BarettaBekker
Mariene ecologie

Deltares
Enabling Delta Life

Assessments of phytoplankton in the Netherlands and neighbouring countries according to OSPAR and WFD

Deltares

Assessments of phytoplankton in the Netherlands and neighbouring countries according to OSPAR and WFD

drs. J.G. Baretta-Bekker
dr. T.C. Prins

1207005-000

Deltares

Title
Assessments of phytoplankton in the Netherlands and neighbouring countries according to OSPAR and WFD

Client	Project	Reference	Pages
RWS Waterdienst	1207005-000	1207005-000-ZKS-0007	69

Keywords
North Sea, Phytoplankton, Chlorophyll, *Phaeocystis, OSPAR, Comprehensive Procedure, WFD*

Summary
In this report the OSPAR assessment and WFD assessment as applied by the Netherlands have been described, differences in used methods have been highlighted and results have been compared with each other.
In the 1-nm coastal zone of the Netherlands both the OSPAR Comprehensive Procedure and the Water Framework Directive have to be applied. In OSPAR the 1-nm zone is only a small part of the area Coastal Waters, which extend from the low-water line to the salinity isoline of 34.5.
As OSPAR distinguishes only two classes and WFD five, it is the boundary between acceptable and unacceptable in both assessment methods that counts. For OSPAR that is the assessment level between Non Problem and Problem Area and for WFD it is the boundary between good and moderate (g/m). The coastal water bodies of the Netherlands are divided into two types: the polyhaline and the euhaline, because of different salinity ranges, caused by different freshwater inflow and hence leading to different nutrient loads.
The value of the OSPAR assessment level for chlorophyll-a in the Coastal Waters is identical to the g/m assessment level of the euhaline WFD water bodies, but lower than that of the polyhaline water bodies. It is to be expected that the OSPAR assessment score for chlorophyll per station is not always identical to the WFD assessment score applied to the same stations, which does not imply that both assessment methods contradict each other, because of the different areas where the assessments are meant to be applied.
To assess *Phaeocystis* OSPAR uses the maximum cell count per year, while WFD calculates the bloom frequency of *Phaeocystis*, defined as the number of months per year with *Phaeocystis* blooms.
Sensitivity analysis with a synthetic time series of daily values, comparing both methods shows that in general a maximum concentration or cell count is an inaccurate metric, even with relatively short sampling intervals of e.g. one week. It also showed that the 90-percentiles are not a very reliable metric with sampling intervals of more than two weeks. With longer sampling intervals, averages are more accurate than 90-percentiles.
In general, it should be realized that at long sampling intervals, all metrics have a relatively low accuracy. The risk of failing to detect changes in status and the risk of misclassification is high in those cases.
the Netherlands monitor *Phaeocystis* only during the growing season. In only a small number of cases (3 out of 88) the WFD indicator for *Phaeocystis* gives a different status assessment then when it is monitored year round.
An overview of the analytical methods for measuring chlorophyll and counting cells of *Phaeocystis* or other phytoplankton species has been included.

Reference
Opdrachtnr 31073086/4500200811

Deltares

Title
Assessments of phytoplankton in the Netherlands and neighbouring countries according to OSPAR and WFD

Client	Project	Reference	Pages
RWS Waterdienst	1207005-000	1207005-000-ZKS-0007	69

Versie	Datum	Auteur	Paraaf	Review	Paraaf	Goedkeuring	Paraaf
	apr. 2014	drs. J.G. Baretta-Bekker		Dr. J. Baretta			
		dr. T.C. Prins					

State
final

Table of Contents

1 Introduction

After the implementation of the European Water Framework Directive (EC-WFD, 2000), which was limited to the coastal waters and did not apply to all marine waters, the European Marine Strategy Framework Directive (MSFD) has been adopted in 2008 to protect the marine environment and its ecosystems, and to promote its sustainable use. The MSFD has been anchored into Dutch legislation in 2010. The objective of the MSFD (EC-MSFD, 2008) is to achieve and / or maintain good environmental status in European marine waters. The MSFD requires the Member States to take the necessary measures to achieve this goal in 2020.

In the MSFD eleven descriptors have been defined, which are relevant for the implementation of the directive. The descriptor of concern in this report is descriptor 5 for eutrophication, which states: *"Human-induced eutrophication is minimised, especially adverse effects thereof, such as losses in biodiversity, ecosystem degradation, harmful algae blooms and oxygen deficiency in bottom waters"* (EC-MFSD, 2008).

The European Commission is directing the process, but the individual Member States are responsible for the implementation of the MSFD and the cooperation and international coordination within a marine (sub)region. Member states shall, as far as possible, build upon relevant existing programmes and activities developed in the framework of international agreements such as Regional Sea Conventions.

In the OSPAR convention eutrophication is one of the points of concern. OSPAR developed the Comprehensive Procedure, which is a harmonized, integral assessment system of the main causes and effects of eutrophication. The OSPAR Common Procedure is 'fit for the purpose' of assessing this descriptor and supports the setting of targets and indicators under MSFD (OSPAR, 2012a; OSPAR 2012b). As OSPAR is in force for the whole North-eastern Atlantic, while the WFD is for ecology only valid in marine waters in the zone within one nautical mile from the coast line (Figure 1.1), the comparability between OSPAR COMP and WFD needs attention. There are differences between OSPAR COMP and WFD assessments in the use of chlorophyll-a and *Phaeocystis* metrics, both in geographical extent and in the aggregation of the metrics into the final assessment.

