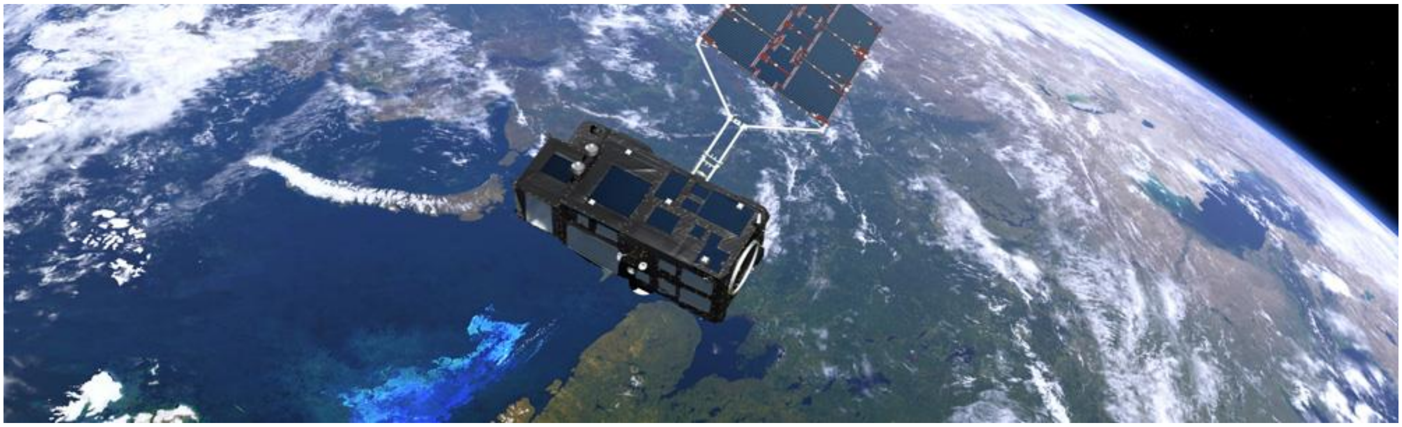
**Joint monitoring programme**
**of the eutrophication of the North Sea**
**with satellite data**

# ACTIVITY 3

## *–Towards a joint monitoring and assessment programme of eutrophication in the North Sea*



May 2019

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
ENVIRONMENT
Directorate D - Water, Marine Environment & Chemicals

Co-funded by the European Commission –DG Envrionment

**This report can be cited as**:

Markager, S., S. Upadhyay, P. Stæhr, H. Parner, H. Jakobsen, P. Walsham, K. Wesslander, D. Van der Zande, and L. Enserink (2019). Towards a joint monitoring and assessment programme for eutrophication in the North Sea. Activity 3 Report. 52pp.

**Authors & contributors**:

Stiig Markager[1], Sanjina Upadhyay[1], Peter Stæhr[1], Hjalte Parner[2], Hans Jakobsen[1], Pamela Walsham[3], Karin Wesslander[4], D. van der Zande[5] and Lisette Enserink[6]

1) Department of Bioscience, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark
*Corresponding author. markager@bios.au.dk, tel.: +45 61710081

2) International Council for the Exploration of the Sea, H.C. Andersens Boulevard 44-46, DK-1553, Copenhagen V, Denmark

3) Marine Laboratory, Aberdeen, Scotland, UK.

4) Swedish Meteorological and Hydrological Institute, SE 426 71, Västra Frölunde, Sweden

5) Royal Belgian Institute of Natural Sciences, Rue Vautier 29, 1000 Brussels, Belgium

6) Rijkswaterstaat, Zuiderwagenplein 2, 8224AD Lelystad, The Netherlands

**Financial support:** Co-funded by the European Union, DG-Environment EMFF Programme Implementation of the second cycle of the MSFD -Achieving coherent, coordinated and consistent updates of the determinations of good environment status, initial assessments and environmental targets (DG ENV/MSFD Second Cycle/2016), Grant Agreement No. 11.0661/2017/750678/SUB/ENV.C2 and the Ministry of Environment and Food of Denmark

**Acknowledgements:** This report is a detailed summary of work performed under Activity 3. We thank the project partners for valuable discussions during the project.

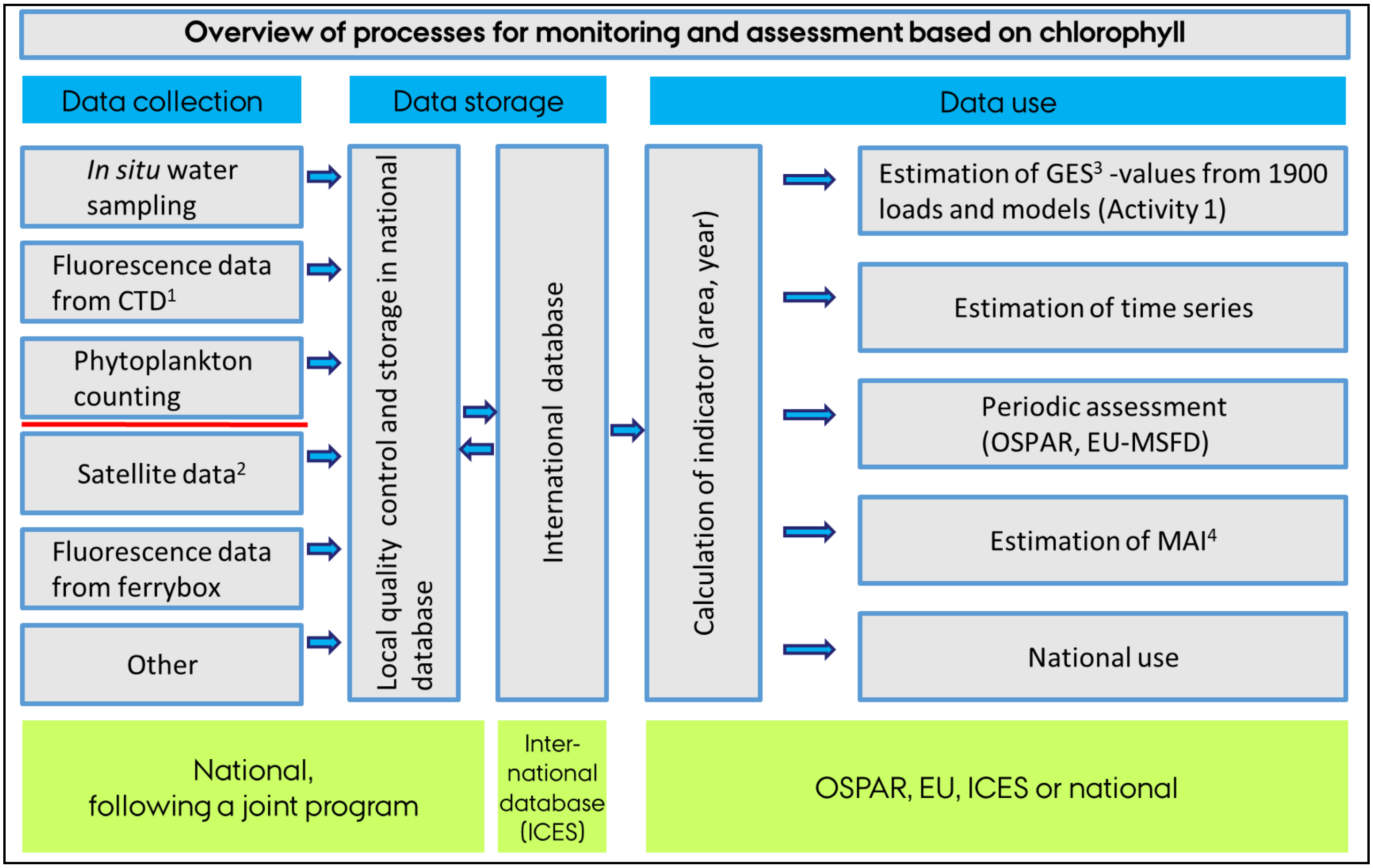
**Key words:** Eutrophication, monitoring, North Sea, OSPAR, chlorophyll, marine ecology, ocean colour, satellite observations, eutrophication assessment

**The JMP EUNOSAT project partners are:** RWS (NL), RBINS (BE), Deltares (NL), AU (DK), Ifremer (FR), PML (UK), MSS (UK), CEFAS (UK), SMHI (SE), IMR (NO), NIVA (NO), UBA (DE), BSH (DE), NLWKN (DE)

## Summary

In this report, we discuss various aspect of environmental monitoring in the North Sea with focus on chlorophyll and the use of remote sensing techniques, specifically estimation of chlorophyll from satellites. We also consider the use of monitoring data in the assessment of environmental status. The background for the analysis is criteria laid down in the GES Commission Decision (2017/848) under the Marine Strategy Framework Directive (MFSD) and the needs to improve OSPAR data assessment across the region. The report presents the results from Activity 3 in the project 'Joint monitoring programme of the eutrophication of the North Sea with satellite data' funded by the EU DG-Environment EMFF Programme 'Implementation of the second cycle of the MSFD' and with co-funding from the Ministry of Environment and Food of Denmark.

The report consists of an introduction treating the background and concepts for environmental monitoring. Eutrophication is described with particular emphasis on the usefulness of chlorophyll as an indicator. In the main part of the report, we try to outline directions for a monitoring and assessment programme for open marine areas such as the North Sea. The key figure (Fig. 6) gives an overview of the activities, and the topic in each box is discussed ending with a conclusion and/or recommendations. Together, these recommendations form the main outcome – an outline of a monitoring and assessment programme for the North Sea –and are summarised below.



See Fig. 6 for explanatory text.

The introduction of new techniques like estimation of chlorophyll from satellites gives rise to two problems/questions – how reliable are the estimates and how do we combine data from different techniques? These problems/questions are discussed in the last chapter using the Kattegat as a case. Here, we combine and analyse time series for *in situ* collected water samples and satellite estimates of chlorophyll.

The report is an integrated part of the project and should be read in connection with the reports from activity 1 and 2 (Blauw et al. 2019 and Van der Zande et al. 2019). Overall, the aims of the three activities are:

1) Derivation of threshold values for Good Environmental Status (GES) for nutrients and algae concentrations with a common a common method for all North Sea countries (Activity 1);
2) Generation and validation of a coherent multi-algorithm satellite-based chlorophyll-a product for the North Sea and the suitability of these products for eutrophication assessments (Activity 2)
3) Definition of coherent assessment areas with similar ecological and physical functioning (Activities 1 and 2 together)
4) Development of a potential design of a future monitoring and assessment programme (Activity 3).

## Recommendations

### Data collection

#### 3.2.1 Combining ship-based in situ sampling with other techniques

*Conclusion: Satellites and other remote sensing or automated techniques can provide valuable data for chlorophyll with high spatial and temporal coverage. The techniques can provide input information about the variability in chlorophyll concentrations that might help in the design of operational* in situ *monitoring. However, there is also substantial uncertainties. For satellites this is particularly in optical complex water such as shallow waters, areas with high CDOM concentrations or high turbidity. As ship based activities will continue,* in situ *sampling for chlorophyll can also continue at low additional cost. Moreover,* in situ *data for chlorophyll are essential for validation of the other techniques.*

*We recommend continued* in situ *sampling, but with adaptations so that it optimally supports calibration of other techniques like satellite observations, for instance by taking surface samples for chlorophyll when there is a satellite overpass.*

#### 3.2.1.2. Combine chlorophyll sampling with other monitoring activities

*Conclusion: We recommend that EU or OSPAR initiates a process that builds on and continues the work by EFARO with the goal of combining sampling for chlorophyll with other monitoring activities and to propose a pilot study and a detailed plan for a coherent monitoring programme covering all environmental issues and indicators, including fish stocks. The goal is to develop a sampling programme that supports the monitoring of eutrophication including the use of satellite observations for chlorophyll and water clarity determination. The idea of regional pilot studies seems to be a realistic first step, involving all stakeholders.*

#### 3.2.1.3 Combining environmental monitoring with other activities at sea

*Conclusion: We recommend that North Sea countries take the initiative to investigate the possibilities of implementing a coherent solution for marine survey activities in order to reduce costs and maximise benefits.*

#### 3.2.1.4 Joint cruises and ownership of the ships

*We recommend that North Sea countries, potentially through a follow up project, analyse the state of the current fleet of monitoring/research vessels operating in the North Sea suggest new possibilities for ownership of research vessels, which may, for instance, be a shared fleet or shared ownership among neighbouring nations.*

#### 3.2.1.6 Analytical procedure for chlorophyll

*Conclusion: Overall, chlorophyll concentration obtained with different method can be used together. It seem like HPLC-values for chlorophyll* a *is about 80 percent of values for total chlorophyll. However, we strongly recommend that this conversion factor is controlled on a regular basis. Similar, inter comparisons should be made if different solvents (acetone or ethanol) are involved. Further, we recommend that laboratories participate in regular external quality assurance schemes. We recommend that EU or OSPAR initiates a working group whose task will be to include in its work programme development of a common procedure for measurement of chlorophyll. As the outcome of the work has significant economic implications, it should involve national authorities. The options for such a working group could be:*

*1: to agree on the use of HPLC – costly, but specific and useful as it can detect changes in pigment composition with changes in nutrient richness.*

*2: to agree on another cheaper but still common method for chlorophyll measurement. Here, fluorescence on extracts might the best option as it is sensitive and easy to perform.*

*3: to allow for the use of different methods for measuring chlorophyll. This will require continued inter-comparisons and submission of metadata on methods and calibration along with values to a common database. Inter-comparison of data is strongly recommended for all options.*

*4: to supplement chlorophyll measurements with direct measurement of pigment absorption and coordinate this with satellite observations.*

### 3.2.2 Fluorescence data from CTD-profiles

*Conclusion: We recommend that CTD-profiles for* $Flu_{Chl}$ *are used in combination with* in situ *sampling in order to obtain average concentrations of chlorophyll that are less affected by the depth for* in situ *sampling than if an average is based on sampling from discrete depths.*

### 3.2.3 Phytoplankton counting

*Conclusion: We recommend that direct counting of phytoplankton remains a part of eutrophication monitoring and that new technologies are developed based on flow cytometry and image analysis.*

### 3.2.5 Fluorescence data from ferrybox systems

*Conclusion: The use of ferrybox systems for chlorophyll monitoring depends on quality control of the dataset, in particular accounting for the problem with the circadian rhythm of the fluorescence signal:chlorophyll concentration ratio. We recommend that resources used to develop this technique further, are directed towards analysis of data before implementation of the technique on new ferry lines.*

## Databases

### 3.3.1 Local quality control and storage in local databases

*Conclusion: We recommend: 1) that local institutions perform the initial quality control and store data in a national database; 2) that data are not censored, i.e. storage of all values in the database with minimum four digits. This might involve changes or clarification at EU or national level of the legal framework for laboratories supplying environmental* *monitoring services; 3) that appropriate metadata on sampling, analytical methods and calibration are stored; 4) that the existing working group (MSFD CIS TG Data) explores pros and cons of a reporting format for metadata, to specify best practice for data handling and propose how this could be implemented; and 5) that EU and its member states ensure that existing environmental metadata are secured. Experience tells us that such information is lost over time along with institutional and personnel changes.*

#### 3.3.2.1 International databases - Hosting of database and data operation

*Conclusions: We recommend that the service currently provided by ICES will be continued either by ICES itself or a similar institution. We strongly recommend that long-term funding is secured and that the efforts to include metadata and quality control are strengthened.*

## Data use

### 3.4.1. Calculation of indicators, (area, year)

*Conclusions: We strongly recommend that the ecological-based areas defined in Activity 1 (Blauw et al. 2019) are used in the future and that further subdivisions of local areas are discussed by neighbouring countries with local knowledge. For each area, we suggest that an annual indicator value is calculated as the mean over the period March-September on non-transformed values for chlorophyll concentrations. The period can then be subdivided for specific areas if phytoplankton growth is limited by for instance phosphorus in the spring and nitrogen later on. That will still allow the calculation on one indicator value for the entire period from March to September.*

### 3.4.4 Periodic assessment (OSPAR, EU-MSFD)

*Conclusion: We recommend that EU and OSPAR adopt a procedure similar to the one used in HELCOM (outlined in Fig. 10).*

### 3.4.5 Estimation of maximum allowable inputs (MAI) of nutrients

*Conclusion: We recommend that quantitative relationships are established between each indicator and the nutrient input to each area.*

### 3.4.6 National use

*Conclusion: We recommend storage of data in open and accessible databases allowing the adjacent countries to utilise all available data in their management of local coastal areas.*

## Evaluation of the use of satellites for estimating chlorophyll concentrations in environmental monitoring for the Kattegat area and a method for combining *in situ* data and satellite data

*Conclusions:*

*1. Satellite observations reflect the chlorophyll concentration only at the surface. The average concentration of the mixed layer is only partly reflected and satellite observations do not allow determination of the deep chlorophyll maximum. However, the latter is not as important in a management perspective as eutrophication is reflected by the chlorophyll concentration in the mixed layer. It is a problem, though, that not even the mixed layer concentration is well represented by satellite estimates.*

*2. There is a random component in satellite observations in the Baltic Sea transition zone (the Kattegat region), probably due to its complex optical conditions created by the high CDOM concentrations, which hampers the use of ocean colour algorithms, thus increasing the error of satellite estimates of chlorophyll. However, for a seasonal average,* *the spatial variation is small for the time series, i.e. the different parts of the basin display the same inter-annual variability.*

3. *There are significant systematic errors in satellite-estimated chlorophyll concentrations in the 19-year time series analysed, especially outside the MERIS period (2003-2011). The MERIS and Sentinel-3 data allow for the use of neural networks (e.g. FUB-WEW), which are well suited for the complex Baltic waters. If such data are not available (e.g. SeaWifs, MODIS), a correction procedure can be applied if* in situ *observations are available for an area, rendering the satellite measurements useful. Therefore, in the current state, satellite observations cannot be used for the Baltic transition zone for monitoring or management purposes without validation with* in situ *observations. In the future, when a long standing and coherent satellite programme has been in operation for a period of, for instance, 10-20 years, sufficient evidence might have been gathered to permit the use of satellites as a standalone technique.*

4. *Based on the observations from coastal stations, satellites seem to have problems with measuring chlorophyll closer to the coast than ca. 5 km. Further limitations may apply for shallow areas or areas with high turbidity.*

## Table of contents

# 1. Introduction

## 1.1. Background – why environmental monitoring?

Marine ecosystems provide a range of important ecological and economic services. It is therefore essential for society to know the environmental state of marine ecosystems such as the North Sea. This necessity has increased over time as the anthropogenic pressures exerted on the sea have grown in the form of, for instance, nutrient loadings, a changing climate and overexploitation of fish stocks. Several of the major biogeochemical cycles for substances such carbon, nitrogen and phosphorus are now disturbed to a level where they essentially are controlled by human activities, affecting not just the health of marine systems but also the functioning of these and the entire biosphere. This involves a decrease in many ecological services, for example clean air, fish stocks, biodiversity and recreational possibilities associated with a healthy environment. As environmental problems often cross national boundaries, in particular those associated with the sea, there is a growing need for comparable, reliable and affordable methods to monitor changes in environmental status to ensure proper knowledge-based management.

Fortunately, our ability to monitor the environment has improved due to the availability of new techniques such as the development of automated systems for data collection on different platforms like buoys, ferrybox systems and autonomous underwater vehicles (AUVs), advancement of satellite based technologies and drones. Our understanding of processes and relationships within marine ecosystem has also increased. Long time series of data have given rise to the development of detailed hydrodynamic-mechanistic ecosystem models and statistical models, which are important new tools (e.g. Almroth and Skogen 2010, Meier et al. 2011, Hinsby et al. 2012, Timmermann et al. 2010, Timmermann et al. 2014, Erichsen et al. 2017). These new methods allow us to define the pressures causing unwanted changes in the ecosystems and ultimately to quantify the relationship between pressures and state of an ecosystem. The latter is essential for a knowledge based political decision process and for efficient management. Figure 1 outlines the processes involved in the monitoring and assessment of marine systems.

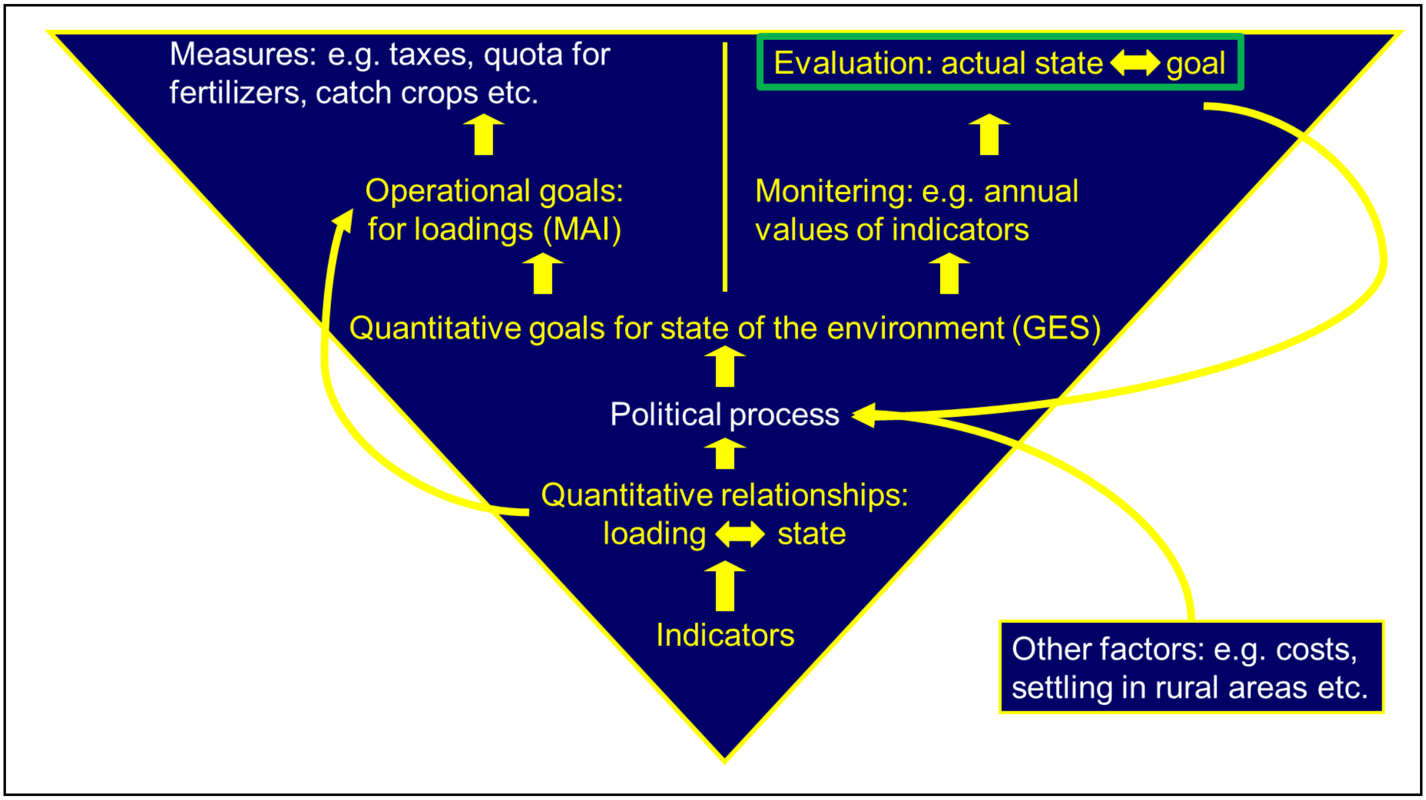


Fig. 1. An outline of the activities in monitoring and assessment and, ultimately, the management of the environmental state of marine systems related to eutrophication. Yellow text represent concepts or processes that were treated in the project and to various degrees described in this report. White text represents concepts or processes that belong to the political part of the process, which is what the concepts developed in this report will support. The 'Quantitative goals for state of the environment' correspond to GES (Good Environmental Status) in the MFSD but in a generic sense it can have other definitions. When the guidelines for determination of GES, for instance based on reference conditions, are decided at the political level, as it is the case for the EU directives, the process become a technical process as described in Fig. 2. The top right box (green frame) corresponds to the activity 'Marine management' in Fig. 2. MAI is the abbreviation for 'maximum allowable inputs', which is the maximum sum of input of nutrients that corresponds to the goal for environmental state. For marine areas is it the sum of riverine inputs, atmospheric deposition and point sources. For marine areas dominated by riverine input, like many estuaries, this correspond to a concentration of nitrogen or phosphorus times the average flow. Inputs of nutrients are equivalent to pressures in the DPSIR-terminology. 'Measures' (top-left) are concrete measures that can change the nutrient inputs.

