100%	100%	100%	100%
	Winter total phosphorus concentration	0.88	0.88	0.92	1.00	1.00
		100%	83%	83%	83%	83%
	Winter DIN concentration	0.58	0.52	0.53	0.85	0.82
		83%	83%	83%	83%	83%
	Winter DIP concentration	0.85	0.82	0.82	0.92	0.96
		83%	83%	83%	83%	83%
	Summer TN concentration	0.70	0.73	0.72	0.85	0.87
		100%	100%	100%	100%	100%
	Summer TP concentration	0.89	0.90	1.00	1.00	1.00
		83%	83%	83%	83%	83%

Decimal numbers are the EQR values, and percentages the HEAT-calculated confidence assessments

Table 7 – Summary of the results in Table 6, showing the mean status score for each Quality Element, as well as the standard deviation of that score, in terms of class boundaries.

	Björkholmen	Alsbäck	Släggö	Å13	Å15	Å17
EQR, status & confidence rating of phyto-plankton	0.60		0.51	0.60	1.00	1.00
	50%		62%	50%	50%	57%
EQR, status & confidence of physico-chemical parameters	0.76		0.76	0.80	0.95	0.96
	90%		88%	88%	88%	84%
Overall classification & confidence	Good		Moderate	Moderate	High	High
	70%		75%	69%	69%	70%
Status colours	Bad	Poor	Moderate	Good		High

Alsbäck was not included, as only Secchi depth data was available.

A3.5 Discussion

At first glance, each of the methods delivers a similar result, with lower status in the coastal zone, and close to pristine offshore at Å15 and Å17. Both the HEAT approach and the Swedish WFD implementation identify the boundary between the high status offshore water and the good or moderate status coastal water as lying between Å13 and Å15. The boundary between Skagerrak inshore and offshore, used in the OSPAR assessment is also close to these stations.

The OSPAR approach identifies nutrient inputs to the coastal zone to be a problem, and this may have been confirmed by the bad status found at Björkholmen for the winter DIN concentration. This bad status was not found in the HEAT assessment – probably because the HEAT assessment was based

on mean nutrient concentration and mean salinity – whereas the Swedish WFD methodology identifies the highest concentration event during a winter, and calculates the salinity-related EQRs based on the individual observations during that event, giving a more conservative (worse) EQR score.

Coastal chlorophyll *a* concentrations and phytoplankton species composition were also flagged as areas of concern in the OSPAR report. An EQR based approach for phytoplankton community structure has yet to be developed in Sweden, and so no such analysis was possible either within the Swedish WFD or HEAT analyses. Chlorophyll *a* was analyzed however: In the Swedish WFD implementation, the chlorophyll *a* assessment is carried out on integrated samples collected by hose from between 0 and 10 metres. Data are only available from Släggö and Å17. In the HEAT approach, chlorophyll *a* estimates from water bottle samples were used, allowing assessment for all stations apart from Alsbäck. HEAT classified Släggö as having moderate status based on the phytoplankton Quality Elements, with moderate confidence. The Swedish approach had high status for biovolume, and moderate status for chlorophyll *a*. The resulting classification was thus “Good”, although the standard deviation in this classification was greater than one class boundary, suggesting that the classification is uncertain. HEAT gave a confidence of 62% for this element, also indicating some uncertainty.

While the OSPAR approach produces a qualitative assessment of the eutrophication status, the lack of guidance about which metrics should be used is concerning. In 2007, Sweden used the 90th percentile nutrient concentration to describe elevated nutrient conditions, although it is not clear whether OSPAR meant “elevated status” to refer to a long term or severely eutrophied state, or to short term events as was interpreted in this case. The clear advice contained in the national guidance for the application of the Water Framework Directive does not allow this sort of freedom of interpretation and is to be preferred.

HEAT uses the individual EQR scores within a quality element to calculate the weighted average EQR, which is then used to set the status for that quality element. It is not yet clear how the Quality Element status is determined by HEAT when two indicators have class boundaries at differing EQR values. Figure 5 illustrates this problem, where two indicators in a Quality Element have the same EQR, but because of differences in acceptable deviation, have different statuses. What is the final status?

Figure 5. Example where both indicators in a quality element have an EQR of 0.47, but the status of indicator 1 is moderate, and indicator 2 is poor.