The objective of this report is to make an inventory, for chlorophyll-a and *Phaeocystis,* of the OSPAR Comprehensive Procedure and of the WFD assessment method and the results of both methods for the Netherlands and the North Sea countries: Belgium, Germany, UK and France.

For the WFD Prins & Baretta-Bekker (2010) gave an overview of the scientific background of the *Phaeocystis* metric, as one of the indicators for eutrophication. It contains the various aspects of the methods from sampling to setting the class boundaries used by the Netherlands and the neighbouring countries. The present report expands on this paper by including the WFD metric for chlorophyll in order to account for the other phytoplankton component.

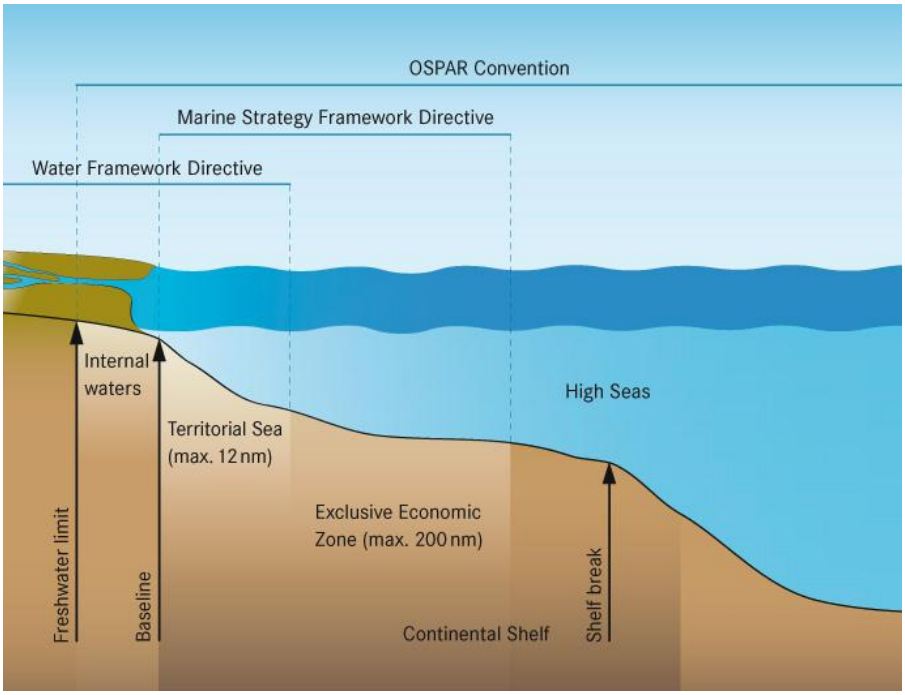

Figure 1.1 Jurisdictional areas of the United Nations Convention on the Law of the Sea, the OSPAR Convention, the EU Water Framework Directive (WFD) and the EU Marine Strategy Framework Directive (MSFD). Source: OSPAR (2010).

Outline The assessment method that will be used for the MSFD, the OSPAR COMP and the one of the WFD, will be described in Chapter 2. The results of the two assessment methods will be presented for the Dutch part of the North Sea in chapter 3 and compared with each other in Chapter 4. In Chapter 5 the OSPAR and WFD results of the adjacent areas of the North Sea countries Belgium, Germany, UK and France, have been compared and Chapter 6 focuses on methodological issues. Discussion and Conclusions can be found in the Chapters 7 and 8, respectively.

2 Assessment methods

2.1 The OSPAR Comprehensive Procedure (COMP)

OSPAR stands for the Oslo-Paris convention which has been amalgamated in 1992 from the Oslo Convention on the dumping of waste into the sea of 1972, and the Paris Convention of 1974 on the pollution of the sea from land-based sources. OSPAR aims to protect the marine environment in the North East Atlantic (including the North Sea) through international cooperation. In 1998, after ratification by the fifteen signatories[1], the treaty entered into force.

Within OSPAR the so-called Comprehensive Procedure has been developed. The Comprehensive Procedure consists of a set of qualitative assessment criteria which are linked to form a holistic assessment and area classification with respect to the eutrophication status of a given maritime area. The holistic approach is reflected in the selection and application of such common assessment parameters which reflect, once inter-linked, the main cause/effect relationships in the eutrophication process. These cause/effect linkages form the essence of the classification process as illustrated by a generic conceptual framework for all categories of surface waters (Figure 2.1). The assessment criteria have been divided into four categories: the causative factors, the direct effects, indirect effects and other possible effects of nutrient enrichment, each with their associated area-specific parameters (Table 2.1).