Management efforts to mitigate human-related perturbations have resulted in several international and national environmental monitoring programmes of which many target the marine environment. International agreements aiming to improve water quality include, among others, EU's Water and Marine Strategy Framework Directives (WFD, MSFD) (European Commission 2000, 2008), the regional Baltic Sea Action Plan (BSAP) from HELCOM (HELCOM 2007) and the OSPAR North-East Atlantic Environment Strategy 2010-2020 (OSPAR Agreement 2010-3), which all stress the need for systematic monitoring and regular assessments of ecosystem status, including open seas like the North Sea. The approaches for assessing water quality status differ according to the method used but generally included a previously defined baseline against which the observed parameters are assessed. The purpose is to determine whether the objective of Good Ecological/Environmental Status (GES) is reached for each parameter and for the ecosystem as a whole (Borja et al. 2013, 2014).

Environmental policy development and management take place at different levels – from the UN-level, for instance through the 17 Sustainable Development Goals, over international fora like HELCOM, OSPAR, ICES and EU, to authorities at national, regional and local levels. Common for the different marine monitoring programmes is that they include a series of generic steps or processes, which are repeated and part of an information cycle (Fig. 2).

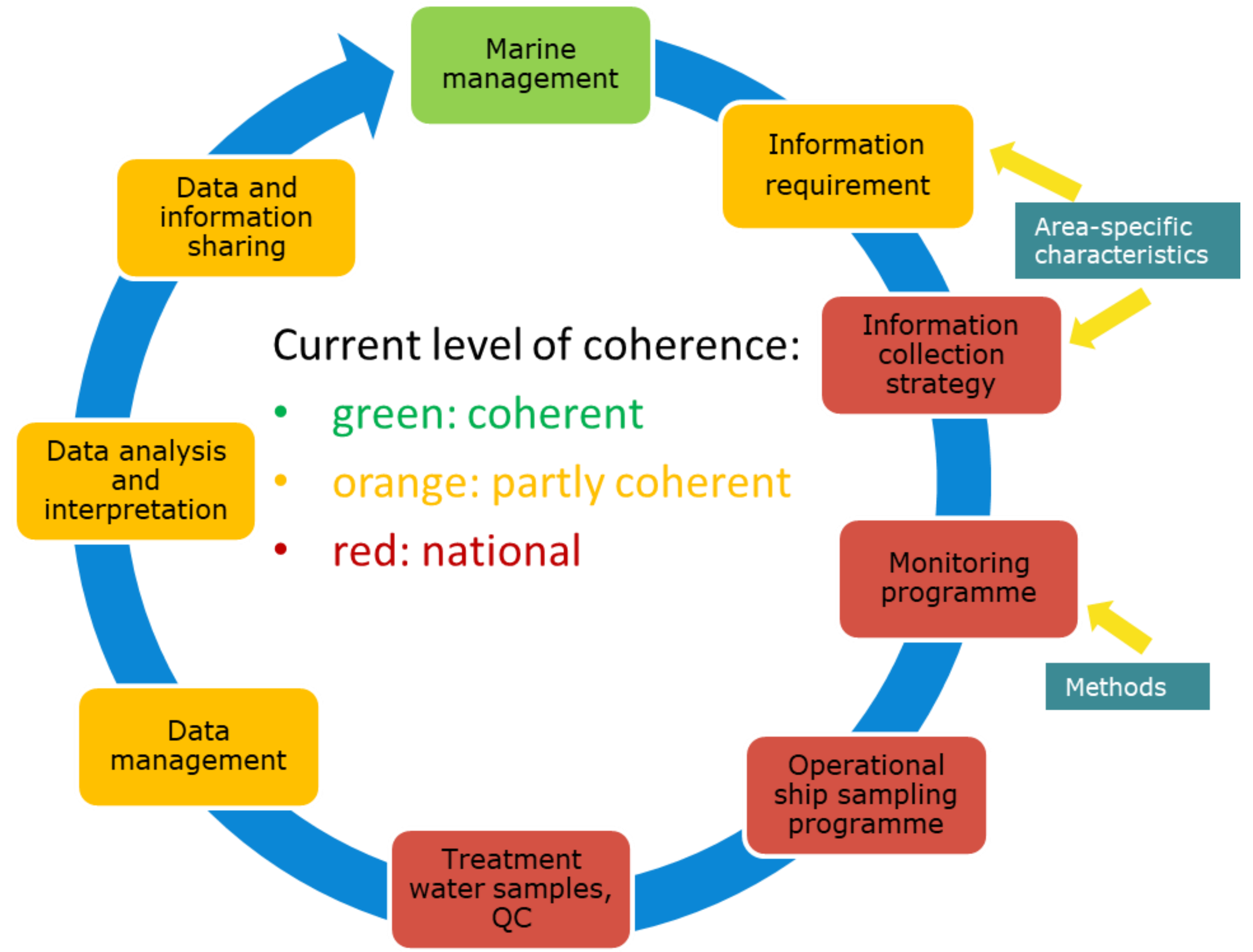


Fig. 2. The technical aspects of monitoring and assessment can be seen as an information cycle where 'marine management' (top box) is equal to the box in Fig. 1 (green frame) about evaluation of the actual state compared to the political decided goals for the environment. The outcome of the evaluation then represents the basis for a political decision when or if actions are needed. The information cycle is generic for all marine environmental indicators. The colours in this figure represent our view of the *current* level of coherence among North Sea countries in monitoring and assessment of chlorophyll. The JMP EUNOSAT project addresses all steps in the information cycle.

In a nutshell, the vision of the project is to make all parts 'green', which is equal to a fully coherent monitoring and assessment program for eutrophication. This will take years, but the aim of the JMP EUNOSAT project is to take steps in that direction for the North Sea.

## 1.2. Types of environmental monitoring

As national and EU-wide demands and regulations of GES assessment are increasing, application of efficient and reliable monitoring methods to assess water quality is of high importance. Environmental monitoring takes different forms depending on the purpose, the problem addressed and the ecosystem (Table 1). The present study focuses on operational eutrophication monitoring for MSFD and OSPAR.

Table 1. Four types of monitoring activities providing data for knowledge-based management of marine waters.

| Monitoring type | Description |
|---|---|
| Background monitoring | Collection of basic information about physical, chemical and biological variables for an ecosystem with the specific purpose of documenting the state of the environment. |
| Operational monitoring | Aimed at a specific and identified problem, for instance eutrophication, which is also the focus of this report. This type of monitoring often includes both variables describing the pressure of, for instance nutrient loadings, and their effects on the ecosystem. |
| Investigatory monitoring | Relationship between pressures and effects are unknown or uncertain. The monitoring often contains a significant research component. |
| Quantitative monitoring | Monitoring of the amount of a given resource, for instance a drinking water reservoir or a fish stock. |

### 1.3. New methods

The technological development impels implementation of new methods in environmental monitoring. One method that is particular relevant for monitoring of chlorophyll is satellite observations of ocean colour or remote sensing (RS). RS has existed since the 1990ties but are developing fast, particular with the launching of the Sentinel satellites. Recent studies indicate that RS has a high potential to supplement and optimise national marine monitoring programmes (Brockmann et al. 2004, Kratzer et al. 2014, Hossain et al. 2015, Brockmann Geomatics AB 2017, Stæhr et al. 2019). EU directives, OSPAR and HELCOM all recommend increased implementation of both RS and modelling data in monitoring and assessment activities. Monitoring by RS is expected to provide cost-efficient additional data that will help to fulfil the requirements/aims of the programmes by increasing the number of observations and improving spatial coverage (Strong and Elliott 2017). However, evaluations and performance assessments of the results in combination with the existing techniques are required before full implementation.

### 1.4. Aim of the report

The aim of this report is to describe how a monitoring programme for eutrophication can be designed for the North Sea and the adjacent areas covered by OSPAR with focus on chlorophyll and optimal use of new technologies such as satellites. Yet, it is important to see this specific activity in a broader context. Many other environmental issues exist – such as fish stocks, climate change and contaminants – and there are many operational advantages in merging the different types of monitoring activities. Thus, the aim of this report is also to place monitoring of chlorophyll in a broader context.

Monitoring of the North Sea was analysed in a previous EU project under the same MSFD support programme (Towards a joint monitoring programme for the North Sea and Celtic Sea, JMP NS/CS, 2015).

## 2. Operational monitoring of eutrophication

Eutrophication is an anthropogenic input of nutrients at a level that affects the structure and functioning of aquatic ecosystems and is a worldwide phenomenon (Conley et al. 2000, Cloern 2001, Boesch 2002, Kemp et al. 2005a,b). The relevant nutrients for the North Sea are phosphorus and nitrogen. Overall, the mechanisms of eutrophication are well described, although the quantitative relationship between nutrient loadings and environmental status for a specific area may vary and be difficult to determine. Other pressures, such as removal of top predators by fishing, damage to the sea floor or climate change, influence the effect of a certain nutrient input, which adds complexity to our understanding of eutrophication. This is a challenge to management as the necessary management measures then might depend on the development of other pressures. This calls for a holistic approach in the management of eutrophication.

### 2.1. Chlorophyll as an indicator of marine eutrophication

The basic sequence of processes in eutrophication is outlined in Fig. 3. The use of chlorophyll concentrations as an indicator of eutrophication is based on observations in a range of aquatic systems where increased loadings of nutrients have stimulated the growth of phytoplankton, leading to a higher phytoplankton biomass. The strength of the relationships between nutrient inputs and chlorophyll concentrations may vary between different regions depending on the physical, chemical and biological characteristics of the water body, but the general pattern is clear. A positive relationship between nutrient loadings and chlorophyll concentrations has been documented for a number of locations (Riemann et al. 2016).

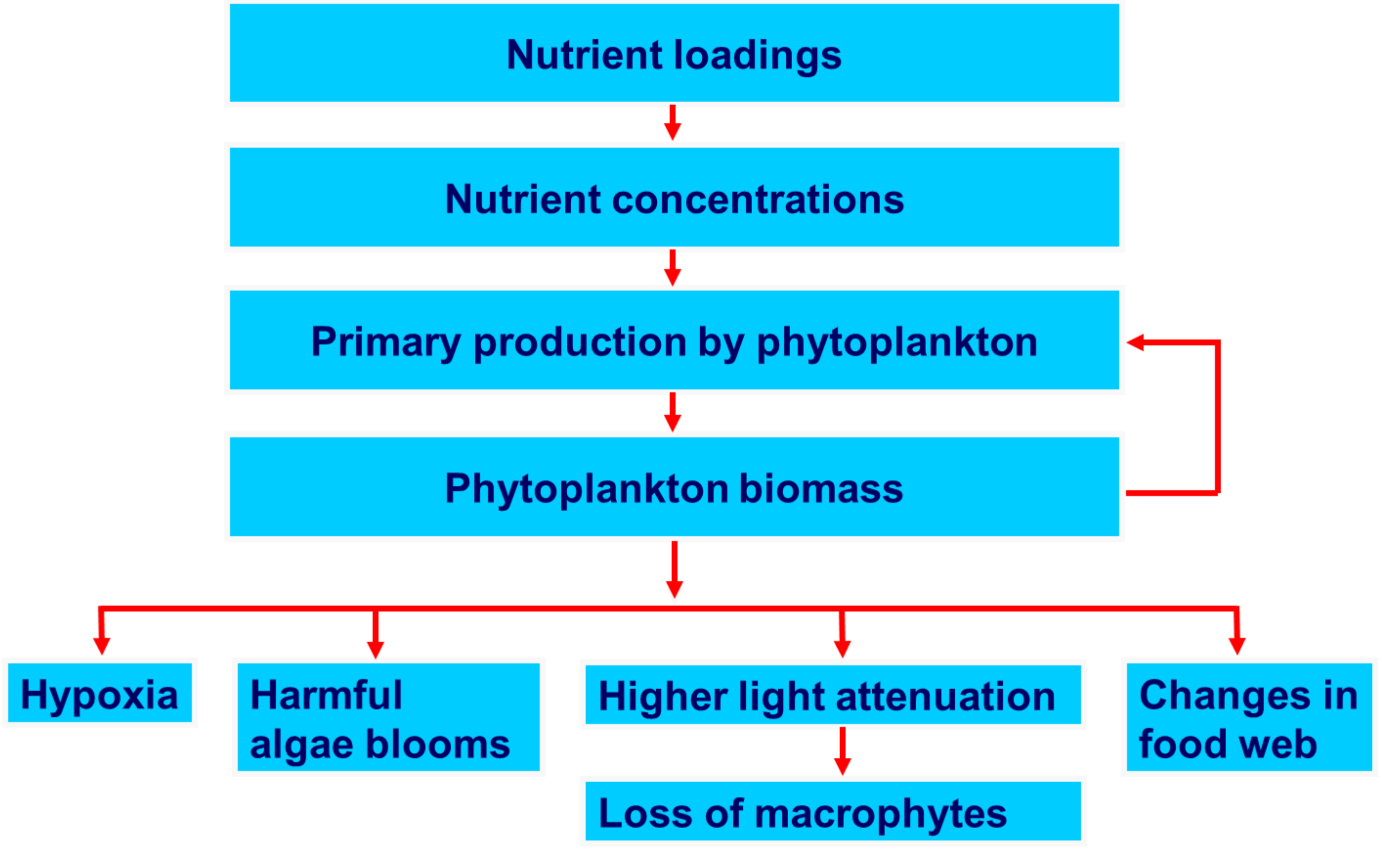


Fig. 3. Conceptual model of eutrophication. Arrows indicate a positive impact.

The situation can be a bit more complex, however. A higher concentration of phytoplankton provides more food for grazers such as zooplankton and benthic filter feeders. Particular long-lived benthic filter feeders can build up a biomass over years and hence increase the grazing rates on phytoplankton. Especially in well-mixed shallow water bodies, this may counteract a build-up of phytoplankton biomass but also in pelagic system can we observe a strong top-down control on the phytoplankton biomass (Lyngsgaard et al. 2017). Thus, minor or moderate changes in nutrient loadings may not necessarily translate into a concomitant change in chlorophyll concentrations. This top-down control is probably most pronounced at intermediate levels of chlorophyll concentrations. This has the implication that chlorophyll concentrations should not be used as the only indicator for eutrophication.

## 2.2. Chlorophyll – a proxy for phytoplankton biomass

### 2.2.1 The ratio between carbon and chlorophyll

The advantage of chlorophyll measurements is that chlorophyll is specific for phytoplankton and relatively easy to measure in a water sample. However, a problematic issue is the fact that a chlorophyll concentrations is only a proxy for phytoplankton biomass. The carbon to chlorophyll ratio (C:Chl) and the chlorophyll content per cell are highly variable depending on the nitrogen status of the cells (Fig. 4), and on other factors as well (Jakobsen and Markager 2016). At high TN concentrations (eutrophic conditions) both the chlorophyll concentration and the carbon biomass will be high However, at lower TN-concentrations, particularly summer chlorophyll concentrations will be lower, but since the C:Chl ratio is higher under nitrogen depletion, up to five fold, the carbon biomass will not be equivalently low. Thus, a chlorophyll-based indicator is likely to display a more pronounced change over time than the true change in carbon biomass.

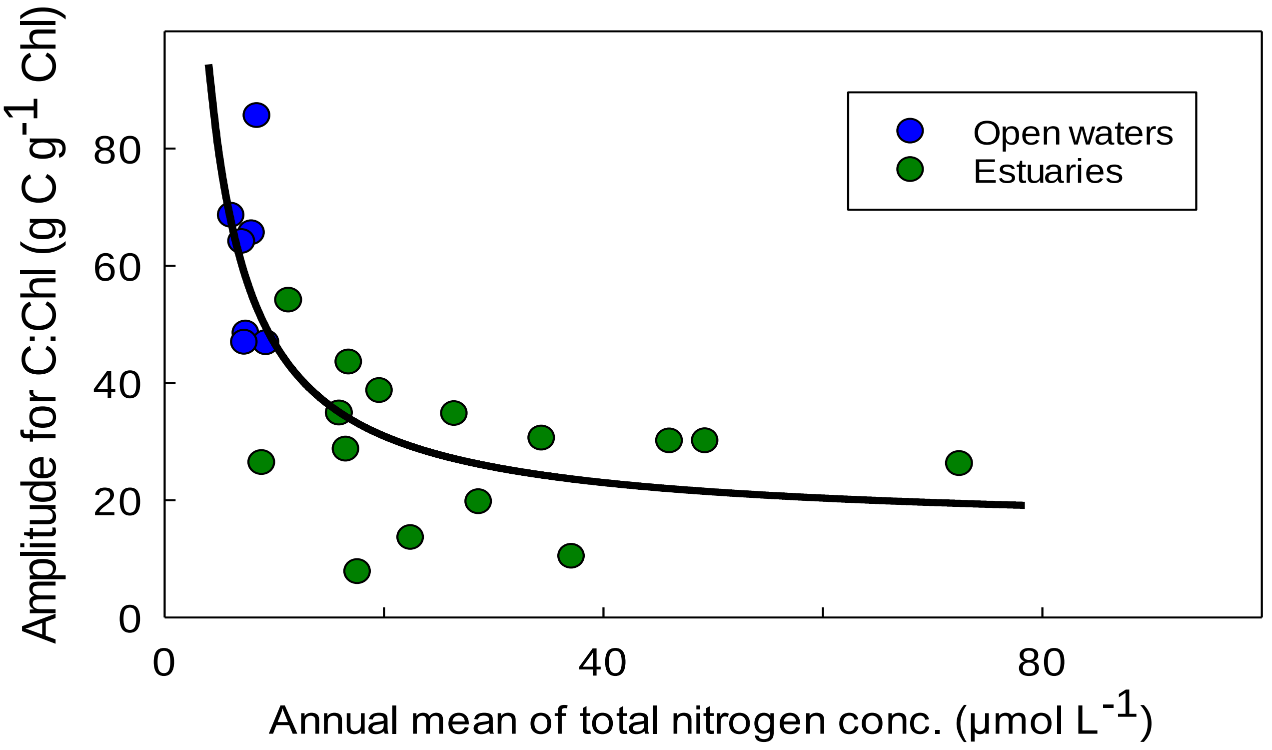


Fig. 4. The amplitude in the seasonal oscillation of C:Chl ratio for Danish estuaries as a function of the total nitrogen (TN) concentration. Adopted from Jakobsen and Markager (2016).

### 2.2.2 Chlorophyll – one among many pigments

Another uncertainty issue linked to the use of chlorophyll as an indicator of eutrophication is the fact that chlorophyll exists in several forms and that the chlorophylls are only one among many groups of pigments in phytoplankton. The only method that provides a specific value for chlorophyll *a* is HPLC. All other methods (see section about techniques) will, to a varying degree, co-measure other forms of chlorophyll than chlorophyll *a*. The common approach to solve this problem is to scale a result to chlorophyll *a*. This is common practice in connection with, for instance, fluorescence and satellite-based techniques where the measured property is light absorption (and for fluorescence the resulting emission at higher wavelengths) not specific for chlorophyll *a*. However, this introduces an uncertainty, particularly when combining data based on different techniques into a single indicator value, and it essential to consider this in the interpretation of the indicator.

In general, there is a good correlation between the different types of chlorophyll and between chlorophyll and other pigments, but in some cases significant variations occur. As an example, dominance of diatoms may result in a high ratio of chlorophyll *c2* to chlorophyll *a* (Fig. 5). This issue is not trivial in a monitoring and assessment context as the composition of the phytoplankton community will change systematically with changes in nutrient loadings and eutrophication status. Thus, a scaling factor, for instance between chlorophyll *a* measured with HPLC and a fluorescence based technique, is likely to change over time, particularly if the nutrient level is changing. Such a change will systematically affect a chlorophyll indicator.