The Swedish WFD approach generates a status value for each parameter within a quality element, and then averages those values to produce the final

status. Using the example in Figure 5, after calculating all the EQR and status values, a status score is calculated using the EQR values of the class boundaries, indicating where within a status class the indicator is. Indicator 1 would have a status score of about 2.4 (almost half way along the “moderate” status) while Indicator 2 would have a value of 1.9 (close to the poor/moderate boundary). The resulting status class is the mean of the two values (2.15) which is above 2, so in the “Moderate” class. In this respect, the Swedish approach is more transparent, as the parameter status is combined. The Swedish approach also has the advantage that even where EQR based methods have not been developed, it is possible to include a classification of, for example oxygen status, in the final assessment using just the status score.

Both HEAT and the Swedish method give a measure of uncertainty in the status scores of the individual quality elements. The HEAT approach is more sophisticated than the Swedish WFD implementation, and allows more “user control” to comment on the quality of the assessment. The information is different however: the Swedish approach gives an idea of the uncertainty within a quality element, which may be caused by poor sampling, poor boundary setting, or a failure to understand how several indicators within a quality element relate to one another. The HEAT approach is more explicit, but works more at the level of individual indicators. The final confidence interval at the Quality Element level reflects only the qualitative judgement used at the indicator level, and says nothing about the variability of the indicator status within the quality element. The ideal may be a combination of the two approaches.

The weighting of the various indicators within a quality element is also useful – particularly when an indicator may only have partial relevance in describing the eutrophication status (for example, oxygen or benthic fauna in an area affected by long retention times or dredging). Use of the weighting tool should be constrained, however, as it is possible to use it to “turn down” the impact of an indicator, even in some cases to improve the overall status score of the quality element. Guidelines on the use of this term are needed.

A3.6 Summary

Implementation of the OSPAR Common Procedure is relatively straightforward, but leaves a great deal of freedom to experts to pick suitable metrics and sampling periods when making the assessment.

The Swedish implementation of the WFD is the opposite: clearly stated exactly what data periods are to be used and which parameters, from which sampling regimes. There is no weighting of parameters, but data availability issues do make it difficult to assess some water bodies without using “Expert Judgement”, and the presence of other activities, such as dredging, means that rather than being a pure eutrophication assessment tool, it is more a water quality assessment methodology. HEAT allows more expert

freedom than the Swedish WFD implementation, through its use of parameter weighting. The quality assessment possibilities are useful though give different confidence information to that obtained by the Swedish WFD implementation method.

A3.7 References

- Aarup, T. 2002, "Transparency of the North and Baltic Sea – a Secchi depth data mining study", *Oceanologia* vol 44 number 3, pages 327 – 337
- European Commission, 2008, "COMMISSION DECISION of 30 October 2008 establishing, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise", *Official Journal of the European Union* 2008/915/EC, available online at: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:332:0020:0044:EN:PDF>
- EU, 2000, "Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the Community action in the field of water policy", available online at: http://ec.europa.eu/environment/water/water-framework/index_en.html
- Hansson M., 2008, "Justering av databehandling av bedömningsgrunder för näringsämnen i Västerhavet", SMHI Report to Naturvårdsverket, SMHI Diary Number 2008/1193/1933; Naturvårdsverket Contract Number 502 0808
- Hansson, M., och B. Håkansson. 2005. Förslag till Vattendirektivets Bedömningsgrunder för pelagiala vintertida näringsämnen och sommartida effekterrelaterade näringsämnen, siktdjup och klorofyll i Kustvatten. (Ej fastställt, får ej åberopas som standard) SMHI Dnr 2005/1278/1933.
- HELCOM, 2009, "Eutrophication in the Baltic Sea – an integrated thematic assessment of the effects of nutrient enrichment in the Baltic Sea region", *Balt. Se. Environ. Proc.* No 115B.
- Håkansson B., O. Lindahl, R. Rosenberg, P. Axe, K. Eilola, B. Karlson, 2007, "Swedish National Report on Eutrophication Status in the Kattegat and the Skagerrak", Håkansson (ed.), SMHI Report RO 36, Diary No. 2006/1182/1933, available online at http://www.smhi.se/content/1/c6/03/00/40/attachments/RO_36.pdf
- Larsson U., S. Hajdu, J. Walve, A. Andersson, P. Larsson and L. Edler, 2006, "Bedömningsgrunder för kust och hav: Växtplankton och näringsämnen", 107 pp
- Miljöanalys, SLU, http://info1.ma.slu.se/ma/www_ma.acgi?StationX?ID=Intro&S=478
- Naturvårdsverket, 2007, "Bedömningsgrunder för kustvatten och vatten i övergångszon: Bilaga till Handbok 2007:4", available (in Swedish) from <http://www.naturvardsverket.se/>