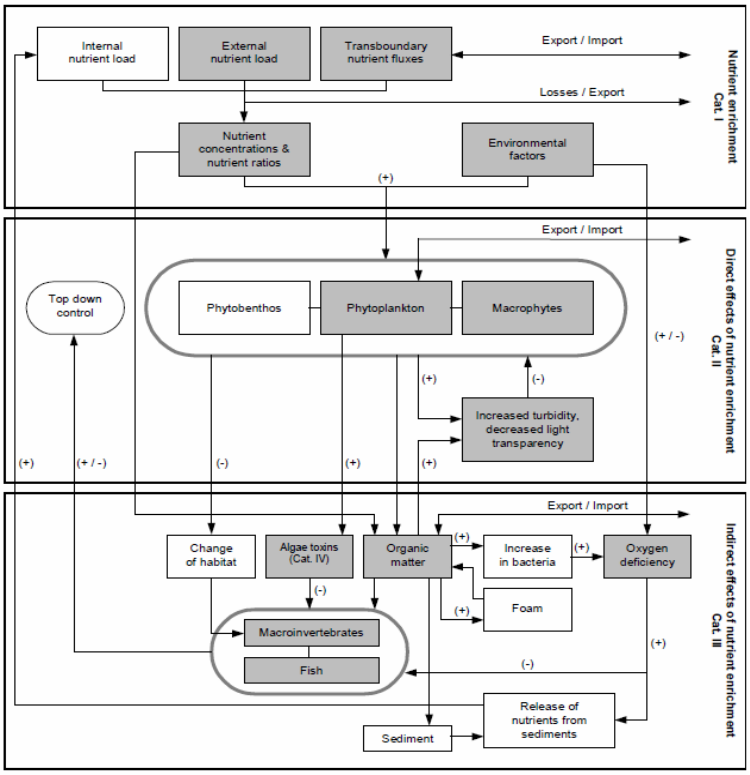

Figure 2.1 Generic conceptual framework to assess eutrophication in all categories of surface waters (OSPAR, 2005). Note: Shaded boxes indicate components relevant for the Comprehensive Procedure. '+' indicate enhancement; '-' indicate reduction; Cat. I = Category I. Degree of nutrient enrichment (causative factors); Cat. II = Category II. Direct effects of nutrient enrichment; Cat. III = Category III. Indirect effects of nutrient enrichment; Cat. IV = Category IV. Other possible effects of nutrient enrichment.

[1] *De fifteen countries are: Belgium, Denmark , Finland, France, Germany, Ireland, Iceland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Swiss en the United Kingdom.*

Table 2.1 Harmonised assessment parameters and related elevated levels. Note that parameters found at levels above the assessment level are considered as "elevated levels" and entail scoring of the relevant parameter category as (+). For concentrations, the "assessment level" is defined as a justified area-specific % deviation from background not exceeding 50%.

Assessment parameters	
Category I	**Degree of nutrient enrichment**
	1 Riverine inputs and direct discharges (area-specific) *ated inputs and/or increased trends of total N and total P* (compared with previous years)
	2 Nutrient concentrations (area-specific) Elevated level(s) of winter DIN and/or DIP
	3 N/P ratio (area-specific) Elevated winter N/P ratio (Redfield N/P = 16)
Category II	**Direct effects of nutrient enrichment (during growing season)**
	1 Chlorophyll *a* concentration (area-specific) Elevated maximum and mean level
	2 Phytoplankton indicator species (area-specific) Elevated levels of nuisance/toxic phytoplankton indicator species (and increased duration of blooms)
	3 Macrophytes including macroalgae (area-specific) Shift from long-lived to short-lived nuisance species (e.g. *Ulva*). Elevated levels (biomass or area covered) especially of opportunistic green macroalgae).
Category III	**Indirect effects of nutrient enrichment (during growing season)**
	1 Oxygen deficiency Decreased levels (< 2 mg/l: acute toxicity; 2 - 6 mg/l: deficiency) and lowered % oxygen saturation
	2 Zoobenthos and fish Kills (in relation to oxygen deficiency and/or toxic algae) *g-term area-specific changes in zoobenthos biomass and species composition*
	3 Organic carbon/organic matter (area-specific) Elevated levels (in relation to III.1) (relevant in sedimentation areas)
Category IV	**Other possible effects of nutrient enrichment (during growing season)**
	1 Algal toxins Incidence of DSP/PSP mussel infection events (related to II.2)

2.1.1 Chlorophyll-a

For chlorophyll the Netherlands use the mean and the 90-percentile (instead of the maximum) concentrations over the growing season from March to September, inclusive, as the assessment parameters. The background values of these parameters and their related elevated levels, which are used as assessment levels, are outlined in Table 2.2. These assessment levels are in correspondence with OSPAR (2011).

The area-specific background concentrations of chlorophyll and the related elevated levels, which are used as assessment levels, for the Dutch marine waters are given in Table 2.2.

Table 2.2. Background concentrations and assessment levels of chlorophyll-a (µg C/l) for each of the Dutch OSPAR areas in the North Sea. The numbers are mean and 90-percentile values of the concentrations in the growing season.

Area	Background concentration		Assessment levels	
	mean	90-perc	mean	90-perc
Coastal zone	**5**	**10**	**7.5**	**15**
Southern Bight offshore	**1.5**	**3**	**2.25**	**4.5**
Oyster Grounds	**1.5**	**3**	**2.25**	**4.5**
Dogger Bank	**1.5**	**3**	**2.25**	**4.5**

2.1.2 *Phaeocystis*

For *Phaeocystis* the assessment parameter is the maximum number of cells found during the growing season from March to September, inclusive. This maximum should not reach values above the assessment level. The assessment level for *Phaeocystis* is set at 10^7 cells/l in all areas. A bloom of 10^7 cells/l or more is considered an extreme bloom.