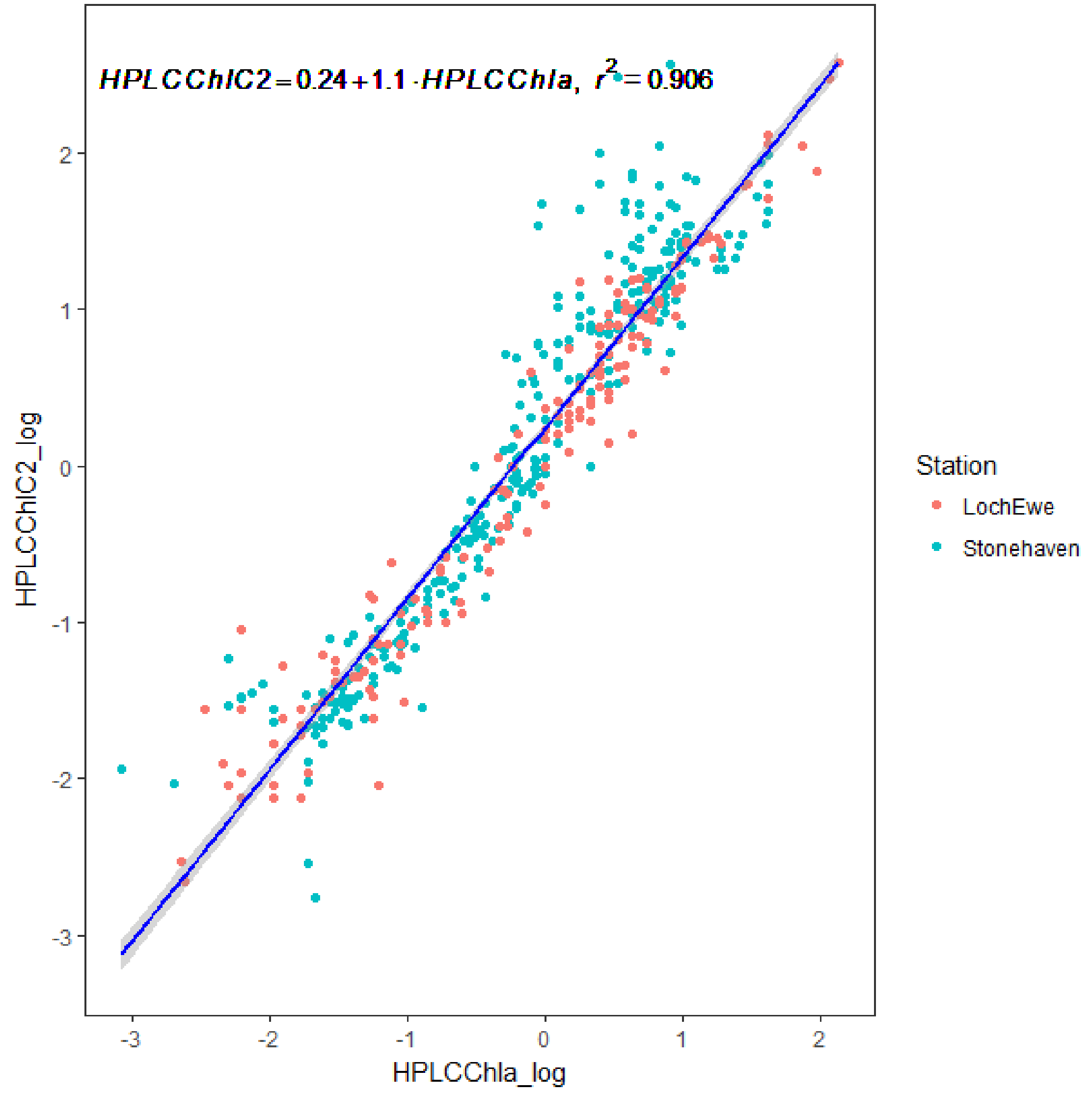


Fig. 5. Relationship (log-log transformation) between chlorophyll *c2* and chlorophyll *a* for a dataset from Scotland. Elevated values of chlorophyll *c2* occur at chlorophyll *a* values below -1.5 for Loch Ewe and for Stonehaven for chlorophyll *a* values (ln transformed) above zero.

### 2.2.3 Cell size and light absorption

A third issue – similar to that above - is that optical techniques such as satellite-derived estimates and fluorescence-based techniques (e.g. ferry box data) are sensitive to conditions that affect the amount of light absorbed by pigments – the pigment specific absorption. It is well documented that the ratio of light absorption to chlorophyll, often denoted a*, increases with decreasing cell size (Stæhr and Markager 2004). Since the cell size of phytoplankton increases with an enhanced nutrient supply (Stæhr et al. 2004, Cloern 2018), a systematic negative bias will occur for optical methods with rising nutrient supply.

### 2.2.4 Vertical distribution of chlorophyll

A fourth issue regarding the use of chlorophyll as an indicator of eutrophication concerns the distribution of the phytoplankton biomass and hence chlorophyll in the water column. Depending on water clarity, satellite-based observation typically detects the concentration of chlorophyll within the first few metres from the surface and a fluorometer on a ferrybox detects the signal from the depth of the intake, usually three to five metres. This contrasts to *in situ* sampling where water samples are collected from the entire water column (see 3.2.1 and 3.2.2). Techniques that only detect chlorophyll close to the surface, such as satellites, adequately measure chlorophyll under eutrophic conditions, because it is the chlorophyll in the mixed surface layer that most closely reflects eutrophication (Lyngsgaard et al. 2014). The reason is that eutrophication cause an increase in light attenuation, which reduce the amount of light at depth. This can reduce both the benthic vegetation (Krause-Jensen et al. 2012) and the occurrence of deep chlorophyll maxima, where phytoplankton otherwise will thrive with access to nutrient from below and light from above (Lyngsgaard et al. 2014). Thus, eutrophication induce a shift, where the biomass of phytoplankton in the mixed layer will increase relative to the biomass around the pycnocline. In consequence, techniques that mainly measure close to the surface will detect the fundamental effect of eutrophication.

It is, nevertheless, essential to account for this effect when merging data from different sources or, for instance, when comparing the present state with a GES-value estimated with a different sampling strategy or when analysing time series. For such tasks, it is very important to clearly define the depth interval over which chlorophyll is estimated and in monitoring guidelines ensure that effects of the effective depth interval are accounted for.

## 2.3 Specific objectives of operational monitoring

Using chlorophyll as an indicator of eutrophication implies that good status is achieved when the chlorophyll concentration during a number of consecutive years is below a GES-value defined from historical data when nutrient loadings were low (see Activity I and II reports, Blauw et al. 2019 and Van der Zande et al. 2019). According to this, the objectives for an operational monitoring of eutrophication in the North Sea are: 1) to follow the development of the chlorophyll indicator over time; 2) to provide the data necessary to set up quantitative relationships between nutrient loadings and GES status and 3) to estimate the required reductions in loadings for areas not in good status. In a management context, the first objective means providing data for the agreed status reports, for instance the common procedure in OSPAR and reporting under the WFD and the MSFD. Under the MSFD there is an explicit requirement to strive for comparability and coherence of GES assessments between EU member states. In addition, there are often national demands for consistent time series. Given that abatement measures are costly, there is a legitimate political focus on the environmental improvements following actions plans for a better marine environment. In practical terms, these objectives mean that we calculate an annual indicator value for each assessment area so that it is possible to follow the changes in the environmental status for an area.

The following requirements must be fulfilled for a joint monitoring programme and data management using chlorophyll as indicator:

1. The chlorophyll indicator values should represent an ecological homogenous area (see Activities 1 and 2, Blauw et al. (2019), Van der Zande et al. (2019). In some coastal areas, we see strong gradients from the coast and outwards. In such cases, the monitoring should be designed so that it describes the gradient in a consistent way (*cf.* Blauw et al. 2019).
2. Within an area, the period covered by a chlorophyll indicator should reflect the ecological effects of nutrient loadings (growing season and effects of nitrogen or phosphorus). This has implications for the period that the indicator covers. As an example, in areas where phosphorus is limiting during part of the growing season and nitrogen is limiting in other parts of the year, the indicator must be divided into the relevant periods in order to directly link pressures (loadings of either phosphorus or nitrogen) to the indicator in a precise manner. (see 3.4.1.2).
3. Sampling and analytical techniques should be comparable between countries sharing the same assessment area. If the techniques differ between countries, which is often the case, and harmonization of techniques is unrealistic, conversion factors should be developed and applied.
4. The indicator must be consistent over time. Thus, if methods have changed over the years, implications for consistency in time series must be tested and accounted for with appropriate corrections. Preferentially, a time series for the North Sea should start in the 1980s and 1990s where nutrient loadings were highest in order to describe the improvements due to already achieved load reductions. In OSPAR, load reductions are compared to the loadings in 1985. The estimates derived from new techniques should therefore be validated against existing data.
5. The data set used to estimate reference conditions and hence thresholds for GES should be based on the same techniques as used for the indicator. If there are discrepancies, the effects should be tested and accounted for through appropriate corrections.
6. The basis for the indicator, i.e. the measuring techniques and methods used to calculate the indicator, should develop over time and allow for incorporation of the newest techniques.
7. The exact way an indicator is calculated should be transparent and based on data accessible in open international databases so that the values can be tested and recalculated over time. An important issue is that the effect of each choice in the procedure can be tested and documented.
8. When an indicator is based on data obtained through different measuring techniques, the combination of data e should be based on scientifically documented techniques and in particular comply with requirements 3 and 4 (see also section 4.3.3).

## 3. Recommendations for a eutrophication monitoring programme for the North Sea

In the following, we present our recommendations for a monitoring and assessment programme of eutrophication using chlorophyll concentrations in the North Sea. The recommendations are based on the work done in the JMP EUNOSAT-project and the discussions among the participants during the four workshops held as part of the project. However, the recommendations represent the views of the authors and are not necessarily agreed upon by all members of the project consortium.

For each element, a number of options are available. In the text we endeavour to describe the pros and cons for each option. However, there is not always a definite answer to what is the best option. The chosen solution will often be a compromise between scientific quality, costs, practical issues and political guidelines, for instance a wish for national control of monitoring and assessment of territories. National interests are an important factor as the future monitoring will build on previously collected data and represents national capacities and interest in maintaining the consistency of national time series.

The project identified the following drivers for change:

- *policy*: improving coherence of assessments (MSFD, OSPAR). This is an explicit requirement under the MSFD and regularly assessed by the European Commission on the basis of Art 11 reporting;
- *money*: the need for cost effective monitoring programs while more data are needed for MSFD implementation;
- *technology push*: more and/or better data through new techniques;
- *science*: improve understanding of ecosystem functioning.

Together, these drivers are expected to enhance the willingness of North Sea countries to adapt their monitoring programmes for eutrophication, in spite of national and institutional barriers. During the projects the ideas and results were discussed with OSPAR working groups and also with HELCOM. We hope that this has provided a foundation to test and potentially accept the JMP EUNOSAT approach for a coherent chlorophyll assessment framework.

### 3.1 Overview of the monitoring process

The monitoring and assessment of eutrophication include three steps: 1) data collection; 2) data storage and 3) data use (Fig. 6). Below, we will treat each element individually and provide recommendations.

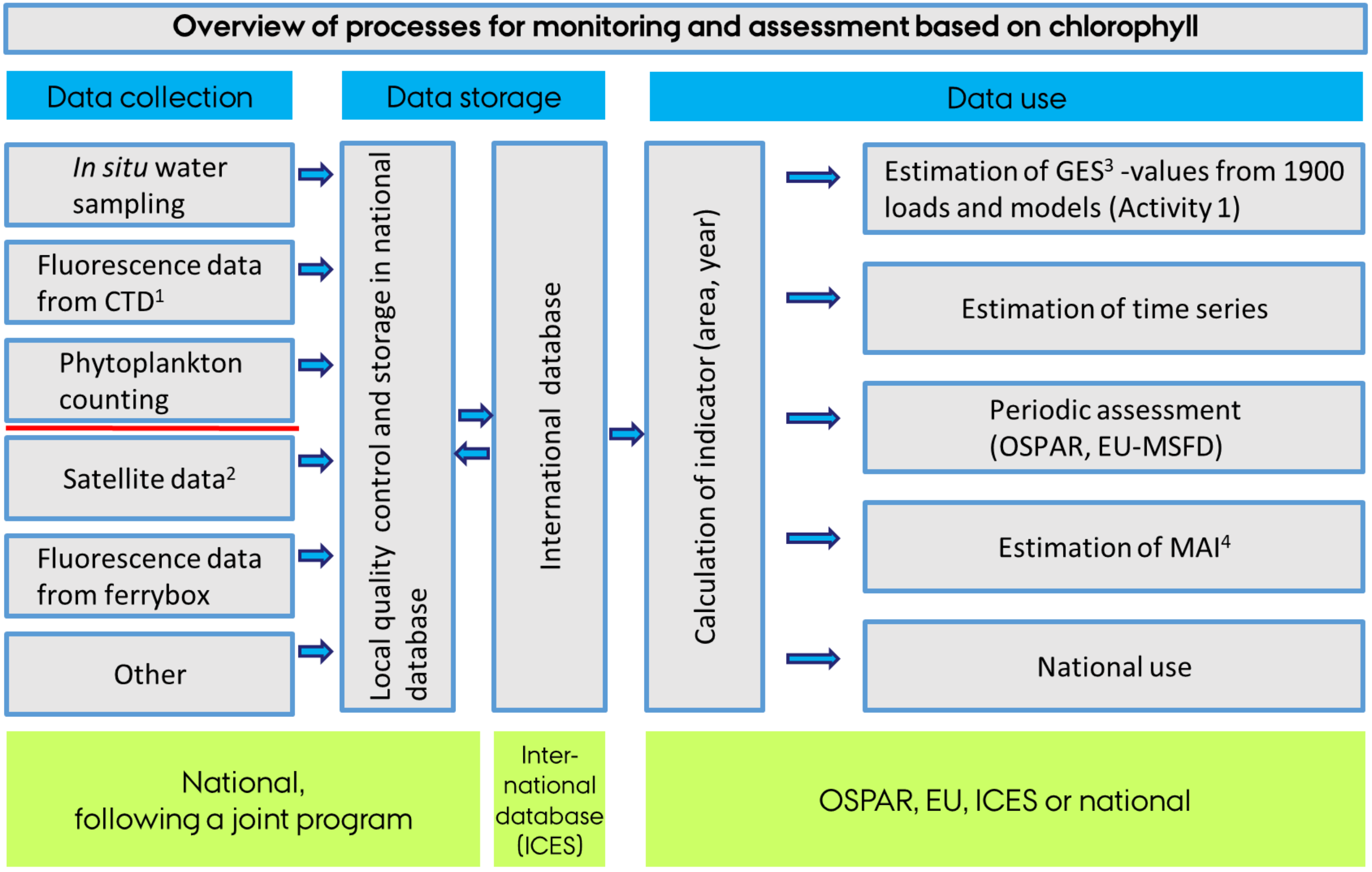


Fig. 6. The main elements in monitoring and assessment of phytoplankton biomass. The top row (blue boxes) shows the three main categories of activities. The middle row shows specific activities within each category. The bottom row shows where/who carry out the activities. Data collection techniques above the red line require operation of monitoring vessels, while techniques below the line do not require *in situ* sampling. 1) CTD is and abbreviation for conductivity, temperature and depth as is nowadays used for an ensemble of electronic instruments there automatically collect data for various variable when lowered down through the water column. 2) The satellites can be

operated internationally, as is the case for the Sentinel satellites, and international organisations like CMEMS are responsible for basic data storage and dispersal. The data are then processed and used for estimate of chlorophyll concentrations for the specific region of interest. The goal of JMP EUNOSAT is to suggest an improved and uniform procedure for estimating chlorophyll concentrations for the North Sea region. 3) GES is 'Good environmental status'. 4) MAI is maximum allowable inputs of nutrients for an area in order to obtain a certain environmental status.

## 3.2 Data collection

Sampling of water for chemical analyses such as chlorophyll or nutrients, the operation of a CTD (see Fig. 6) or other physical sampling, require navigating research vessel on open sea. This is in contrasts to satellite data, fluorescence data from commercial ships (ferrybox data) or other platforms like drones or automated underwater vehicles, there can be done without the operation of dedicated research vessels.

### 3.2.1 Combining ship-based *in situ* sampling with other techniques

*In situ* monitoring using dedicated vessels in open seas such as the North Sea is costly and the average cost is about 20.000 € per day. In coastal areas, smaller and hence cheaper vessels may be used but often at the risk of many idle days in harbour due to rough weather conditions. Therefore any technology that helps to reduce cost of data collection at open sea would be welcomed.

Monitoring activities in the North Sea include monitoring of fish stocks, physical condition like salinity and temperature, sampling of benthic communities in order to monitor biodiversity, marine mammal activity and contaminants in various matrixes. Other important indicators for eutrophication, beside chlorophyll, is nutrient concentrations, light attenuation (can to some degree be estimated from satellites), bottom water oxygen concentrations and benthic vegetation in coastal area. These indicators cannot be sampled remotely or automatically, therefore ship based monitoring activities will still necessary in the future. It is, therefore, unlikely that satellite or ferry box observations can completely replace ship-based sampling. This lead to the first conclusion and recommendation;

*Conclusion: Satellites and other remote sensing or automated techniques can provide valuable data for chlorophyll with high spatial and temporal coverage. The techniques can provide input information about the variability in chlorophyll concentrations that might help in the design of operational* in situ *monitoring. However, there is also substantial uncertainties. For satellites this is particularly in optical complex water such as shallow waters, areas with high CDOM concentrations or high turbidity. As ship based activities will continue,* in situ *sampling for chlorophyll can also continue at low additional cost. Moreover,* in situ *data for chlorophyll are essential for validation of the other techniques.*

*We recommend continued* in situ *sampling, but with adaptations so that it optimally supports calibration of other techniques like satellite observations, for instance by taking surface samples for chlorophyll when there is a satellite overpass.*

#### *3.2.1.1 Ship-based in situ sampling*

Traditionally, ship-based environmental monitoring is done by nationally operated vessels dedicated to research and/or environmental monitoring. Combination with other activities, such as fish surveys, occurs in some countries but is, according to our knowledge, the exception, and there is currently no coordination of the practical monitoring and sampling across national sea territories (de Boois and van Hal 2015). There are four areas offering possibilities of improvement: 1) the combination of sampling for eutrophication/chlorophyll with other monitoring activities, 2) combination of monitoring activities with other activities at sea such as navy surveys or search and rescue capabilities, 3) coordination of monitoring activities across national boundaries and 4) organisation and ownership of a fleet of monitoring/research vessels. The overarching goal is that once a vessel is at sea, it should solve as many tasks as possible in order to minimise operational costs and $CO_2$ emission related to burning of fuel. However, a fundamental requirement for a joint programme is that all participants agree on a common protocol for sampling and sample analysis. The latter poses a challenge for the national programmes in the way they are designed today since a common protocol will inevitably require that some countries change their protocol, which will have consequences for national time series. A common protocol will improve the reliability of the data and quality of the end result – an annual indicator value for chlorophyll.

Opportunities and challenges for collaboration between countries were identified in the former JMP NS/CS project (Birchenough et al. 2015). Regarding joint protocols and survey planning it was suggested that agreements are needed on a top level (e.g. Memorandum of Understanding) and subsequent levels (e.g. monitoring expert to monitoring practitioner), to ensure that the process is cascaded further and is effectively done. Furthermore, joint planning will be supported by sharing actual resources (e.g. staff, vessels, equipment) and subsequently sharing the final outcomes (e.g. data and knowledge) with all parties involved.

The aim of this section is to propose concrete actions to enhance collaboration and stimulate joint organisation of *in situ* monitoring.

#### 3.2.1.2 Combine sampling for chlorophyll with other monitoring activities

National ships coordinated by ICES are used for monitoring of fish stocks. A grid net of 10x10 nautical miles covers the North Sea and each grid cell is visited about six times per year, mainly from January to September. The sampling includes trawling, acoustic surveys and sampling for planktonic fish larvae and zooplankton. The ships are equipped with CTDs and do CTD-profiles in order to calculate the velocity of sound for the acoustic surveys. Trawling is mainly done during daytime and the ships are thus, to some extent, idle during night.

It seems obvious to combine monitoring for eutrophication with monitoring of fish stocks. The ships already cover a large part of the open North Sea during most of the growing season of phytoplankton and the ships already have the basic equipment at their disposal such as CTD, laboratory space and a crew trained for monitoring. Thus, for a small additional cost, compared with the total costs of operating the ships, simultaneous sampling for chlorophyll can be undertaken. In addition, water samples for nutrients and other water constituents of interest, including contaminants, may be made. Furthermore, other activities can be carried out like sampling of benthic communities and non-indigenous species, contaminants in sediments and biota, litter on the seafloor, deployment of buoys and listening devices for monitoring of harbour porpoises and visual observations of seabirds and marine mammals. A disadvantage may be that part of the crew will be periodically idle, for instance the trawling crew when the ship is occupied with other activities. In addition, there is a need for coordinating programmes and activities.

ICES has developed approaches and guidelines for environmental add-on parameters on fish surveys (IBTS, ICES 2016a and b) and also investigated the added value of fully integrated surveys for the purpose of integrated ecosystem assessments (ICES 2016c and ICES 2016d). In spite of these efforts, such as in the JPI Oceans pilot action Multi-use of infrastructure for monitoring, involving six North Sea countries (JPI Oceans, 2016), operational implementation has been only partly successful so far. However, Sweden has implemented concurrent sampling in their open waters combining fish surveys with water sampling. The main obstacle is probably that monitoring of fish stock is managed separately from other environmental monitoring. In some countries fish stock monitoring is undertaken by separate institutions, and in some by institutions belonging to different ministries as is the case in Germany. This means that there is no economic incentive for the institutions to work towards a change. Additional challenges identified in the JPI Oceans pilot action were (Remøe, 2018):

a. MSFD indicator monitoring needs are not sufficiently specific
b. lack of communication between experts in relevant JPI Oceans and ICES working groups
c. lack of additional funding to perform add-on tasks on fish surveys (IBTS)
d. fleet operators are not involved in the process
e. underdeveloped coordination and readiness among national agencies and ministries (as also mentioned above)

Cheaper and more coherent monitoring will be a benefit at both the national and at EU level and increase the chances of success. Initiatives for changes should come from the policy level (EU, OSPAR and national governments). Recently, the EU Marine and Fisheries Directors have agreed to explore possibilities for better collaboration between MSFD and the EU Common Fisheries Policy (CFP). A joint workshop involving fisheries and ecosystem experts was held to discuss joint monitoring as one of four themes and to propose a roadmap for further action (EU MSFD CIS, 2019).

The issue of enhanced collaboration between MSFD and CFP monitoring was initially brought forward by the European Fisheries and Aquaculture Research Organisation (EFARO) in a plea to 'rethink' and redesign fisheries and ecosystem monitoring. In order to facilitate a more effective and efficient design and cooperation in monitoring of the sea a number of questions need to be addressed, such as (EFARO 2017):

a. from the perspective of policy and advice for policy development, which data do we really need?
b. at what quality, quantity and precision?
c. and what will be the most efficient and effective way of collecting, processing and presenting these data?