2.2 Final assessment according to OSPAR

The status of chlorophyll-a and *Phaeocystis* determines the final assessment in this report. In general, the final assessment is an integrated area classification, by integrating the assessment parameters: areas with direct and/or indirect effects are problem areas, regardless of the nutrient concentrations (Table 2.3, groups a and b), areas without direct (chlorophyll; indicator species; macrophytes) and/or indirect effects (oxygen deficiency; benthos; fish; organic carbon) are non-problem areas (Table 2.3, groups c1 and d) on the condition that there are enough Category II and III data available. If that is not the case it is a potential problem area Table 2.3, groups c2). NB Nutrient enrichment in one area may contribute to direct and indirect effects elsewhere. OSPAR also tests for trends over the assessment period, see for example Figure 2.2.

	DIN	DIP	Chl	O2
Coastal W	~ decrease	~ decrease	~ decrease	~ increase
Wad Sea	~ decrease	~ decrease	~ decrease	~ decrease
W Scheldt	~ decrease	~ decrease	~ decrease	~ increase
E-Dollard	~ decrease	~ decrease	no trend	no trend
South Bight offsh	~ decrease	no trend	~ decrease	~ increase
Oyster G	no trend	no trend	? no trend	no trend
Doggersb	no trend	no trend	~ decrease	~ increase

Figure 2.2 Overview of trends per assessment parameter. Green indicates in the right direction and red in the wrong direction (Baretta-Bekker et al., 2008).

Table 2.3 Examples of the integration of categorised assessment parameters (see Table 2.1) for an initial classification (OSPAR, 2005).

	Category I Degree of nutrient enrichment Nutrient inputs Winter DIN and DIP Winter N/P ratio	Category II Direct effects Chlorophyll *a* Phytoplankton indicator species Macrophytes	Categories III and IV Indirect effects/other possible effects Oxygen deficiency Changes/kills in zoobenthos, fish kills Organic carbon/matter Algal toxins	Initial Classification
a	+	+	+	problem area
	+	+	-	problem area
	+	-	+	problem area
b	-	+	+	problem area[2]
	-	+	-	problem area[2]
	-	-	+	problem area[2]
c	+	-	-	non-problem area [3]
	+	?	?	Potential problem area
	+	?	-	Potential problem area
	+	-	?	Potential problem area
d	-	-	-	non-problem area

(+) = Increased trends, elevated levels, shifts or changes in the respective assessment parameters in Table 2.1

(-) = Neither increased trends nor elevated levels nor shifts nor changes in the respective assessment parameters in Table 2.1

? = Not enough data to perform an assessment or the data available is not fit for the purpose

Note: Categories I, II and/or III/IV are scored '+' in cases where one or more of its respective assessment parameters is showing an increased trend, elevated level, shift or change

In this specific case with enough data in each area, but with only data of the direct effects chlorophyll-a and *Phaeocystis*, an area is a **problem area** when chlorophyll-a and/or *Phaeocystis* are above the assessment value, and a **non-problem area** when both parameters are below the assessment level.

2.2.1 The OSPAR monitoring stations in NL

Monitoring in the Netherlands is organized in the so-called MWTL programme of Rijkswaterstaat (RWS) since the 1970s. The monitoring network for the North Sea consists of stations along transects perpendicular to the coast, along the salinity gradient. In the beginning of the 1990s this programme has been adapted to serve the OSPAR monitoring requirements. The marine waters of the Dutch continental shelf are divided into areas based on chemical and hydrological characteristics. These areas in the Dutch continental shelf are: the Coastal waters with salinities below 34.5, and the offshore waters with salinities above 34.5 divided into: The Southern Bight, The Oyster Grounds and the Dogger Bank (Figure 2.3).

The stations along these gradients in the OSPAR areas are given in Table 2.4 and shown in Figure 2.3.

[2] *For example, caused by transboundary transport of (toxic) algae and/or organic matter from adjacent/remote areas.*

[3] *The increased degree of nutrient enrichment in these areas may contribute to eutrophication problems elsewhere.*

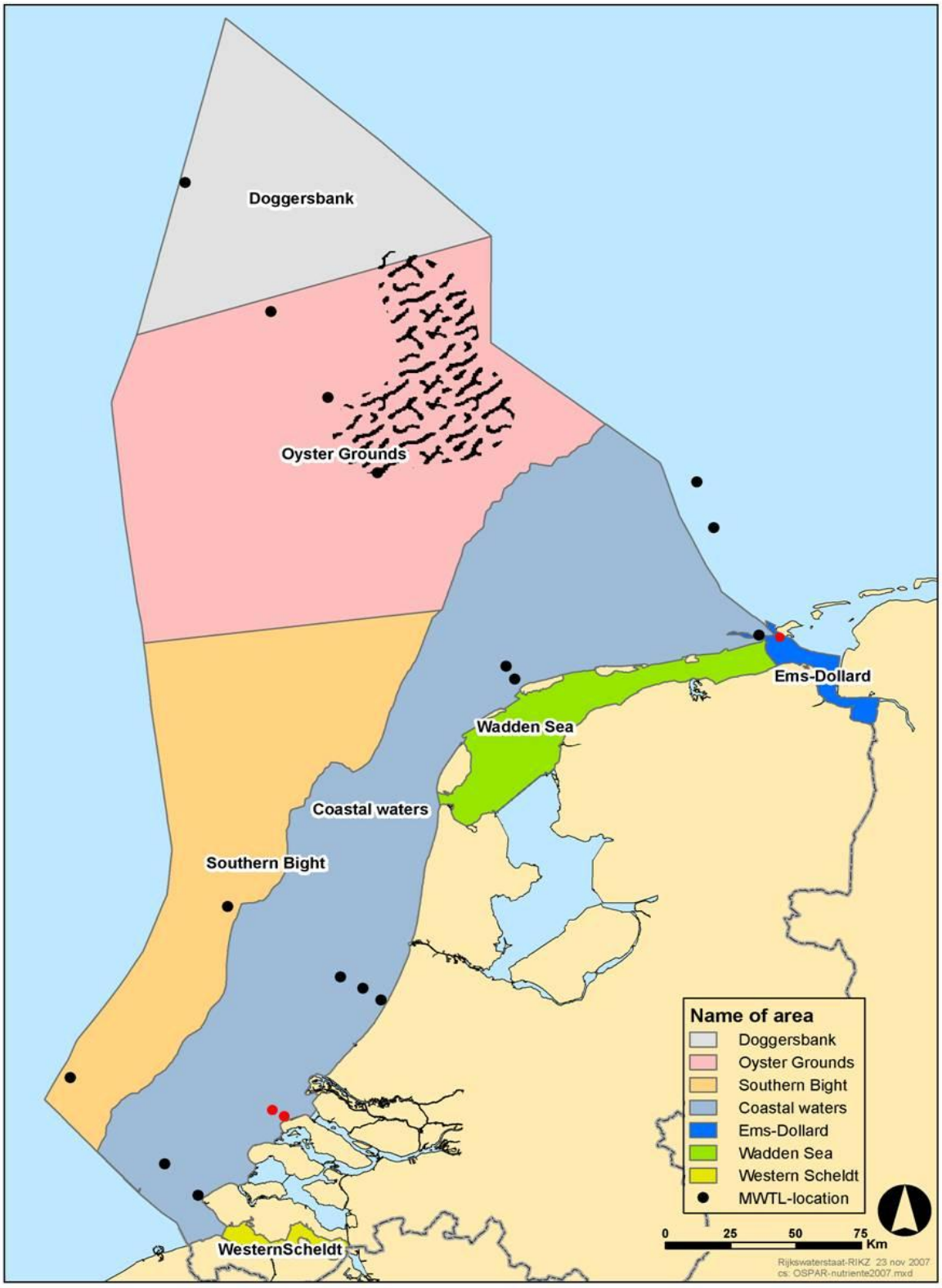

Figure 2.3 The Dutch continental shelf with the four sub areas: Coastal waters (the border of the Coastal waters is the average 34.5 isohaline) and Offshore waters (salinity > 34.5) divided into: Southern Bight offshore, Oyster Grounds and Dogger Bank. Sampling stations in Coastal waters and Offshore water are indicated. The names of the transects with the black stations are from south to north; Walcheren, Noordwijk, Terschelling and Rottumerplaat. Shaded area is the Oyster Grounds proper. The southern red stations are Goeree 2 and Goeree 6 and in the north Huibertgat oost has been added.

Table 2.4 Dutch monitoring stations per transect with the OSPAR area, where in they are situated.

Transect	Station	OSPAR area
Walcheren	Walcheren 2	Coastal waters
	Walcheren 20	Coastal waters
	Walcheren 70	Southern Bight
Goeree	Goeree 6/Goeree 2[4]	Coastal waters
Noordwijk	Noordwijk 2	Coastal waters
	Noordwijk 10	Coastal waters
	Noordwijk 20	Coastal waters
	Noordwijk 70	Southern Bight
Terschelling	Terschelling 4/Boomkensdiep[5]	Coastal waters
	Terschelling 10	Coastal waters
	Terschelling 100	Oyster Grounds
	Terschelling 135	Oyster Grounds
	Terschelling 175	Oyster Grounds
	Terschelling 235	Dogger Bank
Rottumerplaat	Huibertgat oost	Coastal waters
	Rottumerplaat 3	Coastal waters
	Rottumerplaat 50	Coastal waters
	Rottumerplaat 70	Coastal waters

2.3 Water Framework Directive assessment method

In 2000 the European Water Framework Directive (WFD) came into force and was incorporated into Dutch legislation in 2006. The objective of the WFD is to ensure good quality of the European surface and ground waters in 2015.

The WFD applies, for the ecological status, to all fresh waters, to transitional waters and to the 1-nautical mile zone of the coastal waters in the EU Member States (Figure 1.1). In Dutch coastal waters five water bodies are distinguished, in which the biological quality elements phytoplankton and benthos are assessed. This report is limited to phytoplankton. When the indicators for the biological quality element phytoplankton for the Water Framework Directive were determined in the Netherlands, the OSPAR Comprehensive Procedure provided the point of departure. Considerable attention is paid in Van den Berg *et al.* (2004) to the arguments in favour of deviating on a number of points from the OSPAR methodology. The following indicators were selected:

Phytoplankton – Biomass

The biomass of phytoplankton in the marine transitional and coastal waters is assessed on the basis of the 90-percentile chlorophyll-a biomass, during the growing season, which is from March to September, inclusive.