EFARO proposed a set of pilot studies, focusing on three relevant areas (North Sea, Celtic Sea and Bay of Biscay) and requiring eight months and € 650.000. The results of these pilot studies would provide sufficient information to start a proper discussion within the community on a possible redesign of the system for data collection and advice, its efficiency and priorities. In these pilot studies, several scenarios would be evaluated including simulation of a theoretical cut in total available (national) budget in data collection (of 25% and 50%) to designing a new monitoring framework from scratch. The intention of this exercise is not to explore the possibilities for budget cuts, but to create a very strict setting for exploring room for improvement of survey efficiencies, given clear priority settings. EFARO suggested that this analysis is done as a regional approach. Note that this approach contrasts explicitly with current and past efforts exploring possibilities for adding MSFD monitoring to existing CFP/DCF surveys (EFARO, 2017). There was much gratitude for this initiative among Marine Directors, and the ideas were discussed in the joint fisheries and ecosystem experts workshop mentioned above. However, the proposals for pilot studies have not yet been implemented.

*Conclusion: We recommend that EU or OSPAR initiates a process that build on and continue the work by EFARO with the goal of combining sampling for chlorophyll with other monitoring activities and to propose a pilot study and a detailed plan for a coherent monitoring programme covering all environmental issues and indicators, including fish stocks. The goal is to develop a sampling programme that supports the monitoring of eutrophication including the use of satellite observations for chlorophyll and water clarity determination. The idea of regional pilot studies seems to be a realistic first step, involving all stakeholders.*

#### 3.2.1.3 Combining environmental monitoring with other activities at sea

Other marine activities are search and rescue capabilities, fighting of oil spills and maintenance of sovereignty over areas. Today, these activities are often undertaken by the navy or coast guard. The countries also have dedicated research vessels that have the same the equipment that is needed for marine monitoring. Thus a potential exists for combining these activities with the above-described marine monitoring with the opportunity to also host smaller research projects. Some countries have integrated environmental monitoring and research on navy ships in the Arctic with good results.

*Conclusion: We recommend that North Sea countries take the initiative to investigate the possibilities of implementing a coherent solution for marine survey activities in order to reduce costs and maximise benefits.*

#### 3.2.1.4 Joint cruises and ownership of the ships

The optimal solution is to aim for coordination of monitoring activities across national territories as is currently the case for the monitoring of fish stocks of which nationally owned ships are currently responsible. Often, these ships also operate as national research vessels. The benefits of a simple solution when monitoring cruises crossing national boundaries were analysed in de Boois van Hal (2015).

*We recommend that North Sea countries, potentially through a follow up project, analyse the state of the current fleet of monitoring/research vessels operating in the North Sea suggest new possibilities for ownership of research vessels, which may, for instance, be a shared fleet or shared ownership among neighbouring nations.*

### 3.2.1.5 Sampling strategy and analytical procedures

As mentioned above, a common protocol for sampling and analytical procedures is necessary for a joint sampling programme that includes cross-boundary cruises. Such a protocol will also be advantageous for national cruises as it will enhance the comparability of data among neighbouring countries. A disadvantage is, though, that such a protocol will most likely entail changes from their present procedure. The consequence is that time series will be interrupted and introduce differences compared with other data collected in national coastal waters unless a common agreement can be made both at the national level and with other international fora, for instance HELCOM. Particularly the interruption of time series can be a problem as consistent time series are crucial for analysing of the effectiveness of measures introduced to reduce nutrient loadings. Below we have listed suggestions for a common procedure:

1. Samples are taken at 1, 5, 10, 15, 20 and 30 m depth as well as in the deep chlorophyll maximum. Deeper samples can be included, if necessary.
2. Samples are taken with a rosette/CTD equipped with a chlorophyll fluorometer in order to giver a profile for chlorophyll with depth (see section 3.2.2.).
3. Samples are filtered on board and the filters kept in a minus $80^0$ freezer or in liquid nitrogen.
4. 1 m samples are collected at satellite overpass times to contribute to a common match-up database for ocean colour product validation in different water types. Samples should be analysed according to an agreed protocol (see below).

Countries that maintain national cruises and follow national procedure should regularly test the comparability of their procedure and the common protocol. All data must be submitted to the common database (see below).

### 3.2.1.6 Analytical procedure for chlorophyll

A critical issue is the analytical procedure to be used for chlorophyll analysis. The options are spectrophotometric detection after extraction in ethanol or acetone, fluorometric detection after extraction in ethanol or acetone or HPLC after extraction in acetone. The latter method is the most specific as it can separate the different pigments and types of chlorophyll, but its costs are significantly higher than for other techniques. A shortcoming of the other methods is that the detection efficiency of the different forms of chlorophyll (Chl *a*, Chl *b*, Chl $c_1$ and $c_2$) varies. Therefore, a result will to some degree be affected by the composition of the phytoplankton community (Fig. 5) and should be given as total chlorophyll or just chlorophyll but not as chlorophyll *a*. This problem cannot be entirely solved by inter calibration exercises, as the conversion from one method to another depends on the composition of the phytoplankton community. On the other hand, there is reasonable agreement between the methods, except that HPLC-values for chlorophyll *a* are lower than chlorophyll values obtained using the other methods, which is due to the fact that the HPLC is more specific than the other methods (see next paragraph).

Previous reports (Beretta-Bekker et al. 2015) have concluded that 'converting chlorophyll values from one method to another is considered not reliable' and that 'the use of fixed factors is no option'. This study was based on the idea that a single conversion factor would be able to combine *in situ* data obtained by between different techniques, for instance for the assessment of a transboundary area. However, our analysis does not support this conclusion. We have analysed the data from two Scottish sites (that also was used in Beretta-Bekker et al. 2015) and found a good correlation between values for chlorophyll *a* concentrations obtained HPLC and chlorophyll determined with fluorescence on acetone extracts (Fig. 7). Note that the slope is close to one (log-transformed values). The offset of 0.23 transform into a slope of 1.26 in absolute values, which imply that the HPLC-method determinate a chlorophyll *a* content of 0.8 of the total chlorophyll content measured by fluorometry on extracts for this data set. A similar conclusion was also found in the annex II of the report (Walsham 2015, see their figures 5 and 6). Our conclusion is therefore, that values obtained with different method can be combined and used together in an indicator for an area. The opposite conclusion, as cited above, would also imply that most of the scientific literature on pelagic ecology would be incorrect, since chlorophyll, which is usually used as a scaling parameter, is analysed with a variety of methods in the literature. We fully acknowledge that the relationship between results from different methods will vary. In fact, we could show that the variability in Fig. 7 could be explained by the variation in the content of chlorophyll $c_2$ (p<0.001). The effect can also be seen when comparing Fig. 5 and Fig. 7, where a similar deviation from the line can be seen at low and high values a among the two stations (note that the x-values are the same).

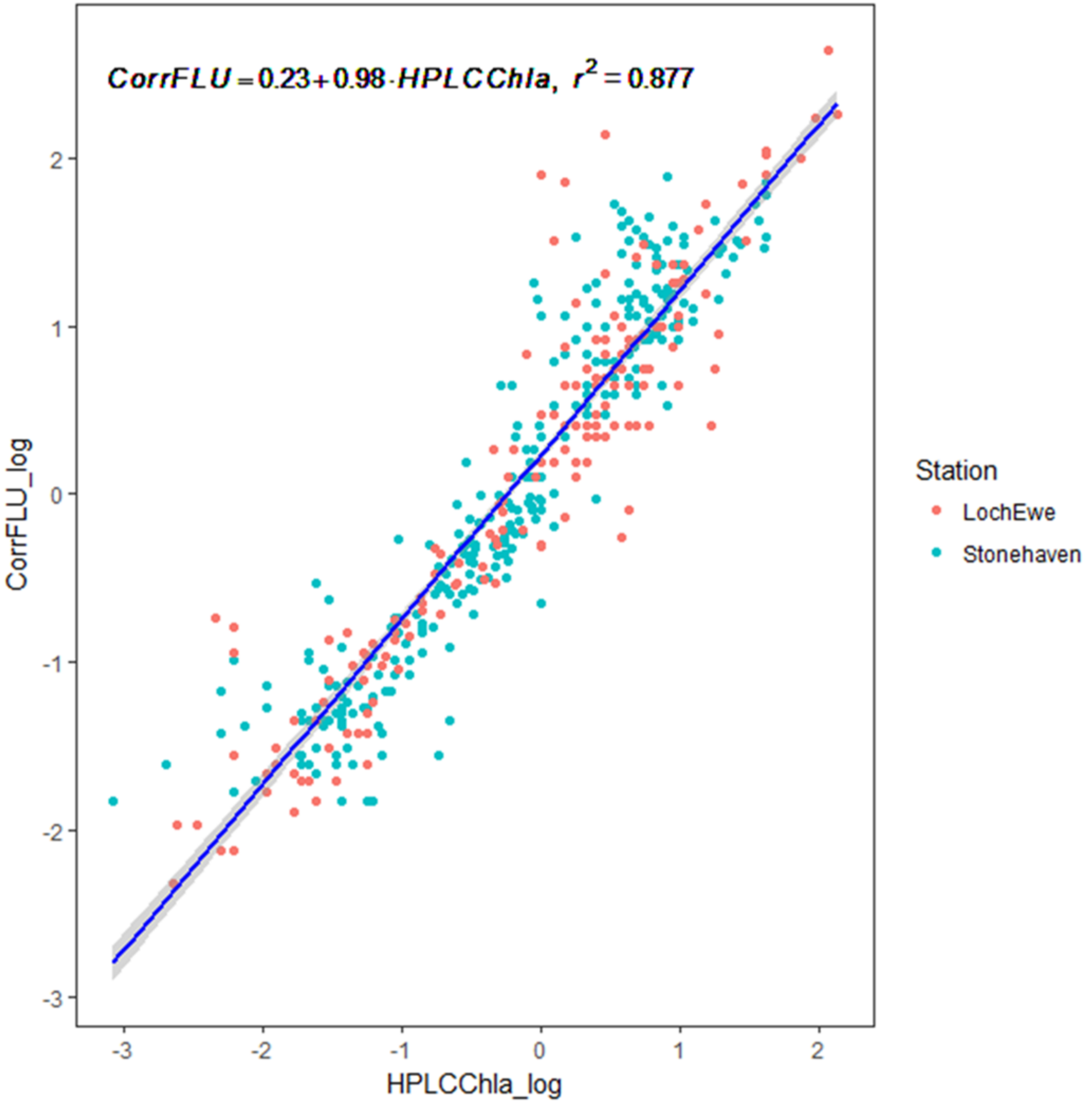


Fig. 7. Relationship (log-log transformation) between Chl *a* (HPLC) and total chlorophyll Chl (fluorescence on extract) for a dataset from Scotland. Positive deviations from the overall relationship is found for low and high values, as was the case in Fig. 5 (same x-values).

Our conclusion is that as a first approach, chlorophyll values obtained with different method are comparable and can be combined. However, values for chlorophyll *a* from the use of HPLC will be lower than values for total chlorophyll obtained with other method. The result based on the Scottish data set suggest a conversion factor of 0.8.

As to remote sensing, the scientific community favours the use of HPLC for validation, which offers some advantages. However, the statement that only HPLC-derived data can be used for validation of satellite based observations seems not well justified, as satellite observations basically detect the light absorption by the phytoplankton – a variable that is highly influenced by factors such as cell size, species composition etc. (Stæhr and Markager 2004). The absorption detected by satellites is either in the blue part of the spectrum or at the red absorption peak of chlorophyll. Particular in the blue part, but also to some degree in the red part, the absorption is the combination of the sum of overlapping absorption spectra from different pigments, where chlorophyll *a* is only one of them.

An alternative to HPLC is to abandon the use of extraction methods and measure phytoplankton absorption directly by the filter pad technique (Kishino et al. 1985, Cleveland and Weidemann 1993). This technique is comparable with the HPLC-technique as to costs and has the advantages that it measures the same characteristic – light absorption by pigments – that is estimated by satellites. Moreover, light absorption is a variable that is more closely linked to primary production than a chlorophyll concentration. Enhanced primary production by phytoplankton is the ultimate driver of eutrophication. However, such a shift will interrupt the existing time series for chlorophyll.

In conclusion, all methods for assessment of chlorophyll have advantages and disadvantages. Since the overall budget for a monitoring activity often is a fixed amount, the use of more expensive methods, like HPLC, will be at a cost of fewer samples. Thus, the benefits from a better and more accurate methods might by offset by a lower spatiotemporal coverage. However, the combination of satellite observation and *in situ* sampling might lower then number of chlorophyll measurements that is necessary. In general, it is important to recognise that chlorophyll, other pigments or light absorption are only proxies for phytoplankton carbon biomass, and caution should therefore be taken when assessing the responses in relationship with eutrophication and GES status (see section 2.2.2 and 2.2.3).

*Conclusion: Overall, chlorophyll concentration obtained with different method can be used together. It seem like HPLC-values for chlorophyll* a *is about 80 percent of values for total chlorophyll. However, we strongly recommend that this conversion factor is controlled on a regular basis. Similar, inter comparisons should be made if different solvents (acetone or ethanol) are involved. Further, we recommend that laboratories participate in regular external quality assurance schemes. We recommend that EU or OSPAR initiates a working group whose task will be to include in its work programme development of a common procedure for measurement of chlorophyll. As the outcome of the work has significant economic implications, it should involve national authorities. The options for such a working group could be:*

*1: to agree on the use of HPLC – costly, but specific and useful as it can detect changes in pigment composition with changes in nutrient richness.*

*2: to agree on another cheaper but still common method for chlorophyll measurement. Here, fluorescence on extracts might the best option as it is sensitive and easy to perform.*

*3: to allow for the use of different methods for measuring chlorophyll. This will require continued inter-comparisons and submission of metadata on methods and calibration along with values to a common database. Inter-comparison of data is strongly recommended for all options.*

*4: to supplement chlorophyll measurements with direct measurement of pigment absorption and coordinate this with satellite observations.*

### 3.2.2 Fluorescence data from CTD-profiles

*In situ* sampling is usually done with a combined CTD and rosette sampler where the CTD is equipped with a chlorophyll-fluorometer. This gives continuous profiles of chlorophyll fluorescence ($Chl_{flu}$) that can be used to estimate the average chlorophyll concentration in different strata of the water column (Fig. 8). The challenge is that the $Chl_{flu}$:Chl ratio varies depending on cell size, the density of chlorophyll in the cells and previous light exposure. Thus, the fluorescence profile must be scaled to *in situ* samples in order to provide a chlorophyll profile (see Lyngsgaard et al. 2014 for details). Figure 8 shows the pattern in the fluorescence yield (fluorescence:µg $l^{-1}$ Chl) with depth for stratified stations in the Kattegat and Belt Sea region. All six stations exhibit the same pattern where the yield increases with depth from the surface down to a certain depth after which the yield decreases again. However, the general pattern is that the yield becomes higher with depth than at the surface, which is also what should be expected from the physiology of algae. The maximum yields (annual average for the station) are between 16 and 75% higher at deeper depth than at 1 m and the maximum yield often corresponds to the depth of the deep chlorophyll maximum. It is clear that continuous scaling of the fluorescence signal from chlorophyll fluorescence instruments is necessary and critical for monitoring activities.

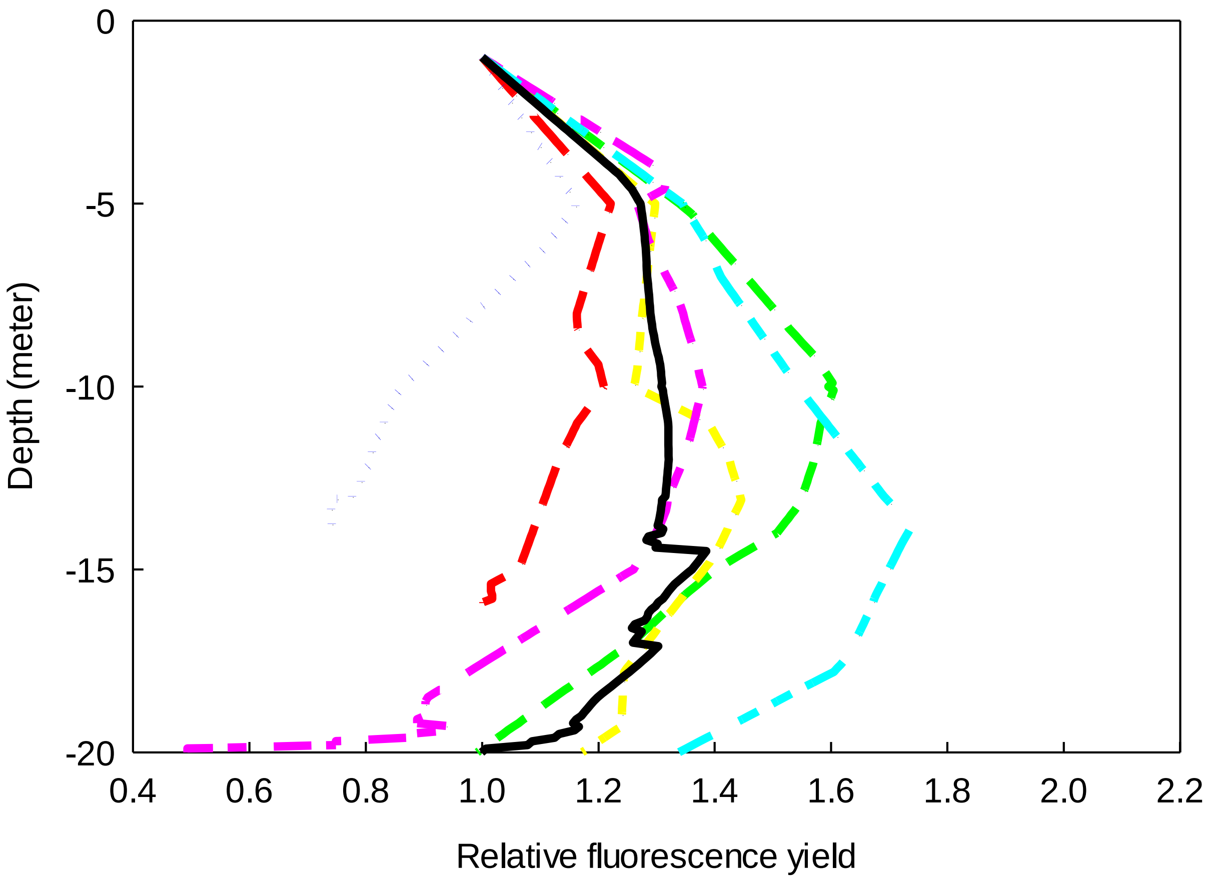


Fig. 8. Relative fluorescence yield with depth for six stations in the Kattegat and Belt Sea region in 2018. The coloured lines are annual means (approximate 26 profiles) for each station and the black line is the mean for the six stations. The ratio $Chl_{flu}$:µg $l^{-1}$ Chl is set to 1 at the depth of 1 m. The data for each station are obtained using the same instrument and calibration throughout the year.

Monitoring programmes in Belgium and Denmark systematically explore this opportunity to estimate the average chlorophyll concentrations in different strata of the water column based on CTD-profiles.

*Conclusion: We recommend that CTD-profiles for $Flu_{Chl}$ are used in combination with* in situ *sampling in order to obtain average concentrations of chlorophyll that are less affected by the depth for* in situ *sampling than if an average is based on sampling from discrete depths.*

### 3.2.3 Phytoplankton counting

Direct counting of cell number in combination with measurements of cell size gives the most direct value for phytoplankton biomass in term of cell number and cell volume (Harrison et al. 2015). The cell volume can then be converted into carbon biomass assuming a cell volume to carbon ratio (Menden-Deuer & Lessard 2000). This can further be summarised into group-specific biomass or total community biomass (Jakobsen et al. 2015). This data set may, in turn, be used to establish various indices used to identify changes in ecosystem services. An analysis of 30 years of monitoring data showed that the variability in the carbon biomass of phytoplankton was lower than the variability for chlorophyll concentrations when the nitrogen availability changed. The reason is that phytoplankton accumulates carbon in the cells under nitrogen depletion (Jacobsen and Markager 2016). Thus, the decline in chlorophyll concentration observed at decreasing loadings will overestimate the decrease in phytoplankton biomass in carbon units.

Counting and species identification are expensive so the method is not suitable on a large scale, although there are alternatives – see next paragraph. However, monitoring of species composition and the carbon:chlorophyll ratio on a few stations in a region is valuable for development and implementation of the MSFD regarding food webs and eutrophication where relevant criteria are:

1) D4C1 The diversity (species composition and their relative abundance) of the trophic guild is not adversely affected due to anthropogenic pressures)
2) D4C2 The balance of total abundance between the trophic guilds is not adversely affected due to anthropogenic pressures)
3) D4C3 The size distribution of individuals across the trophic guild is not adversely affected due to anthropogenic pressures)
4) D4C4 Productivity of the trophic guild is not adversely affected due to anthropogenic pressures
5) D5C3 The number, spatial extent and duration of harmful algal bloom events are not at levels that indicate adverse effects of nutrient enrichment. Information about species composition is also valuable for detecting changes in food web structure and harmful algae blooms.

Microscopic counting, but to a lesser extent taxonomical identification of phytoplankton, is replaceable by flowcytometric analysis and image flow cytometry. The improvement of flow-cytometers has occurred along with the development of laser technologies, and these instruments are now widely used in medical diagnostics. The technology, include pulse-shape recording flow cytometry that has shown promising results. This technology has demonstrated a potential in the analysis of phytoplankton community composition because of its rapid analysis time, that it can estimate particle volumes and because it has a counting precision comparable to microscopic counting (Haraguchi et al. 2017). PFCM data has already demonstrated that the technology provides reliable data with a time frequency as low as minutes. Image flow cytometry is conducted by a video camera that focuses on the fluid stream. Figure 9 show examples of automatic collected images. The images are afterwards post-processed by image analysis, and when combined with calibration information from previous processed samples, volume, cell numbers and some information of taxonomy is obtained (Sosik & Olson 2007, Lombard et al. 2019). Such flow cytometers are currently tested on Ferrybox lines.

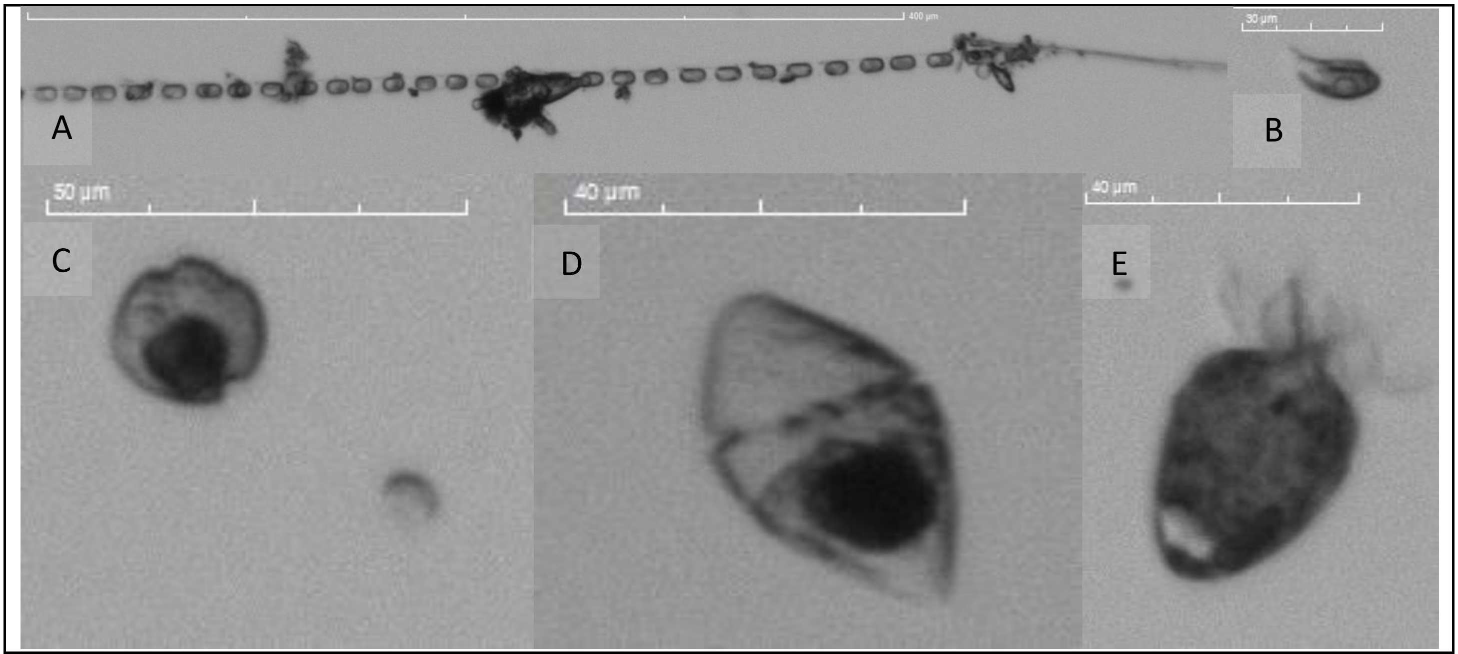


Fig. 9. Collage of plankton images collected by Cytosense (CytoBuoy, NL) using the optional image flow camera. A: the chain-forming diatom *Skeletonema marinoi*, B the cryptophyte *Teleaulax* sp., C: the dinoflagellate *Gymnodinium* sp. and D: An oligotrich ciliate.

*Conclusion: We recommend that direct counting of phytoplankton remains a part of eutrophication monitoring and that new technologies are developed based on flow cytometry and image analysis.*

### 3.2.4 Satellites observation of chlorophyll

*Retrieval of satellite data and operational estimation of a chlorophyll indicator* – The launching of the Sentinel programme means there is free and open access to satellite data. However, the correct algorithms for estimating chlorophyll for the certain water types is challenging, and the diversity of complication factors such as turbidity and CDOM makes it essential to use different algorithms adjusted to local conditions. This subject is treated in depth in Activity 2 and reported in Van der Zande et al. (2019).

### 3.2.5 Fluorescence data from ferrybox systems

Fluorometers, similar to those on a CTD, measuring chlorophyll fluorescence are used in ferrybox systems, for instance between Oslo and Kiel and Helsinki and Kiel. The fluorometer continuously measure the signal from an intake at the bow. The same circumstances apply as when such instruments are used on a CTD, i.e. that the signal is affected by cell size, the density of chlorophyll in the cell and thereby self-shading and hence the nutrient status of the cell. Furthermore, the previous light history affects the signal, i.e. how much light did the cells receive just before the measurement. The latter means there will be a systematic circadian rhythm in the ratio of the fluorescence signal to the chlorophyll concentration during day and night (this can to some extend be avoided by dark acclimatization of the sample). On a ferry route with one passing per day, this will transform into a fixed spatial variation.

The problem with the seasonal and circadian rhythms in the signal is considered on some routes (HELCOM 2019). The initial plan was to include ferrybox data from the Kattegat in the analysis. However, it was not possible to obtain a calibrated dataset during the project period, so we had to leave out this technique from the case study analysis in Kattegat in this project (see chapter 4).

*Conclusion: The use of ferrybox systems for chlorophyll monitoring depends on quality control of the dataset, in particular accounting for the problem with the circadian rhythm of the fluorescence signal:chlorophyll concentration ratio. We recommend that resources used to develop this technique further, are directed towards analysis of data before implementation of the technique on new ferry lines.*

## 3.3 Databases

### 3.3.1 Local quality control and storage in local databases

After collection and initial processing of samples, the results will be stored in local national or regional databases. These databases will often also contain environmental data covering other matrixes and national water covered by the Water Framework Directive. It is essential that the local institutes responsible for the practical planning and operation also perform the initial quality control.

#### *3.3.1.1 Metadata*

Metadata are all other data than the value of the variable – here chlorophyll concentration (µg $l^{-1}$), time, position and depth of sampling. Metadata may describe the analytical method, calibration data, gear for collection of samples, algorithms used for satellite observations, instrumentation on ferrybox systems etc. Such data are essential when the data are used later in the evaluation of environmental status and assessment of an area. For instance, comparison of different analytical methods can only be made if the database contains information on the methods used. The OSPAR JAMP Eutrophication Monitoring Guidelines: Chlorophyll a in Water (OSPAR, 2012) describe sampling and analytical procedures and allow for the use of different methods. Sometimes metadata about the actual methods are reported to ICES but this is not always working and ICES also have problems handling these data. Today, metadata are collected via personal contact, which takes time and often prevents an efficient analysis. The experience from the current project is that even in a project with dedicated partners, access to particularly metadata is a major obstacle that prevents efficient analysis. However, we have shown in this project that thorough understanding of the methods used is a prerequisite for comparing data from different sources and for a coherent assessment result.

Definition and structuring of metadata are not straightforward processes. For *in situ* water sampling, a code list for common analytical procedures (http://vocab.nerc.ac.uk or http://seadatanet.maris2.nl/v_bodc_vocab_v2/welcome.asp) is available. However, as the list contains information about a wide range of procedures for many different purposes, it might not be optimal. Instead a regional list for environmental monitoring in Europe may be more efficient, but can fairly easy be made as a sub list. Several EU-member states report to two regional sea conventions. For instance, Germany, Sweden and Denmark report to both OSPAR and HELCOM and similar reporting is done for areas around the Atlantic coast and in the Mediterranean. Thus, in order to ensure an efficient reporting procedure, a common EU-procedure is recommended.

#### *3.3.1.2 Significant digits and handling of data below detection limits*

Chlorophyll concentrations, as well as values of other parameters, are sometimes low or even close to zero. The handling of low values, in particular data below the detection limit, often differs among laboratories and varies between countries, but often only digits considered significant are reported. Values below the detection limit are therefore either not reported, reported as equal to the detection limit, as 50% of the detection limit or even as zero. The arguments are that insignificant digits in a result or values below the detection limits contain no information and might give a false impression of precision. Moreover, it is good scientific practice to only report significant digits. For commercial laboratories, legal issues may arise if they report values below the detection limit, as the reporting is part of their certification. However, all these arguments assume that a value, i.e. the concentration of chlorophyll in a water sample, is a final result. This is, however, not the case. *All values for environmental data are only intermediate results and should be regarded as such and reported without any filtering.* In contrast, the final result is, for instance, an assessment of environmental status, a trend in a time series or a similar 'high level' result, that form the basis for an overall judgement or conclusion about the environmental state. For chlorophyll, low values often occur in winter or in summer during periods with severe nutrient limitation. The latter is important for assessment of environmental status. If data are censored, i.e. low values are omitted, set to zero or a detection limit, changes may be introduced over time depending on analytical procedures or the laboratory used – and the conclusion about environmental status might be significantly affected or wrong. For instance, use of an improved analytical procedure with a lower detection limit will be reflected as a decrease in the seasonal mean value in the cases where values are reported as the detection limit or, the other way around, as an increase, if values below the detect limits are omitted or reported as zero.

A related problem is if only significant digits are reported, for instance as one decimal point for chlorophyll values in the unit µg $l^{-1}$ or mg $m^{-3}$. The calculation of ratios, for example between a value for a water sample and a value derived from satellite observation, will then be very uncertain. Experience tells us that four digits (regardless of the decimal point) ensure that all possible information is retained for concentrations, as it is rarely possible to measure any concentration with a higher accuracy (for salinity and temperature, more than four digits might be applicable).

*Conclusion: We recommend: 1) that local institutions perform the initial quality control and store data in a national database; 2) that data are not censored, i.e. storage of all values in the database with minimum four digits. This might involve changes or clarification at EU or national level of the legal framework for laboratories supplying environmental monitoring services; 3) that appropriate metadata on sampling, analytical methods and calibration are stored; 4) that the existing working group (MSFD CIS TG Data) explores pros and cons of a reporting format for metadata, to specify best practice for data handling and propose how this could be implemented; and 5) that EU and its member states ensure that existing environmental metadata are secured. Experience tells us that such information is lost over time along with institutional and personnel changes.*

### 3.3.2 International databases

#### *3.3.2.1 Hosting of database and data operation*

The fundamental requirement for an efficient assessment procedure is easy access to quality ensured data for an area including the actual values, in this case chlorophyll, metadata and other environmental data, for instance, on temperature and salinity. Thus, the data collected for monitoring purposes must be stored in a common database. As environmental issues, potential regulations and abatement measure can have severe consequences, it is essential that the procedure is open and transparent for all stakeholders. Thus, public and open access to the database is important.

Currently, data are stored in national databases and by the international organisations ICES and EMOD-net. Several issues must be addressed in connection with storage and access of data. These include: 1) access to data including the freedom to analyse and report results based on the stored data; 2) quality control of data submitted from local/national databases; 3) storage and access to metadata on methods etc., as described above; 4) secure and stable operation including the necessary funding and stability of personnel is critical.

ICES currently supplies this database service for HELCOM and has procedures for all the necessary steps, including the assessment of eutrophication. For satellite observation of chlorophyll in the HELCOM area, the data are first retrieved from the Copernicus Marine Environment Monitoring Service (CMEMS) and then transferred to a quality controlled chlorophyll product by the Finnish institute SYKE, which in turn submits chlorophyll growing season means per grid cell to ICES.

For OSPAR the role of ICES so far has been the collection of *in situ* eutrophication data from OSPAR countries. However, unlike HELCOM, the assessments were performed by the OSPAR countries themselves. OSPAR recently has tasked the ICES data centre to apply the HELCOM tool to OSPAR data, also adding chlorophyll products from satellite observation. The retrieval of standard satellite chlorophyll products and the application of the JMP EUNOSAT quality control procedure (see van der Zande, 2019) can be done either by one institute for the entire North Sea (cf. HELCOM approach) or by more institutions. The latter option would require additional efforts to ensure coherence but will most likely improve the quality over time as different approaches regularly will be tested against each other.

An important issue is quality control, which is currently done in an iterative way where the contracting parties submit their data. Subsequently, ICES performs a quality control and reports potential problems. These steps are repeated until no issues remain. The next level of quality control is in connection with assessments where the result for each area is analysed. Here, unexpected results may occur, which initiates a new scrutiny of the data and often resubmission of data due to missing or erroneous data. Such an iterative procedure seems to work well within HELCOM. Other options are that the service is hosted by an EU-institution or in a looser network such as EMODNET. The latter does not yet offer the necessary quality control and latter cannot be recommended at the moment. The set-up with ICES as the formal host may be advantageous for the collaboration with non-EU countries since this will keep the data and the service in an impartial institution.

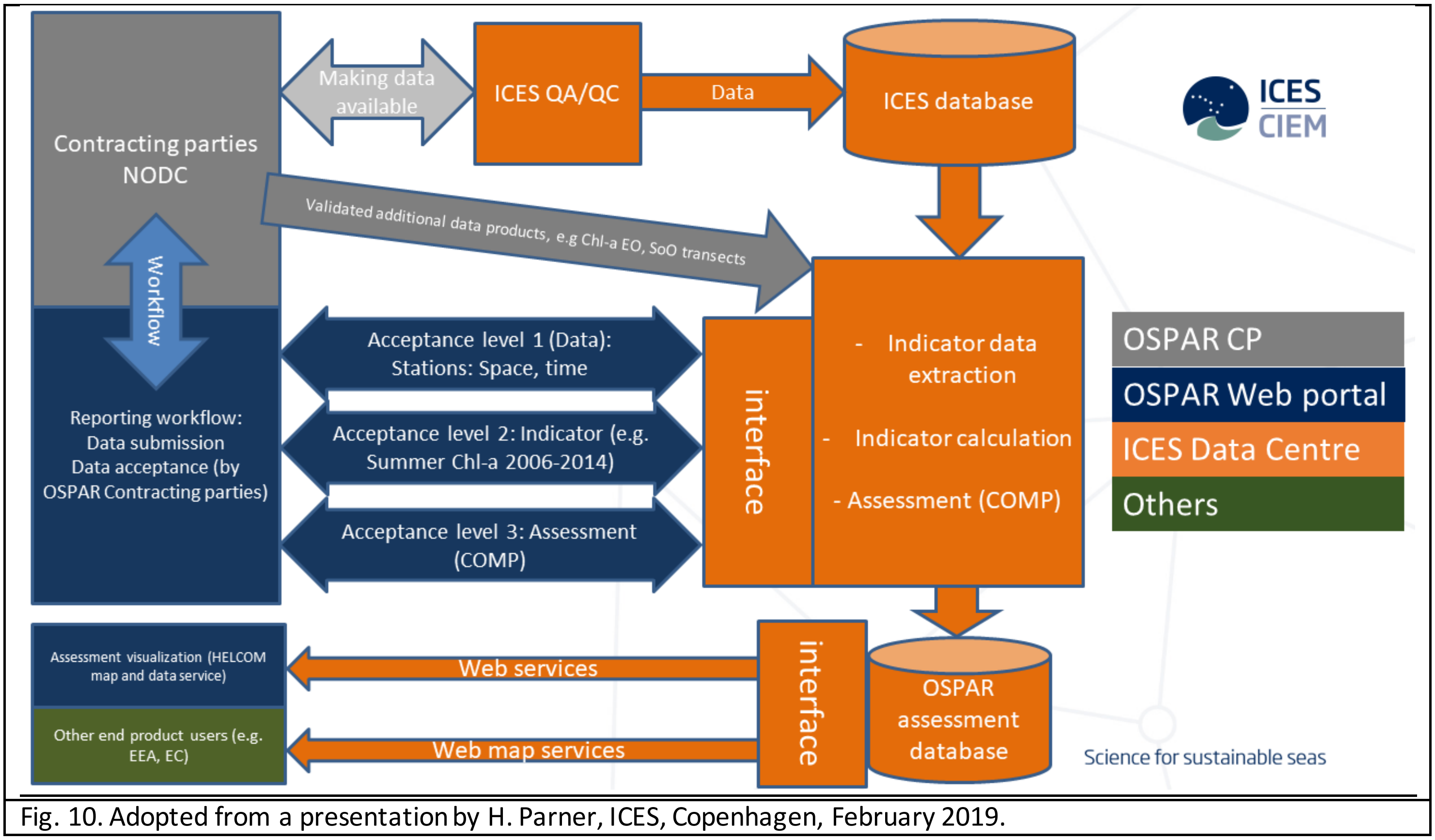


Fig. 10. Adopted from a presentation by H. Parner, ICES, Copenhagen, February 2019.

*Conclusions: We recommend that the service currently provided by ICES will be continued either by ICES itself or a similar institution. We strongly recommend that long-term funding is secured and that the efforts to include metadata and quality control are strengthened.*

### 3.4 Data use

The use of data involves a variety of processes as depicted in Fig. 6. The fundamental difference between data collection and data storage on one side and 'data use' on the other, is that the processes in 'data use' can be changed and redone with relatively little effort. In contrast, if raw data are not collected or analysed in the correct way, or data are not securely stored, the information is lost and cannot be redone. Thus, the procedures for data use are adjustable and the recommendations for changes can be tested, when necessary.

### 3.4.1. Calculation of indicators, (area, year)

Single values for chlorophyll are seldom used directly with the exception that chlorophyll is used as a scaling factor for processes like primary production. When chlorophyll is used as an indicator of eutrophication, the values are aggregated for a certain area and period. Often, the mean value over the growing season is used, given as an annual value for the area. Another common procedure is to calculate the 90% percentile (P90). The following issues must be considered when calculating chlorophyll indicator values: assessment area, season, log-transformation and statistical parameters.

#### *3.4.1.1 Assessment area*

Until now, the national boundaries have defined the eutrophication assessment areas, although within national boundaries smaller assessment areas have been identified based on for instance freshwater influence. National borders have no significance for eutrophication processes and may divide an area with similar conditions into two. In addition, assessment areas within national boundaries may still span across large ecological gradients. A major aim of this project was to define new assessment areas entirely based on ecological characteristics. The results are described in the report for Activity 1 (Blauw et al. 2019).

The optimal number of areas is a balance between the following considerations.

With many areas you can:

a. resolve gradients and can have similar eco-hydrodynamic properties;
b. facilitate detection of time trends when combined with sufficient sampling stations and sampling frequency to capture the spatial and temporal variability of the indicator;
c. provide more precise information on inputs of nutrients and the effectiveness of management measures;
d. threshold setting and performance of the assessment is more laborious;

Fewer areas has the advantages of

a. facilitate presentation of assessment outcomes on a wider geographic scale;
b. less laborious assessment;
c. less costly to ensure that *in situ* monitoring covers all areas, but note that detection of trends require sufficient samples to distinguish the trend from natural variation and random sampling error.

Within each assessment area a finer resolution grid can be applied, for instances based on satellite observation. In the case of gradients within an assessment area (e.g. along the Belgian and Dutch coast), seasonal means of chlorophyll can be calculated for each grid cell and even assessment thresholds can be calculated for each grid cell (see Activity 1 report).

#### *3.4.1.2 Season*

The main growing season for phytoplankton is approximately from February to October/November (Lyngsgaard et al. 2017). The seasonal distribution of chlorophyll show a peak already in February and high values well into November (see Fig. 11 with Kattegat as an example). The higher chlorophyll concentrations at the start and at the end of the growing season is due to low grazing by zooplankton, which allows a build-up of phytoplankton biomass. The grazing in the spring is reduced due to low temperatures and in the autumn due to predation by fish larvae (Maar et al. 2014). In wintertime growth is controlled by the light conditions and hence the combination of surface irradiance, light attenuation and stratification of the water column. An indicator reflecting nutrient availability should ideally cover the part of the growing season where phytoplankton growth is controlled by the availability of nutrients. Currently, this period is defined as from March to September by OSPAR (Northern regions) where HELCOM uses the period from May to September. May to September is also used in the Baltic region for the Water Framework Directive where Norway use February to October. An analysis of the problem (Carstensen et al. 2015) recommended the use of the entire year for the indicator and not the growing season. This has the clear advantages that an annual value is not so likely to be affected by the exact date for sampling in a particular year around the cut-off time for a seasonal defined indicator. However, if we assume that winter blooms are less affected by nutrient availabilities, compared to concentrations in the growing season, then an annual indicator is like to be affected by random variability in winter light conditions.

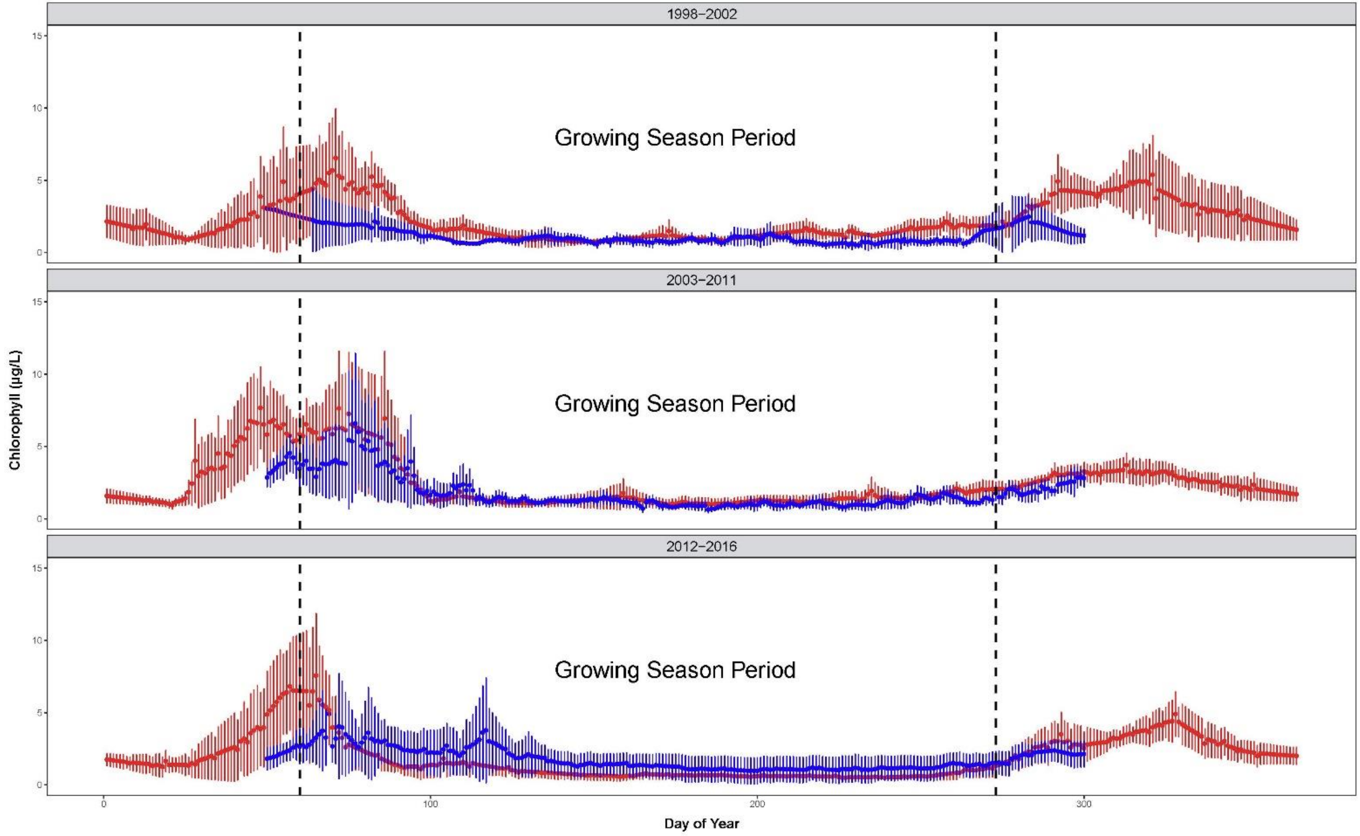


Fig. 11. Chlorophyll concentrations at a station in Kattegat (925, Gniben) for three periods based on *in situ* measurements (red) and satellite observations (blue). A growing season from 1. of March to 1. of October is indicated. The three periods correspond to the periods for availability of different satellites data (Van der Zande et al. 2019).