Phytoplankton – Species composition

From the OSPAR list of indicator species, which includes the nuisance phytoplankton species *Phaeocystis* as well as a number of potentially toxic dinoflagellates only the abundance of *Phaeocystis* is used as an indicator, as the causal relation between nutrient enrichment and the occurrence of toxic algal blooms is not clear (Prins & Baretta-Bekker, 2010).

The WFD phytoplankton classification tool has two indicators: chlorophyll-a and *Phaeocystis* .

2.3.1 Chlorophyll-a

The reference concentration for chlorophyll in the Dutch coastal waters is based on the AMOEBE approach elaborated in Baptist & Jagtman (1997), but also the values used by OSPAR, expert

[4] *From 2007 on the new station Goeree 2 will be used for the WFD, because Goeree 6 is outside the 1-nautical mile zone.*

[5] *In 2007 Terschelling 4 has been replaced by Boomkensdiep.*

judgement and the EU intercalibration procedure (Carletti & Heiskanen, 2009) have been taken into account. However, it should be noted that the results of Carletti & Heiskanen (2009) have not been accepted by the EC, because they did not follow a proper comparison protocol and they should be re-evaluated before 2016. The Dutch coastal water bodies can be divided into two groups, based on their salinity ranges during the growing season: the polyhaline and the euhaline type. The Holland coast and the Northern Delta coast with larger salinity ranges and lower salinities belong to the polyhaline type, while the Zeeland coast and the Wadden coast with smaller salinity ranges, are of the euhaline type (see Figure 2.4).
In the framework of the “Watersysteemverkenning” (Water System Exploration) in the Netherlands, so called reference values, representing the upper boundary of the good status, for a number of functional groups and individual species have been calculated (Baptist & Jagtman, 1997). For chlorophyll-a the “reference” value has been expressed as the 90-percentile value of the concentration. The calculation of this value has been based on model simulations and expert judgement and is 14.3 µg/l for the Holland coast, and hence also for the other Dutch polyhaline water bodies. This value agrees well with the value deduced from Cadée & Hegeman (2002), see Carletti & Heiskanen (2009).

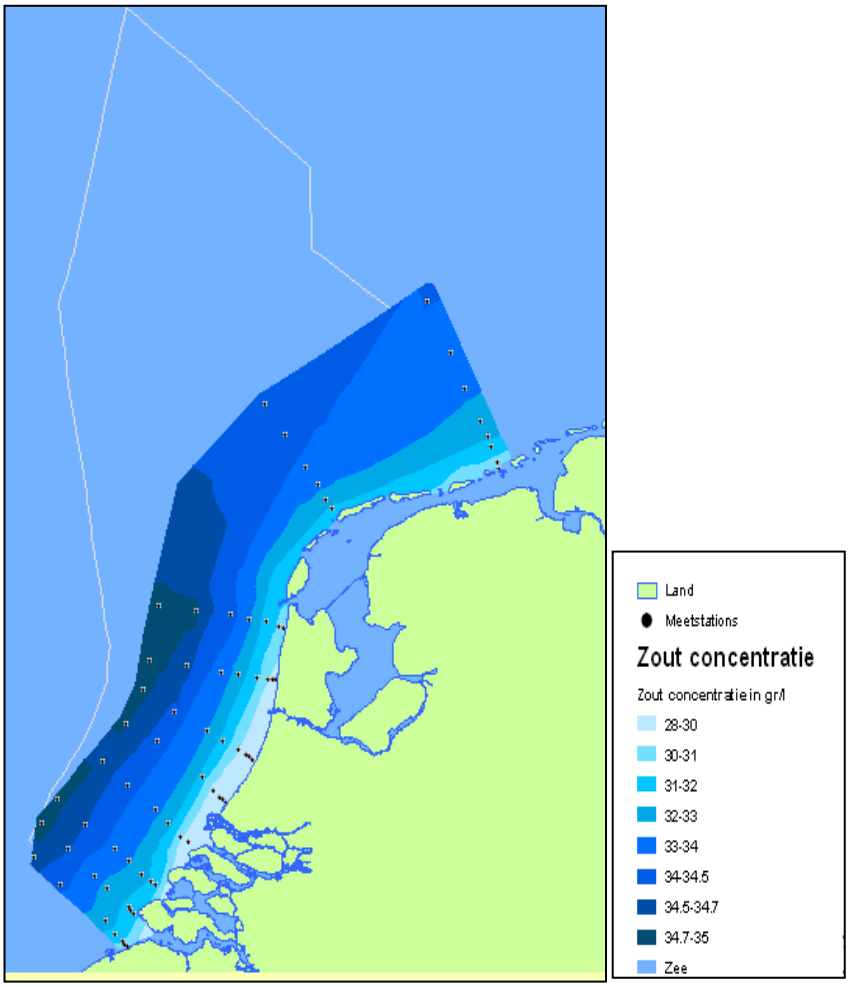

Figure 2.4 Salinity distribution along the North Sea coast of the Netherlands

90-percentile of concentration Chl-a						
Coastal waters		h/g	g/m	m/p	p/b	
ref		10	15	30	60	
Zeeland coast	6.7	14	21	42	84	
N. Delta coast	9.3	14	21	42	84	
Holland coast	9.3	10	15	30	60	
Wadden coast	6.7	10	15	30	60	
Ems coast	6.7					
	high	good	moderate	Poor	Bad	
EQR	1.0	0.8	0.6	0.4	0.2	

Figure 2.5 Chlorophyll reference value (ref; µg/l), class boundaries (µg/l) and standardisation to achieve the EQR for the coastal waters.

This “reference” 14 µg/l has been taken as boundary between the WFD high and good status for the polyhaline waters. The boundary between good and moderate in this classification, which is the target, is 21 µg/l, one and a half times higher than the high/good boundary (Figure 2.5). This factor of 1.5 has been used in OSPAR also, but there it is applied to the “background” value, and the relation between the “reference” value, as used in the “Watersysteemverkenning” and the OSPAR background value has not been unequivocally defined. For the euhaline water bodies lower values have been set. The good/moderate class boundaries for the 90-percentile summer chlorophyll-*a* concentrations for the euhaline areas are set on 15 µg/l, identical to the class boundaries set by Belgium, UK and France.

2.3.2 *Phaeocystis*

Blooms of *Phaeocystis* are a natural phenomenon, but extremely abundant (>> 10^6 cells/l) and long-lasting blooms are considered an effect of eutrophication. The Dutch metric for *Phaeocystis* takes bloom frequency as criterion for eutrophication (Van der Molen and Pot, 2007), using the monitoring data from seven months (March – September, inclusive) considered as the growing season for phytoplankton in Dutch coastal waters. As *Phaeocystis* only rarely reaches bloom densities during the winter months (October-February), the assumption is that it does not bloom during winter.

Comparison of monitoring data from the growing season (March – September, inclusive) with data of the whole year, as agreed in the intercalibration process, shows that the difference in assessment score is small. In 5.3% of the cases blooms occur between October and February (see Prins & Baretta-Bekker, 2010), but in only 2.3% of the cases this results in a different score.

The bloom frequency is determined by looking at the number of months in a year with more than 10^6 *Phaeocystis* cells/l. The frequency is expressed as a percentage of 12 months. A frequency of more than one month per year with a bloom (>10^6 cells/l) is assumed to be the boundary between “high” and “good” status. More than two months per year is the boundary between “good” and “moderate” status. More than four months per year is the boundary between moderate and poor status (Figure 2.6). This methods guarantees that blooms lasting over one month are counted twice or more.

Frequency (%)		10	17	35	85	
	high	Good	moderate	Poor	Bad	
EQR	1.0	0.8	0.6	0.4	0.2	

Figure 2.6 Phaeocystis class boundaries and standardisation to achieve the EQR for the coastal waters.

2.3.3 Combination of EQRs of chlorophyll-a and *Phaeocystis*

Both assessments are expressed in a value between 0 and 1, for which the value has been offset against the reference value. Normalisation of this value results in the Ecological Quality Ratio (EQR). This is shown in Figures 2.4 and 2.5 for both classifications.

The EQR of the quality element phytoplankton is calculated by taking the average of the EQRs for the status of chlorophyll-a and for the status of *Phaeocystis,* in those cases where the EQR of *Phaeocystis* is smaller than the EQR of chlorophyll-a. If this is not the case, the final EQR is equal to the EQR of chlorophyll-a.

EQR_{phyto}= Minimum((EQR_{chl}+EQR_{Phaeo})/2, EQR_{chl})

Or in words: the final assessment is the smallest of

1. the average of the two assessments and
2. the assessment that is based on chlorophyll-a alone;

In other words, *Phaeocystis* can only make the end result worse, but not improve it, so even a total absence of *Phaeocystis* cannot make the final score better than the chlorophyll-a score. See the examples in Chapter 3.2.3.

2.3.4 The WFD monitoring stations in NL

For the WFD assessment a subset of the Dutch monitoring programme has been used. The coastal zone of the Netherlands is influenced by the rivers Scheldt, Meuse, Rhine and Ems (from south to north). The catchment areas are the basis for the division into water bodies (Figure 2.7a; Table 2.5). For each water body one monitoring station is used. Because the Northern Delta coast had no station within the 1 nautical mile zone, a new station, Goeree 2, has been added to the MWTL programme in 2007. Until that year the station Goeree 6 situated 6 km off the coast, has been used. A comparable situation occurs for the Wadden Coast, the area north of the Wadden Islands, that is influenced by the coastal river. Until 2007 the station Terschelling 4 has been used, but in 2007 this was replaced by Boomkensdiep.