### 3.4.1.3 Log-transformation

Chlorophyll concentrations vary greatly over the season often with pronounced peaks in spring and autumn (Fig. 11), which means that the values often deviate from a normal distribution. Log-transformation can make up for this and provide normal-distributed data. However, log-transformation also lends more weight to lower than higher values, implying that the inter-annual variation in the indicator may reflect the variability in low rather than high values over the years. This is problematic since the nature of the chlorophyll analysis entails that a number of factors (degradation of chlorophyll, leaking during filtration, incomplete extraction) produce erroneously low values. As contamination is not possible, there is no systematic bias towards high values. Moreover, the extent of the seasonal peaks reflects the maximum availability of nutrients in the system, which is what we aim to describe with the indicator. Thus, use of log-transformation is not recommended.

#### *3.4.1.4 Statistical parameter*

The statistical parameters in use for the chlorophyll indicator are the mean or the 90% percentile (P90). HELCOM uses mean values where P90 and means values are used by OSPAR. A P90 value is sensitive to the frequency of sampling as the value decreases significantly if peak values are missed. This favours the use of a mean value over the period. However, the mean is sensitive to one or a few high values, particularly if the number of samples is low. On the other hand, the mean value represents the true impact of a high phytoplankton biomass on the ecosystem. We discussed the issue at the final workshop in February 2019 but were not able to reach an agreement. In our view, the use of 90% percentile for the indicator seems arbitrary and the risk of sensitivity to the frequency of sampling is high. Carstensen et al. (2015) showed that the P90 indicator is clearly more uncertain than an indicator based on mean values. We therefore recommend the use of mean values as the best statistical indicator. The use of median values was also discussed as an option.

*Conclusions*: *We strongly recommend that the ecological-based areas defined in Activity 1 (Blauw et al. 2019) are used in the future and that further subdivisions of local areas are discussed by neighbouring countries with local knowledge. For each area, we suggest that an annual indicator value is calculated as the mean over the period March-September on non-transformed values for chlorophyll concentrations. The period can then be subdivided for specific areas if phytoplankton growth is limited by for instance phosphorus in the spring and nitrogen later on. That will still allow the calculation on one indicator value for the entire period from March to September.*

#### *3.4.1.5 Combination of different data types*

This issue has been the focus of most of our efforts in Activity 3. It is therefore treated separately in the chapter below.

### 3.4.2 Estimation of GES-values from 1900 load and models (Activity I)

We do not have observations of chlorophyll concentrations from the period before human impacts on the biogeochemical cycling of nutrients became significant. We know from other sources that the surplus of nitrogen and phosphorus has increased with a factor of 6 to 8 over the last 120 years (Conley et al. 2007, Ellegaard et al. 2006, Kyllingsbæk 2008). Ideally, we should use chlorophyll data from before the year 1900 but such data are unfortunately not available. Therefore, relatively undisturbed conditions must be estimated indirectly from data for nutrient loadings nutrient concentrations and models that link nutrient inputs to chlorophyll concentrations. This is the aim of Activity 1 in this project (Blauw et al. 2019).

The basis for a GES-estimation is a data set on chlorophyll for each assessment area, covering a range of loading conditions. At present, the only existing data source is the data collected over the last decades with a variety of methods. The validity of an assessment comparing the current state with a GES-value hinges on comparability between the methods used in the modelling of year 1900 chlorophyll values, and the methods used today for the current assessment. It essential that the values are comparable, implying that an assessment based on this approach should include documentation for comparability.

### 3.4.3. Estimation of time series

A time series of an indicator, in this case chlorophyll, can be used to document the temporal trends of the environmental status of an area/region. It is often a critical component of the assessment supporting management decisions and therefore receives high political attention. Consistency in sampling approach and analytical procedures, as well as other components, over time is essential for a time series to show a valid pattern. This leads to what is sometimes called 'the tyranny of time series' as any change in methods will affect the time series. This can essentially hamper or preclude the introduction of new methods such as the use of satellite observations. The solution is to store the values in a database together with metadata and to ensure sufficient overlap when methods are changed. This allows for sound statistical analysis of the effects of changes in methods. This subject is treated in detail in chapter 4.

### 3.4.4 Periodic assessment (OSPAR, EU-MSFD)

OSPAR, HELCOM and EU all perform periodic assessments of environmental status with a frequency of six years. The main formal purpose of monitoring activities is to provide the basis for such assessments. Before, common practice was that each country did the assessment for their national areas and reported the outcome. This is often a closed process that does not ensure transparency and consistency. Furthermore, it is not suitable for ecologically based areas that cross national borders. It is therefore highly desirable to introduce a common and transparent procedure as the one described in Fig. 10 adopted from ICES, which is similar to the procedure applied in HELCOM. This allows all the technical aspects of the assessment to be run automatically. The responsibility of the organisations, for instance OSPAR, then is to establish the guidelines for the process.

*Conclusion: We recommend that EU and OSPAR adopt a procedure similar to the one used in HELCOM (outlined in Fig. 10).*

### 3.4.5 Estimation of maximum allowable inputs (MAI) of nutrients

Assuming that nutrient inputs are the main cause of eutrophication and an inadequate environmental status is observed for an area, the next step is to establish quantitative relationships between the value of an indicator and the nutrient input. Then, the values for MAI can be estimated by the use of empirical models (Dinesen et al. 2011, Hinsby et al. 2012, Erichsen et al. 2017, HELCOM 2007) or by mechanistic models as done under OSPAR's eutrophication modelling group (ICG-EMO). MAI values form the basis for political decisions on abatement measures.

*Conclusion: We recommend that quantitative relationships are established between each indicator and the nutrient input to each area.*

### 3.4.6 National use

The processes outlined above focus on the assessments done in international fora like OSPAR, EU and HELCOM. However, there might also be national agendas where each country sets environmental targets for an area. An important issue is then how much of the nutrient input to an estuary or a coastal area that originates from local sources and how much that is imported from adjacent marine areas (Timmerman et al 2010, Jørgensen et al. 2014). In addition, the open seas are the boundaries for coastal areas managed under the Water Framework Directive. The conditions in open areas, for instance the nutrient concentrations, affect the coastal ecosystems and *vice versa*. Hence, it is essential to have access to all data for open and coastal areas in order to optimise the management of all marine areas.

*Conclusion: We recommend storage of data in open and accessible databases allowing the adjacent countries to utilise all available data in their management of local coastal areas.*

# 4 On the use of satellites for estimating chlorophyll concentrations for environmental monitoring in optical complex water – a case study for Kattegat

## 4.1 Introduction

The combination of different observation methods for chlorophyll has significant advantages but also raises the question of how to combine the data into a common indicator. As mentioned above, chlorophyll is a proxy for phytoplankton biomass, but there is no constant relationship between chlorophyll and other proxies like cell number or carbon biomass. In addition, satellites measure light absorption by pigments where *in situ* sampling measures the concentration of one or more of the chlorophyll pigments. The source of the variation is related to factors such as species, cell size and phytoplankton physiology. From a monitoring perspective, it is essential that it is not just a random variation but systematic variations related to nutrient availability in the system – the factor we aim to monitor. Thus, we cannot assume a constant conversion between methods.

The effects of the observation method used also relate to the spatial and temporal restrictions that apply to the different monitoring platforms. Satellite observations are recommended to be used from mid-March to September in the region (see Fig. 11) and only surface concentrations are measured. Outside this period the sun angle is so low that the sensitive is reduced. Further south, the period for satellite observations is longer. A major advantage of satellite observations is their high spatial coverage. Until recently, the rather coarse resolution of satellite images limited their use in the nearshore, more shallow waters due to significant land-sea interaction. The temporal coverage is potentially high, but in some areas cloud cover limits the number of observations. A more detailed assessment of the suitability of satellite observations for eutrophication monitoring is provided in the Activity 2 report. *In situ* sampling allows monitoring down through the water column, but the spatial and temporal coverage is limited.

In practice, a simple averaging of all data means that only the satellite observations that determine the final indicator value due to the large spatial coverage and hence number of observations. The problem also applies to other automated techniques like ferrybox systems with a fluorometer, which records the signal from chlorophyll at the depth of the intake and along the line of the route. Thus, it is necessary to merge data from the two – or more – techniques into a common indicator in a way that ensures inclusion of all techniques.

In this section, we have developed a method for merging satellite observations with *in situ* sampling. The aim was to combine the exactness of the direct determination of chlorophyll in water samples with the high spatial coverage of satellite observations. The second objective was to test the correspondence between the two methods over long time periods (19 years). The Kattegat is a difficult area for satellite observations due to the relatively high concentrations of coloured dissolved organic matter (CDOM, Stedmon et al. 2000). Thus, the correspondence is better in less CDOM-influenced areas (Van der Zande et al. 2019). The analysis covers a period with three different satellites in use, permitting us to evaluate also the usefulness of the different satellites.

### 4.2 Data

The technique developed in Activity 2 was used to estimate chlorophyll concentrations for the OSPAR area. We compared the satellite values with values from ship-based *in situ* sampling at four stations in the Kattegat, including one station in the Great Belt just outside the OSPAR area.

The data set comprises four stations, including three Danish stations (Aarhus Bight, Great Belt and Gniben in the southern Kattegat) and one Swedish station (Anholt East). Aarhus Bight and Great Belt are coastal stations with land within 5-10 km, but with deep water, 17 and 35 m, respectively. Gniben and Anholt East are open water stations (Table 1).

Table 1. Growing season mean values (µg Chl $l^{-1}$ ± standard deviation) for the four stations and the two types of measurements.

| Station | | Ship-based *in situ* sampling | Satellite estimates | Difference | Relative deviation, percentage |
|---|---|---|---|---|---|
| Aarhus Bight (St170006) | Coastal | 2.45 ± 0.74 | 3.64 ± 2.38 | 1.18 | + 48 % |
| Great Belt (St6700053) | Coastal | 2.56 ± 0.83 | 3.00 ± 1.33 | 0.44 | + 17 % |
| Gniben (St925) | Open water | 1.76 ± 0.81 | 1.45 ± 0.72 | - 0.31 | - 18 % |
| Anholt East | Open water | 1.82 ± 0.62 | 1.21± 0.48 | - 0.61 | - 34 % |

| | | | | | |
|---|---|---|---|---|---|
| (St32002) | | | | | |
| **Mean** | | **2.15** | **2.32** | **- 0.17** | **- 3.43** |

Sampling was done at several depths about 20 to 40 times per year from 1998 to 2016, which are the years for which satellite estimates are available. Satellite estimates were averaged for grid cells varying from 1x1 km to 20x20 km. Figure 12 shows the estimate concentrations for three periods using data from the satellites SeaWifs, MERIS and MODIS, respectively.

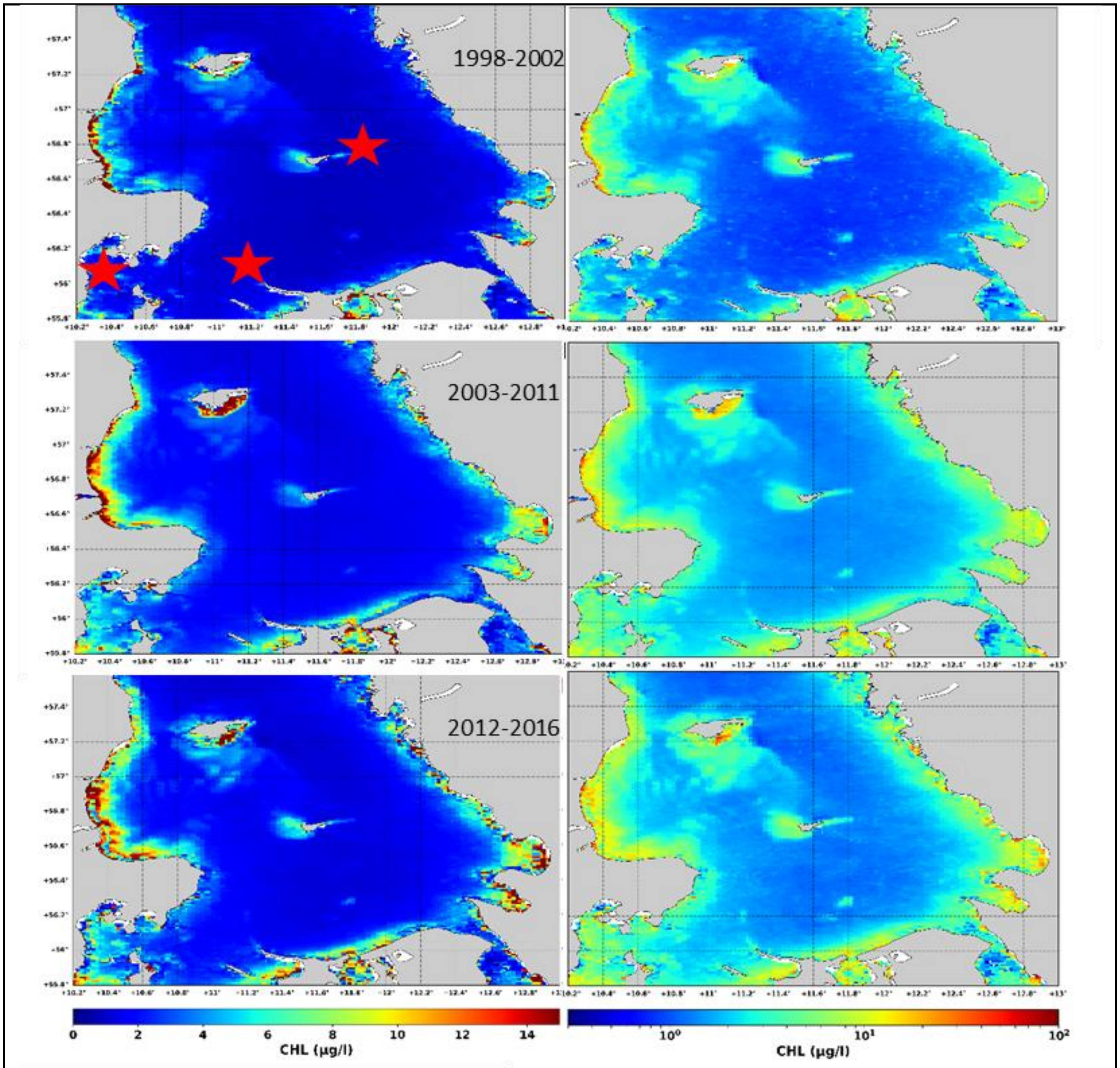


Fig. 12. Satellite estimates of chlorophyll concentrations (left: absolute values, right: log-transformed values) for three periods using data from three different satellites (see Van der Zande et al. 2019 for details). Three of the four sampling stations are shown on the upper left map. The fours station is placed in The Great Belt just South of the border of the maps.

## 4.3 Results and Discussion

### 4.3.1 What does the satellite see?

The initial analysis showed that the best match between the two techniques was obtained with *in situ* values from 1 metre. Comparisons with measurements from deeper depths gave a poorer relationship. Hence, we only used surface (1 m) values in the remaining analysis.

### 4.3.2 Spatial variability

Our initial hypothesis was that a close correspondence in time and space would give the best match between satellite and *in situ* observations. Thus, our initial analysis was based on a grid cell size of 3x3 km and a true matchup (within a day) between the individual sample points. This gave an $r^2$-value of about 0.6 for station 925 (Gniben); however, for the other stations, very few data were available for a true matchup/validation.

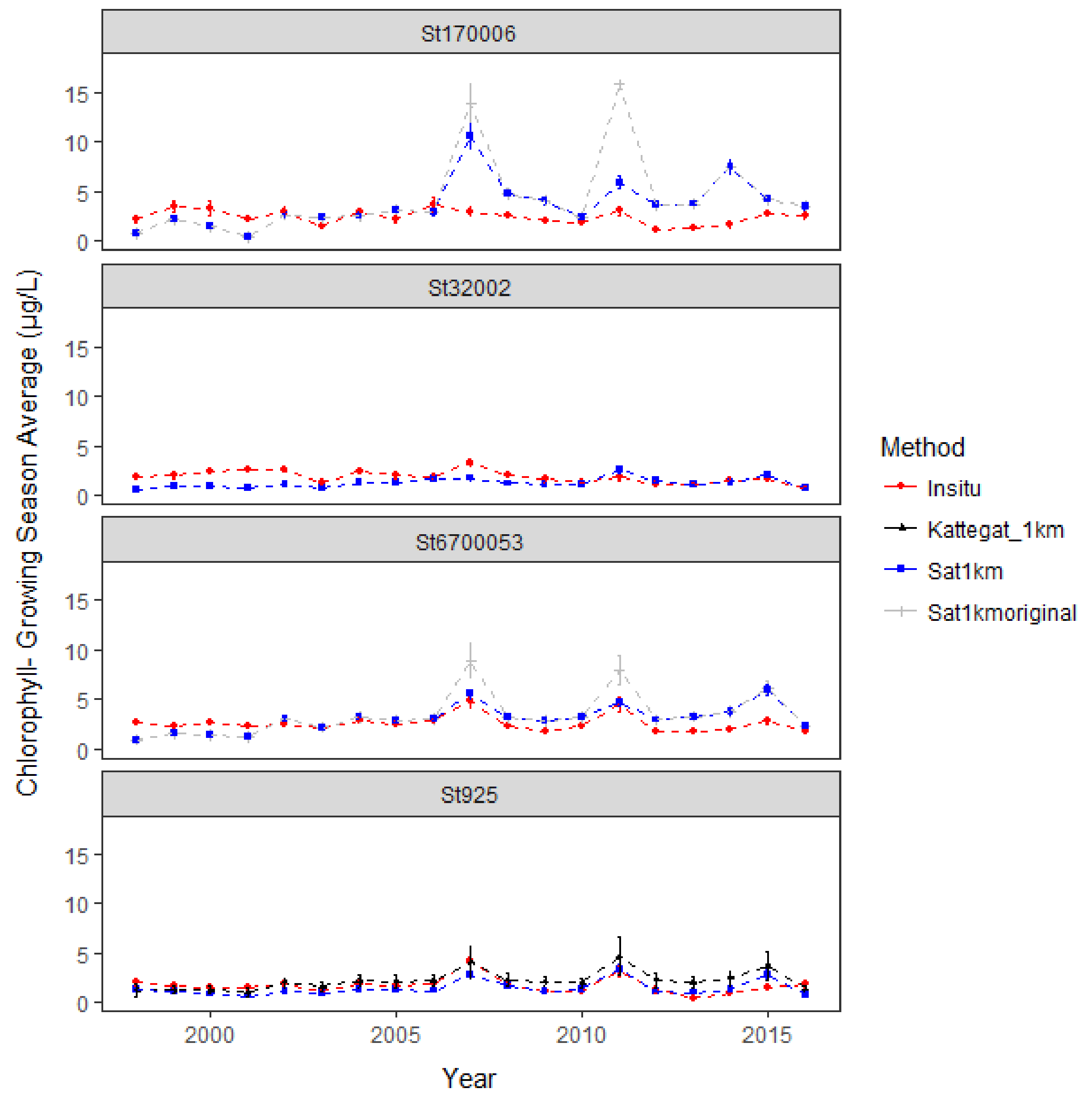


Fig. 13. Growing season mean values for four stations in the Kattegat–Belt Sea area. Red lines are estimates based on *in situ* values. Grey lines are satellite estimates (Van der Zande et al. 2019) for 1x1 km grid cells. In coastal areas, within ca. 5 km from the coast, a number of erroneously high values occur (up to 6 standard deviations above the highest recorded *in situ* value over 19 years with about 600 observations in total). The blue line is corrected satellite-estimated values where values above one standard deviation for the month are removed. The black line in the lower panel is a time series based on 1x1 km grid cells covering the entire Kattegat outside 5 km from the coast.

Based on the results of our analysis, we hypothesised that there is a high degree of random variability in the satellite estimates and to some extent in the temporal and spatial distribution of *in* situ observations of chlorophyll as well, which introduces noise in the relationship between the two techniques when comparing individual observations. However, assuming that this is random variation, much of this variability will disappear with the use of average values. Moreover, keeping the focus on monitoring and management, our primary interest is annual values for a chlorophyll indicator for the basin.

Figure 13 shows growing season average values for the four stations. Both the time series for *in situ* and satellite estimates show several common features, for instance high values in 2007, 2011 and 2015. These values are due to high runoff in the months before the growing season. Based on this, we conclude that despite spatial variation, there is a common pattern in inter-annual variability across the basin using the two techniques.

Our conclusion was further tested by comparing time series for the four stations based on satellite estimates using 1, 5 and 20 km grid cells, all including the position of the stations (Fig. 14). We found a high degree of correspondence between these and almost the same correspondence with the time series based on *in situ* values. Thus, we conclude that grid cell size is not important for the inter-annual variability and hence not for the matchup between the two techniques in the open parts of the Kattegat.

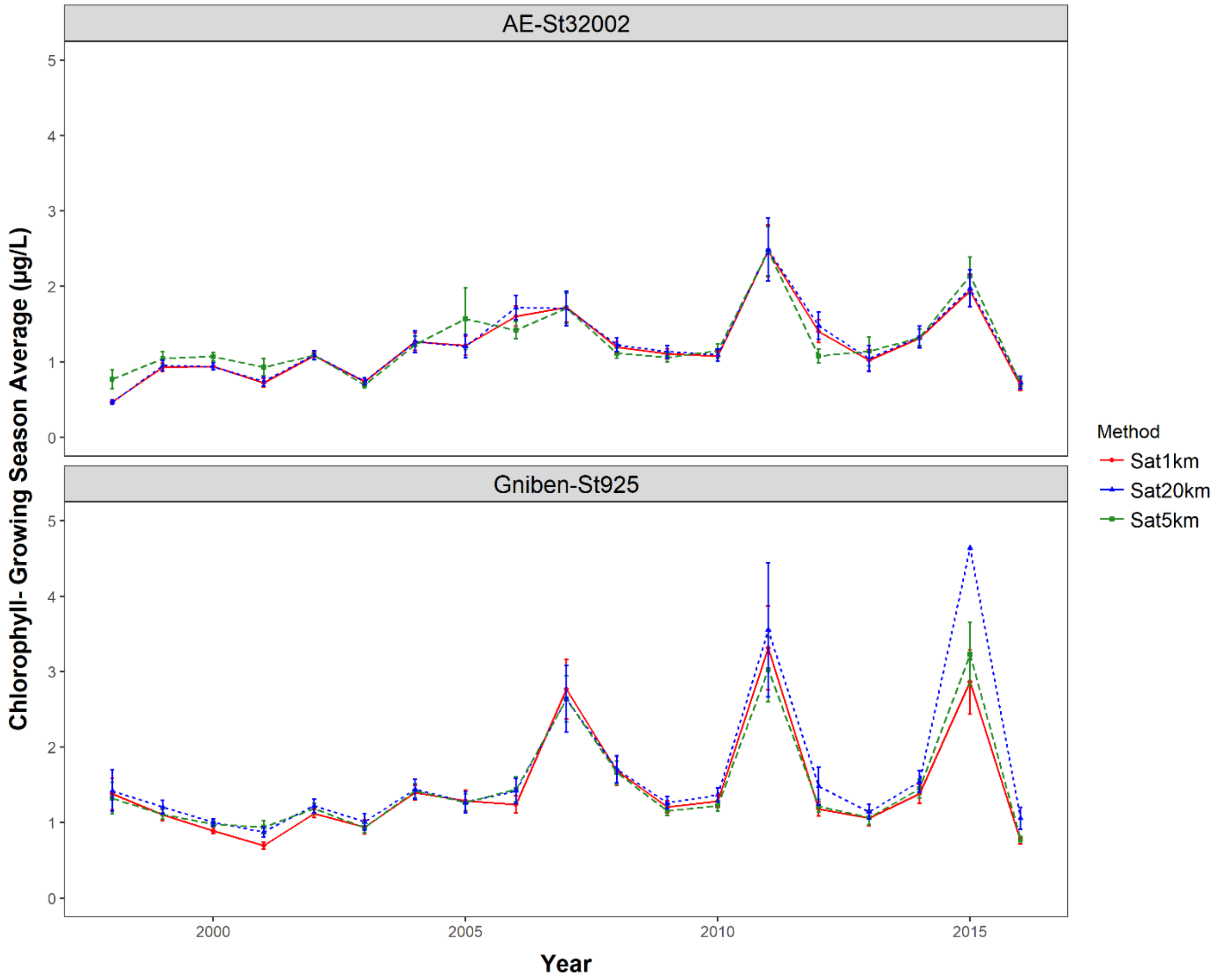


Fig. 14. Growing season mean values estimated using three different grid cells (1 km, 5 km and 20 km) from satellites for two stations in the Kattegat – St32002 (Anholt East) and St925 (Gniben).

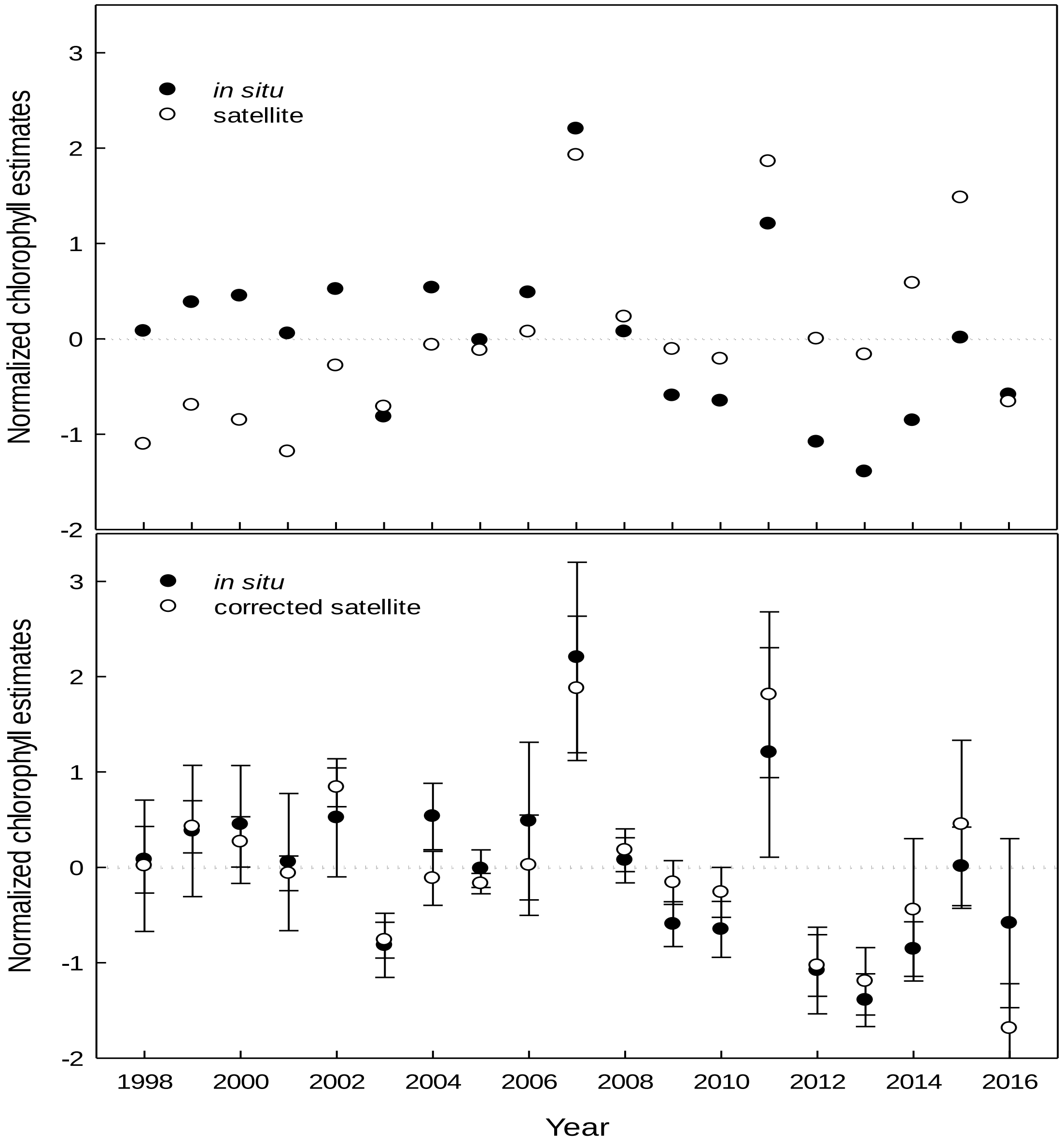


Fig. 15. A) Normalised (mean=zero, standard deviation=1) for the chlorophyll concentration from four stations in the Kattegat and Belt Sea area from 1998 to 2016 (error bars are omitted for clarity). B) As for the panel A, the satellite estimates are corrected according to Eq. 1 and error bars (standard deviation) are shown.

Spatial variability in surface chlorophyll values occurs, primarily as gradients with high values close to the coast, in the vicinity of large freshwater sources at the permanent front zone in the North close to the adjacent Skagerrak. However, at least for the area analysed here, there is a common pattern in the inter-annual variability across the area using the two techniques.

### 4.3.3 Analysing time series of mean values for growing season

Based on the analysis and conclusion above, we calculated normalised time series for the four stations and the two techniques with a mean of zero and a standard deviation of one. The result is shown in Fig. 15.

Figure 16 shows the relationship between the techniques.

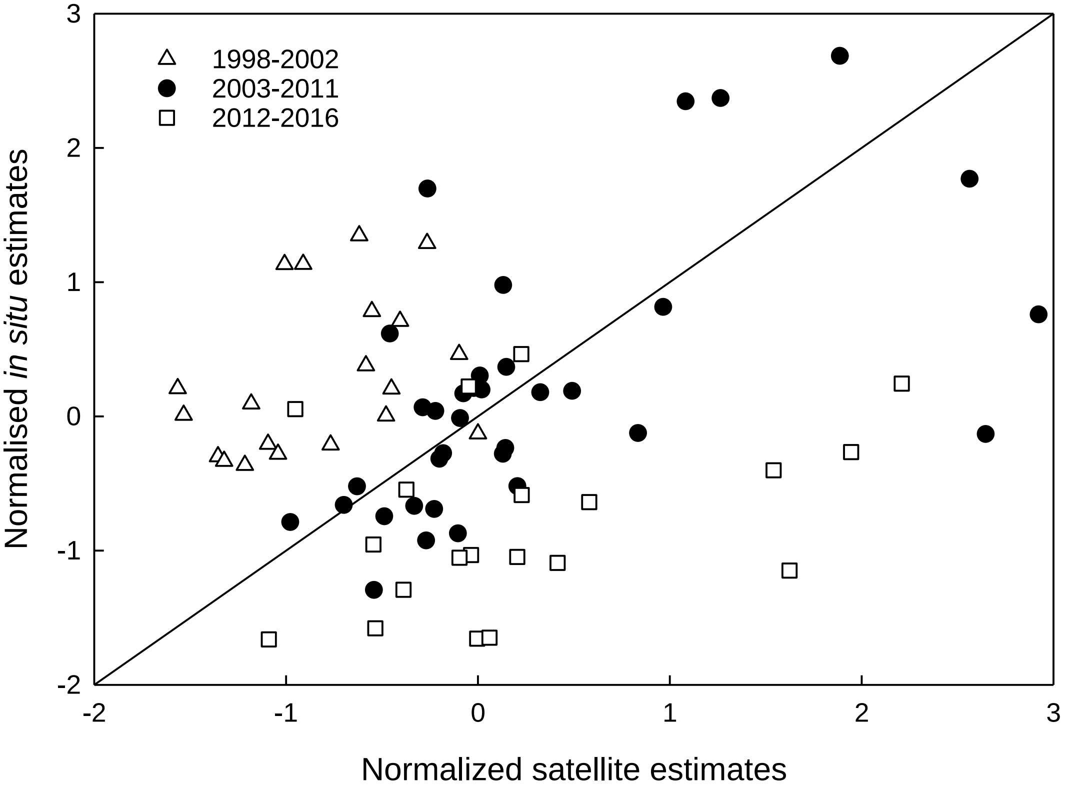


Fig. 16. Normalised *in situ* versus satellite values for all years and stations. Values are normalised to zero mean and one standard deviation of one by subtracting the mean values for the station across years for each type (satellite or *in situ* observation) and dividing by the standard deviation ($x' = (x - xmean)/sdev$).The 1:1 line is shown.

Figure 16 shows a systematic deviation between the techniques – the satellite estimates being higher than the *in situ* values from 1998 to 2002 and lower than the values from 2012 to 2016.

Based on the normalised values it is possible to establish an equation correcting for the satellite observation:

Corrected satellite obs = 0.05 + 0.99 x raw satellite obs. + bias (eq. 1)

where the bias is equal to 1.17 (1998-2002) and -0.98 (2012-2016). The result of this correction is shown in Fig. 15B.

The systematic deviations of satellite estimates compared with *in situ* data have important implications for the detected trend in chlorophyll concentrations and hence for management. Figure 15a shows the change over the years for the normalised values for growing season mean values. The values based on *in situ* sampling are positive (above the mean) from 1998 to 2008 with 2003 as an exception. From 2008 to 2013, there is a decreasing trend with 2007 and 2011 as outliers. Thereafter, values increase again. Overall, there is a negative trend of -0.06 standard deviations per year (p=0.006). The inter-annual variation is closely linked to the nitrogen loading to the area, and a similar pattern over the years also emerged in a much larger data set for the same area. This is a key observation within a management context as the decreasing trend in chlorophyll is the result of an approximately 50% reduction in the nitrogen loading to the area.

In contrast, the raw satellite estimates gave the opposite exhibiting at overall positive trend ($p<0.0001$) due to a negative bias at the beginning of the period and a positive bias at the end. These changes in bias correspond to the satellite data available, where SeaWifs was used from 1998 to 2002 (negative bias), MERIS between 2003 and 2012 (no bias) and MODIS from 2012 and onwards. The better performance of the MERIS-based data might be because of the satellite's 709 nm band that allows us to use FUB-WEW Neural Network analysis (see Van der Zande et al. 2019). Our hypothesis is that the wrong trend in the satellite estimates is because of the CDOM-rich waters of the Kattegat. In the North Sea, the same methods gave better results for the entire period (Van der Zande et al. 2019).

The Kattegat case study shows that monitoring based only on satellite observations would not yield the true pattern in these CDOM rich waters. Therefore, it is essential that satellite observations are compared with *in situ* measurements and, if necessary, corrected. The optimal approach is to improve the basic algorithms for satellite data so that post corrections, as done here, become unnecessary. With the launch of the Sentinel satellites as from 2016, we can expect a higher quality of satellite data, including a 709 nm band, potentially promising betting matching of satellite estimates with *in situ* data in the future.

### 4.3.4 The use of satellite observations

Based on the analyses above, we have calculated a time series for the offshore parts of the entire Kattegat area using 1x1 km grid cells (Fig. 13, lower panel - black dotted lines). This raw time series thus utilised all satellite observations over the time span and area. Such a time series will, in our opinion, combine the better *in situ* technique, i.e. precision, with the strength of satellite observations, i.e. higher spatial and temporal coverage.

*Conclusions:*

1. *Satellite observations reflect the chlorophyll concentration only at the surface. The average concentration of the mixed layer is only partly reflected and satellite observations do not allow determination of the deep chlorophyll maximum. However, the latter is not as important in a management perspective as eutrophication is reflected by the chlorophyll concentration in the mixed layer. It is a problem, though, that not even the mixed layer concentration is well represented by satellite estimates.*

2. *There is a random component in satellite observations in the Baltic Sea transition zone (the Kattegat region), probably due to its complex optical conditions created by the high CDOM concentrations, which hampers the use of ocean colour algorithms, thus increasing the error of satellite estimates of chlorophyll. However, for a seasonal average, the spatial variation is small for the time series, i.e. the different parts of the basin display the same inter-annual variability.*

3. *There are significant systematic errors in satellite-estimated chlorophyll concentrations in the 19-year time series analysed, especially outside the MERIS period (2003-2011). The MERIS and Sentinel-3 data allow for the use of neural networks (e.g. FUB-WEW), which are well suited for the complex Baltic waters. If such data are not available (e.g. SeaWifs, MODIS), a correction procedure can be applied if* in situ *observations are available for an area, rendering the satellite measurements useful. Therefore, in the current state, satellite observations cannot be used for the Baltic transition zone for monitoring or management purposes without validation with* in situ *observations. In the future, when a long standing and coherent satellite programme has been in operation for a period of, for instance, 10-20 years, sufficient evidence might have been gathered to permit the use of satellites as a standalone technique.*

4. *Based on the observations from coastal stations, satellites seem to have problems with measuring chlorophyll closer to the coast than ca. 5 km. Further limitations may apply for shallow areas or areas with high turbidity.*

# 5 Perspectives

The focus for this report is the current situation of chlorophyll monitoring and assessment and together with the reports from the other two Activities in the project (Blauw et al. 2019, Van der Zande et al. 2019) it gives a number of suggestions and recommendations about how to increase the coherence of monitoring and assessment for the North Sea region. In this final section we will give a longer term perspective, beyond the next couple of years. In addition, we look at the procedures in HELCOM. HELCOM covers the Baltic Sea and three countries, Sweden, Germany and Denmark, participate in both OSPAR and HELCOM. These countries will have an interest in similar procedures for monitoring, analytical methods, reporting of data to international data bases and the assessment procedure. Moreover, the Kattegat is covered by both organisations. Finally is it logic to learn from each other and copy procedures that have been shown to work well. Thus, there are a number of arguments for aiming a similar development for the two areas on the technical aspects.

## 5.1 Comparison with HELCOM procedures

In HELCOM the ship based sampling is done by national owned ships, although countries in some cases hire ships from each other when they have problems with their own ships or equipment. A ferrybox line is active between Helsinki and Kiel driven by Finland and satellite observations are used by a number of countries where particular Finland and Poland are active and promote the use of satellite observations.

For the analytical procedure, the HELCOM guidelines prescribe the use of ethanol as solvent for extraction. The concentration in the extract can then be measured either with a spectrophotometer or a fluorometer (HELCOM 2019). As also suggested in this report, there is freedom to use different methods (but only with ethanol as solvent in HELCOM). If a country use a method that deviate from the standard guidelines, it is the responsibility of the laboratory/contracting party to validate the results in a comparison with the standard method.

Ferrybox solutions are used. However, it is emphasised by HELCOM that changes in the ratio $\mathrm{Flu}_{\mathrm{Chl}}$:Chl must be fully accounted for before the results can be used for assessment, e.g. according to Babin (2008).

Satellite observations are used to assess chlorophyll concentrations. Due to the high concentration of coloured dissolved organic matter (CDOM), HELCOM recommends the use of the Case 2 regional Coast Colour algorithm and neural networks (Doerffer and Schiller 2007, Brockmann et al. 2016). This subject is treated in depth in the Activity 2 report in sections 2.2 and 2.3 (van der Zande et al., 2019).

## 5.2 Merging of data from different sources into one indicator for chlorophyll

As described above (chapter 4) combining different techniques is a major challenge. Here, two issues need to be distinguished.

### *5.2.1* Spatial integration

HELCOM has adopted an approach where data are reported for 20x20 km grid cells. As shown above (section 4.3.2, Fig. 13) this will work well in open areas where conditions are rather uniform and the procedure will probably work well and serve the needs for a coherent assessment procedure. However, the technique does not exploit the full potential of the spatial resolution of satellite and ferrybox observation. In a longer perspective, we foresee that the use of dynamic coupled 3D-hydrodynamic ecological models will be solution. These models can perform data assimilation and interpolation in time and space of data with different density. Such models are widely used for research in the Baltic Sea and in the North Sea and can be adapted to do assessments with some further development.

### 5.2.2 Integration of different data sources

Within a grid cell time series of annual means of chlorophyll concentration, obtained by different techniques can be shown as absolute values (Fig. 13) or normalised values (Fig. 14 and section 4.3.3). The latter was used in an approach to correct historic satellite data obtained by satellite sensors that did not perform well in CDOM rich waters. For each grid cell a time series of the chlorophyll indicator can be produced, and if *in situ* data are available for the grid cell, the time series can be aligned with the procedure described in section 4.3.3. The time series for a basin can then be estimated. In the future, with improved sensors and better algorithms, we expect the satellites data to correspond closely to *in* situ values. The technique for validation of time series should still be applied as a validation, but ideally it will a show a complete march. By combining time series of grid cells on a basin scale we obtain an indicator value and a time series with annual values as shown in Fig. 13 (lower panel).

Currently, HELCOM use at weighting approach for combing data from different techniques, at the moment *in situ* data, satellite observation and ferrybox data. Within each grid cell, seasonal means of chlorophyll from the different techniques are given weights (for example: factor 0.45 for satellite data and 0.55 for *in situ* data) before calculating the chlorophyll seasonal mean. Discussions in HELCOM suggest that the rather arbitrary weighting factors should be improved, preferably on the basis of scientific considerations such as statistical confidence in the data. In our view, this is not an optimal solution, as you cannot solve a discrepancy between measurements, which is supposed to measure the same quantity, by a weighting. Therefore, we have suggested the solution outlined in section 4.3.3., which align the time series to the measured *in situ* values for grid cells where both data types are available, as we regard the *in situ* technique to be the most accurate. Again, as mentioned above, the ideal situation is that the different techniques gives the same result for the annual indicator value. In that case, no correction is done, but it is then proven that the satellite estimates are correct. After a validation, and correction if necessary, the satellite data can by applied for all grid cells and hence on a basin scale, and the superior spatial coverage by satellite data are fully exploited.

## 5.3 Ecological coherent areas, GES and scaling of variability

We see the development of ecological coherent areas and related coherent thresholds (or GES values) for chlorophyll a and nutrients as a major step forward (Activity 1, Blauw et al. 2019). The general approach has in principle been approved in OSPAR, and application will be trialled for the next 2022 eutrophication assessment by a dedicated group (OSPAR, 2019). This will include testing of the modelled outcomes (in particular the background and threshold concentrations for nutrients and chlorophyll a and the assessment areas), comparison with curre nt OSPAR and WFD assessments and improvement of the approach where needed. We foresee that this will be an iterative process where (North Sea) countries gradually approach each other and agree about GES-values for areas that are shared among several countries and where all contracting parties gradually agree on the optimal procedure.

A problem related to GES is the inter-annual variability in an indicator. This is not an issue when only one threshold value for acceptable/non acceptable conditions is used, *cf*. MSFD. However, with a scale for status with five classes, as in the Water Framework Directive, it is essential to know the natural variability of the indicator and this has to be estimated for each area. For instance will the natural variability for the chlorophyll concentration be much higher in an estuary compared to open waters. This issue is debated in HELCOM, but so far without a consensus. We foresee that this also needs to be addressed for the North Sea. The seasonal variability (climatology) is used in Blauw et al. (2019) to define coherent areas. However, for scaling of the environmental status into five classes relative to a reference condition, we need to know the natural inter-annual variability under reference conditions. This is a major challenge that has not been resolved yet.

### 5.4 Coherency in monitoring and assessment

Regarding the organisation of the ship-based sampling, we hope that the countries bordering the North Sea will see the potential in working together both among the countries and when it comes to integrating monitoring of different parameters and purposes of monitoring and integrating monitoring with other marine activities. However, this involved many issues of economic, political and organisational nature.

Retrieval of satellite data is rapidly becoming easier and cheaper and is now done by many institutions. Activity 2 has developed a formal procedure for quality assurance and selection of the optimal algorithms for different marine areas in the North Sea (Van der Zande at al. 2019). We see this as a major step forward and we foresee that this will proceed in the future so the diversity of the algorithms used will increase and be better adjusted to local conditions. We foresee that all countries will participate in this work and have their own national capacity. This may lead to diverging methods and less coherence when compared to the approach used in JMP EUNOSAT and also in HELCOM, *i.e*. one institute that collects and quality assures satellite data. Hence, depending on future developments a coordinating body, maybe through CMEMS, should keep an eye on coherence between algorithms and methods related to satellite observation.

The results in activity 2 show that satellite observation works well for open areas. These are also the areas where *in situ* sampling is most expensive, so we suggest that satellite observations are adopted for open areas and then gradually utilised in coastal areas where conditions are more complex. However, it is essential to maintain *in situ* based sampling, but this should be adjusted to work together with satellite observations in a joint monitoring design. Also *in situ* sampling programmes and ferryboxes can contribute to a joint match up database by sampling surface water at satellite overpass.

When the use of satellite observations become widely used and with several players, it is essential to document the methods. This is also stressed by HELCOM (draft guidelines for monitoring of chlorophyll *a*, annex A). Therefore, all algorithms and procedures used for estimating chlorophyll (including the basic scripts) must be open source and stored in public libraries and information about the methods must be supplied as metadata along with the results when reported to international data bases.

### 5.5 Outcome of the project

Figures 2 shows the current situation (duplicated in Fig. 17a) for monitoring and assessment of chlorophyll for the North Sea. Implementation of many of the recommendations will in our judgement bring us in the situation depicted in fig. 17b. The process associated with data storage, management and analysis will be coherent among the countries and also transparent. In addition, satellite data will be an integrated part of chlorophyll assessment. There is still room for improvements, for instance a closer operational collaboration at sea and also the use and further development of other techniques like ferrybox systems and in a longer perspective the use of other platforms like gliders and other automated underwater vehicles.

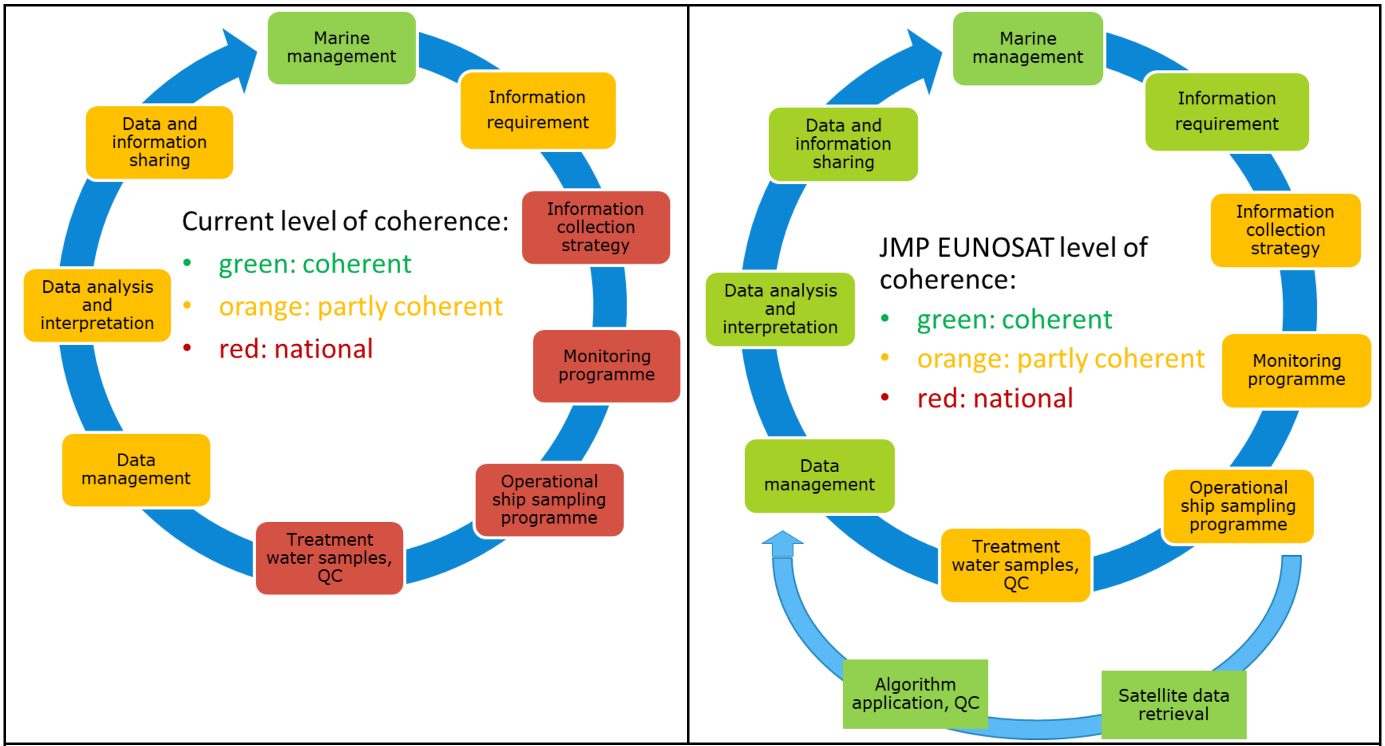


Fig. 17. The JMP EUNOSAT project has addressed all steps in the information cycle and proposes options to improve coherence. The left panel shows the current level of coherence among North Sea countries with regard to *in situ* monitoring and assessment of chlorophyll. The right panel estimates the situation in the (near) future where all JMP EUNOSAT recommendations have been implemented. It also shows the added value of satellite observations (lowest parallel cycle), which are coherent by nature.

## 5.6 Final remarks

Monitoring and assessment of the environment is a complex activity with a diversity of issues. For marine areas the inherent trans-national aspect adds to the complexity. We hope the JMP EUNOSAT project has contributed to move the subject a little bit forward. During the process we have realised that there is seldom one solution that fits all needs and also that we are far from having all the necessary knowledge to make the right decisions. Under such conditions it is important to be flexible and to adopt new solutions even though they are not perfect. When it comes to the physical sampling and storage of data, which cannot be redone, it is important to be cautious and only adopt well proven solutions. When it comes to data analysis and assessment based on stored data, activities can be repeated, and we should be open for new solutions.

# 6 References

Almroth E. and M. D. Skogen (2010). A North Sea and Baltic Sea model ensemble eutrophication assessment. Ambio, 39(1), 59–69. https://doi.org/10.1007/s13280-009-0006-7

Babin M. (2008). Phytoplankton fluorescence: theory, current literature and in situ measurement. In: Babin M, Roesler CS, Cullen JJ (eds.) Real-time coastal observing systems for marine ecosystem dynamics and harmful algal blooms: Theory, instrumentation and modeling. Unesco, Paris, p. 237-280.

Baretta-Bekker H., A. Sell, F. Marco-Rius, J. Wischnewski, P. Walsham, Lynda, M. Mohlin, K. Wesslander, H. Ruiter, F. Gohin, L. Enserink (2015). The chlorophyll case study in the JMP NS/CS project. Document produced as part of the EU project: 'Towards joint Monitoring for the North Sea and Celtic Sea' (Ref: ENV/PP 2012/SEA). 72 pp.

Birchenough, S.N.R., Maes, T., Malcolm, S., and Enserink, L. (2015). Opportunities and Barriers for Joint Monitoring - Outcomes of the workshops organised as part of the EU project: 'Towards joint Monitoring for the North Sea and Celtic Sea' (Ref: PP/ENV D2/SEA 2012). Activity D Report. 51 pp.

Blauw, A., M. Eleveld, T. Prins, F. Zijl, J. Groenenboom, G. Winter, L. Kramer, T. Troost, A. Bartosova, J. Johansson, R. Capell, K. Eiola, A., G. Tilstone3, P. E. Land, V. Martinez-Vicente, S. Pardo, D. van der Zande (2019). Coherence in assessment framework of chlorophyll a and nutrients as part of the EU project 'Joint monitoring programme of the eutrophication of the North Sea with satellite data' (Ref: DG ENV/MSFD Second Cycle/2016). Activity 1 Report. 84 pp.

Boesch D.F. (2002). Challenges and opportunities for science in reducing nutrient over-enrichment of coastal ecosystems. Estuaries 25: 886– 900.

Borja, A., M. Elliott, J.H. Andersen, J. Carstensen, J.G. Ferreira, A.S. Heiskanen, J.C. Marques, J. Neto, H. Teixeira, M.C. Uyarra, L. Uusitalo, & N. Zam- poukas (2013). Good environmental status of marine ecosystems: what is it and how do we know when we have attained it? Mar. Pollut. Bull. 76, 16–27.

Borja, A., T.C. Prins, N. Simboura J.H. Andersen, T. Berg, J.C. Marques, J.M. Neto N. Papadopoulou, J. Reker, H. Teixeira & L. Uusitalo (2014) Tales from a thousand and one ways to integrate marine ecosystem components when assessing the environmental status. Front. Mar. Sci., https://doi.org/10.3389/fmars.2014.00072

Brockmann, C., S. Dury, G. Hesselmanns, R. Duin, M. Costello, R. Jordans, and K. Stelzer. (2004). Operational coastal water quality monitoring: are space borne products an alternative to in- situ measurements - Where are we now? In Proceedings MERIS User Workshop. Vol. SP-549. European Space Agency.

Brockmann C., R. Doerffer, M. Peters, K. Stelzer, S. Embacher & A. Ruescas (2016). Evolution of the C2RCC neural network for sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. In: Proceedings of the ESA Living Planet. ESA, Prague, Czech Republic, 6 pp.

Brockmann Geomatics AB (2017) Satellitbaserad statusklassificering. Web page: Accessed 19 Dec 2017 http://gisportal.brockmann -geomatics.se/webapp/wq/wq_wfd/?type=ch

Carstensen, J., A-T. Skjevik & L.J. Naustvoll (2015). Assessment method for chlorophyll *a*: Developing a common assessment method for chlorophyll *a* between Denmark, Norway and Sweden. https://pure.au.dk/portal/files/94152950/Assessment_method_for_chlorophyll_a_2015.pdf

Cleveland L.S. and A. D. Weidemann (1993). Quantifying absorption by aquatic particles: a multiple scattering correction for glass-fiber filters. Limnology and Oceanography 38, 1321? 1327.

Cloern J.E. (2001). Our evolving conceptual model of the coastal eutrophication problem. Marine Ecology-Progress Series 210: 223–253

Cloern J. E. (2018). Why large cells dominate estuarine phytoplankton. Limnology and Oceanography, 63(S1), S392-S409.

Conley D.J. (2000). Biochemical nutrient cycles and nutrient management strategies. Hydrobiologia 410: 87–96.

Conley D. J., J. Carstensen, G. Ærtebjerg, P.B. Christensen, T. Dalsgaard, J. L. S. Hansen and A. B. Josefson (2007). Long-term changes and impacts of hypoxia in Danish coastal waters. Ecological Applications, 17(5 SUPPL.), 165–184. https://doi.org/10.1890/05-0766.1

De Boois I. J. and R. van Hal (2015). Towards a joint monitoring programme for the North Sea and the Celtic Sea: Activity C Multidisciplinary monitoring. Document produced as part of the EU project: 'Towards joint Monitoring for the North Sea and Celtic Sea' (Ref: ENV/PP 2012/SEA). 58 pp.

Doerffer R., H. Schiller (2007) The MERIS Case 2 water algorithm. International Journal of Remote 28:517–535.

Dinesen G.E., K. Timmermann, E. Roth, S. Markager, L. Ravn-Jonsen, M. Hjorth, M. Holmer and J. Støttrup (2011). Mussel Production and Water Framework Directive Targets in the Limfjord, Denmark: an Integrated Assessment for Use in System-Based Management, Ecology and Society 16(4). doi: 10.5751/ES-04259-160426.

EFARO (2017). Synergies in CFP and MSFD monitoring strategies. Marine Directors meeting document MD2017_1-5. https://circabc.europa.eu/d/a/workspace/SpacesStore/84176070-3e02-416b-b011-dc666dc8cc9a/MD2017-1-5_EFARO.pdf

Ellegaard M., A. L. Clarke, N. Reuss, S. Drew, K. Weckström, S. Juggins, … D. J. Conley (2006). Multi-proxy evidence of long-term changes in ecosystem structure in a Danish marine estuary, linked to increased nutrient loading. Estuarine, Coastal and Shelf Science, 68(3–4), 567–578. https://doi.org/10.1016/j.ecss.2006.03.013

Erichsen A. C. (Ed.), K. Timmermann (Ed.), J. P. A. Christensen, H. Kaas, S. Markager, F. Møhlenberg (2017). Development of models and methods to support the Danish River Basin Management Plans. Scientific documentation. Aarhus University, Department of Bioscience and DHI, 191 pp.

EU MSFD CIS (2019) Outcomes of the marine-fisheries expert workshop. MSCG meeting document MSCG_24-2019-09. https://circabc.europa.eu/d/a/workspace/SpacesStore/49225513-4413-43cb-af4e-3e1c2644c5ce/MSCG_24-2019-09_Joint%20marine-fisheries%20workshop.docx

Haraguchi, L., H. H. Jakobsen, N. Lundholm, & J. Carstensen (2017). Monitoring natural phytoplankton communities: a comparison between traditional methods and pulse-shape recording flow cytometry. Aquat Microb Ecol 80: 77-92.

HELCOM (2007). Toward a Baltic Sea unaffected by eutrophication, HELCOM Overview 2007, HELCOM Ministerial Meeting.

HELCOM (2019). Guidelines for measuring chlorophyll a.pdf. presented at STATE & CONSERVATION 10-2019, 3MA-5 Att.1.

Hinsby K., S. Markager, B. Kronvang, J. Windolf, T.O. Sonnenborg and L. Thorling (2012). Threshold values and management options for nutrients in a catchment of a temperate estuary with poor ecological status. Hydrology and Earth System Sciences: 16, 2663-2683. doi: 10.5194/hess-16-2663-2012.

Hossain MS, Bujang JS, Zakaria MH, Hashim M (2015) The application of remote sensing to seagrass ecosystems: an overview and future research prospects. International Journal of Remote Sensing 36:61–114. doi: 10.1080/01431161.2014.990649

ICES. 2016a. Working Group on Integrating Surveys for the Ecosystem Approach (WGISUR), 26-28 January 2016, Hamburg, Germany. ICES CM 2016/SSGIEOM:27. 45 pp. http://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/SSGIEOM/2016/WGISUR/WGISUR%202016%20FINAL.pdf#search=wgisur

ICES. 2016b. Working Group on Integrating Surveys for the Ecosystem Approach (WGISUR), 26-28 January 2016, Hamburg, Germany. ICES CM 2016/SSGIEOM:27. 45 pp. http://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/SSGIEOM/2016/WGISUR/WGISUR%202016%20FINAL.pdf#search=wgisur

ICES. 2016c. Report of the Working Group to Demonstrate a Celtic Seas wide ap-proach to the application of fisheries related science to the implementation of the Ma-rine Strategy Framework Directive (WGMSFDemo), 16–18 February 2016, Glasgow, Scotland. ICES CM 2016\SSGIEA:12. 20 pp. http://ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/SSGIEA/2016/WGMSFDemo/WGMSFDemo%20Report%202016.pdf

ICES. 2016d. Final Report of the Working Group on Integrated Assessments of the North Sea (WGINOSE), 14-18 March 2016, ICES Headquarters, Copenhagen, Den-mark. ICES CM 2016/SSGIEA:06. 52 pp. http://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/SSGIEA/2016/WGINOSE/WGINOSE%202016.pdf

Jacobsen H.H. and S. Markager (2016). Annual dynamics of carbon to chlorophyll ratio in temperate coastal waters. Limnology and Oceanography, 61, 1853-1866, doi: 10.1002/lno.10338

Jørgensen, L., S. Markager & M. Maar (2014) On the importance of quantifying bioavailable nitrogen instead of total nitrogen. Biogeochemistry 117, 455-472. doi: 10.1007/s10533-013-9890-9.

JMP NS/CS (2015). Towards a joint monitoring programme for the North Sea and Celtic Sea. Project website: https://www.informatiehuismarien.nl/projecten/joint-monitoring/

JPI Oceans (2016) Multi use of infrastructure for monitoring; joint action facts. http://www.jpi-oceans.eu/multi-use-infrastructure-monitoring

Kemp W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Boicourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith and J. C. Stevenson (2005a). Eutrophication of Chesapeake Bay: historical trends and ecological interactions. Marine Ecology Progress Series 303: 1–29.

Kemp W. M. et al. (2005b). Eutrophication in Chesapeake Bay: historical trends and ecological interactions. Marine Ecology Progress Series 303: 1–29. doi:10.3354/meps303001.

Kishino M., M. Takahashi, N. Okami, and S. Ichimura (1985). Estimation of the spectral absorption coefficients of phytoplankton in the sea, Bull. Mar. Sci., 37(2), 634–642.

Kratzer S, Harvey ET, Philipson P (2014) The use of ocean color remote sensing in integrated coastal zone management A case study from Himmerfjärden, Sweden. Marine Policy 43:29–39

Krause-Jensen D., S. Markager & T. Dalsgaard (2012) Benthic and pelagic primary production in different nutrient regimes. Estuaries and Coasts, 35, 527-545. doi: 10.1007/s12237-011-9443-1.

Kyllingsbæk A. (2008). Landbrugets husholdning med næringsstoffer 1900-2005. Kvælstof- Fosfor - Kalium. DJF Intern Rapport Markbrug Nr. 18.

Lombard F., E. Boss, A.M. Waite, M. Vogt, J. Uitz, L. Stemmann, H.M. Sosik, J. Schulz, J.B. Romagnan, M. Picheral, J. Pearlman, M.D. Ohman, B. Niehoff, K. O. Möller, P. Miloslavich, A. Lara-Lpez, R. Kudela, R. M. Lopes, R. Kiko, L. Karp-Boss, J.S. Jaffe, M. H. Iversen, J. O. Irisson, K. Fennel, H. Hauss, L. Guidi, G. Gorsky, S.L.C. Giering, P. Gaube, S. Gallager, G. Dubelaar, R. K. Cowen, F. Carlotti, C. Brisno-Avena, L. Berline, K. Benoit-Bird, N. Bax, S. Batten, S. D. Ayata, L.F. Artigas & W. Appeltans (2019). Globally consistent quantitative observations of planktonic ecosystems. Frontiers in Marine Science 6:196

Lyngsgaard M. M., S. Markager and K. Richardson (2014). Changes in the vertical distribution of primary production in response to land-based N-loading. Limnology and Oceanography 59, (5) 1679-1690.

Lyngsgaard, M.M., S. Markager, K. Richardson, E.F. Møller & H. H. Jakobsen (2017) How well does chlorophyll explain the seasonal variation in phytoplankton activity? Estuaries and Coasts, 40, 1263-1275, DOI 10.1007/s12237-017-0215-4.

Maar, M., A. Rindorf, E.F. Møller, A. Christensen, K.S. Madsen & M. van Deurs (2014). Zooplankton mortality in 3D ecosystem modelling considering variable spatial-temporal fish consumptions in the North Sea. Prog. Oceanogr. 124, 78–91.

Meier H. E. M., H. C. Andersson, K. Eilola, B. G. Gustafsson, I. Kuznetsov, Müller-Karulis, T. Neumann and O. P. Savchuk (2011). Hypoxia in future climates: A model ensemble study for the Baltic Sea. Geophysical Research Letters, 38(24), 1–6. https://doi.org/10.1029/2011GL049929

Menden-Deuer S. and E. J. Lessard (2000). Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. Limnology & Oceanography 45: 569–579. doi:10.4319/lo.2000.45.3.0569

OSPAR (2012) JAMP Eutrophication Monitoring Guidelines: Chlorophyll a in Water. Agreement 2012-11 (replaces Agreement 1997-04). https://www.ospar.org/documents?d=32933

Remøe, S.O. (2018) Multi-use of infrastructure for monitoring: experiences from the JPI Oceans pilot action. EOOS 2018 Conference, presentation. https://eoosconference2018.eu/sites/default/files/EOOSConf18_Session3_SOtto.pdf

Sosik H. M., R.J. Olson (2007). Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. Limnology & Oceanography Methods 5:204-216

Stæhr P. A. and S. Markager (2004). Parameterization of the chlorophyll a-specific in vivo light absorption coefficient covering estuarine, coastal and oceanic waters. Journal of Remote Sensing, 25, 5117-5130.

Stæhr P. A., S. Markager, K. Sand-Jensen (2004). Pigment specific in vivo light absorption of marine phytoplankton from different trophic regimes. Marine Ecology Progress Series. 275:115-128

Stæhr P. A., G. B. Groom, D. Krause-Jensen, L. B. Hansen, S. Huber, L. Ø. Jensen, M. B. Rasmussen, S. Upadhyay and S. B. Ørberg (2019). Use of remote sensing technologies for monitoring Chlorophyll-a and submerged aquatic vegetation in Danish coastal waters. Aarhus University, DCE – Danish Centre for Environment and Energy, no 139, 62 pp. Technical Report.

Stedmon C.A., S. Markager and H. Kaas (2000). Optical properties and signatures of chromophoric dissolved organic matter (CDOM) in Danish coastal waters. Estuarine, Coastal and Shelf Science, 51, 267-278.

Strong J.A., M. Elliott M. (2017). The value of remote sensing techniques in supporting effective extrapolation across multiple marine spatial scales. Marine Pollution Bulletin 116:405–419.

Timmermann K., S. Markager and K. Gustafsson (2010). Streams or Open Sea? Tracing sources and effects of nutrient loadings in a shallow estuary with a coupled 3D hydrodynamic-ecological model. *J. Marine Systems*, 82, 111-121.

Timmermann K., G. E. Dinesen, S. Markager, L. Ravn-Jonsen, M. Bassompierre, E. Roth and J. G. Støttrup (2014). Development and use of a bioeconomic model for management of mussel fisheries under different nutrient regimes in the temperate estuary of the Limfjord, Denmark. Ecology and Society, 19(1). https://doi.org/10.5751/ES-06041-190114

Van der Zande D., H. Lavigne, A. Blauw, T. Prins, M. Eleveld, F. Gohin, S. Pardo, G. Tilstone (2019). Coherence in assessment framework of chlorophyll a and nutrients as part of the EU project 'Joint monitoring programme of the eutrophication of the North Sea with satellite data' (Ref: DG ENV/MSFD Second Cycle/2016). Activity 2 Report. 107 pp