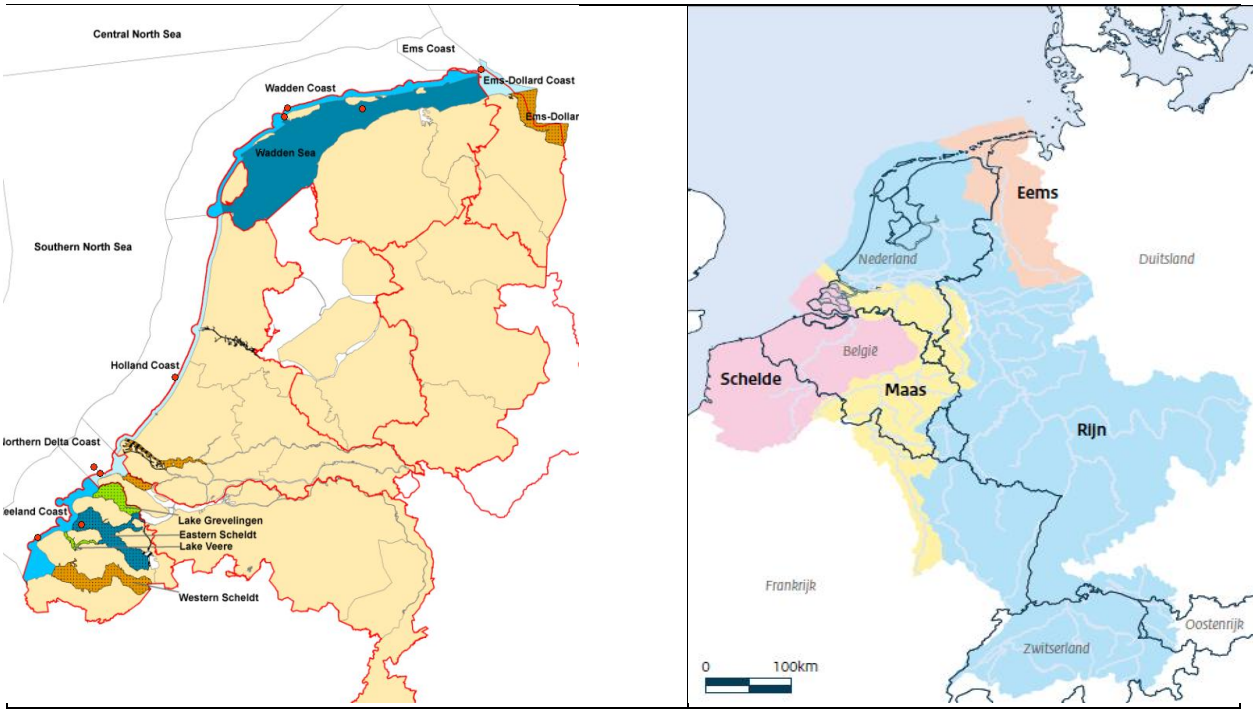

Figure 2.7 Left: Map of the Netherlands with the boundaries of the coastal (and transitional) WFD water bodies. The stations, listed in Table 3.1 are given as red dots. Right: The different catchment areas in the Netherlands (source: Rijkswaterstaat, 2009).

Table 2.5 WFD Water bodies with monitoring station(s).

Water body	Monitoring station	River catchment
Zeeland coast	Walcheren 2	Scheldt
Northern Delta coast	Goeree 6/Goeree 2	Meuse
Holland coast	Noordwijk 2	Rhine
Wadden coast	Terschelling 4/Boomkensdiep	Rhine
Ems coast	Rottumerplaat 3	Ems

3 Assessment results of OSPAR and WFD assessments

3.1 The OSPAR assessment for the Dutch continental shelf

The OSPAR assessment for phytoplankton in the Dutch marine and estuarine waters over the period of 1990 to 2010 is based on the assessment of the chlorophyll-a mean, chlorophyll-a 90-percentile and the maximum concentration of *Phaeocystis* in the growing season.

3.1.1 The OSPAR assessment applied to individual Dutch monitoring stations

The status of the individual stations, based on the assessments of chlorophyll-a (mean and 90-percentile) and *Phaeocystis* is presented in Table 3.1. An area (or station) is a problem area when one or both of the parameters are above the area-specific assessment level.

Table 3.1 OSPAR assessment for the Dutch North Sea stations, arranged to the transects. Blue cells indicate Non Problem station. Red cells: problem station; p= Phaeocystis is above assessment level; c= chlorophyll-a is above assessment level; and cp= both indicators are above assessment level

.

monitoring station	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Goeree 06/02	c	cp	cp	cp	cp	cp	cp	cp	cp	cp	c	cp	c	cp	cp	cp	c	cp	cp	c	cp
Noordwijk02	p	cp	c	c	cp	cp	cp	cp	cp	cp	c	cp	c	cp	c	c		cp	cp	cp	cp
Noordwijk10	cp	cp	p	cp	c	cp	cp	cp	cp	cp	cp	cp	c	p	p	cp		cp	cp	P	p
Noordwijk20	cp	p	p	c	c	cp	p	p	cp	cp	p	cp	c	p				cp	p	P	p
Noordwijk70	p	cp		cp	c	c	cp	p	p	p		cp	c	cp	cp	cp		cp	p	P	c
Huibertgat oost		cp		cp	c	cp	c	p	cp	cp	c	cp	c		cp	cp	p	cp	cp	c	p
Rottumerplaat03		cp	c	c	c	cp	c	c	c	c	c	c	c	c	cp	p	c	cp	cp	c	cp
Rottumerplaat50	p				p		c														
Rottumerplaat70															p						
T4/boomkensdiep	cp	cp	c	cp	cp	cp	cp		cp			cp	p	cp	cp	cp		cp	cp	P	
Terschelling 10	p	p		cp	c	p			cp	p		p	p		p	p		p	p	P	
Terschelling 100																					
Terschelling 135		c																			
Terschelling 175							c														
Terschelling 235						c	c				c						p				
Walcheren 02	c	cp	c	cp	c	c	cp	cp	cp	c	c	cp	c	cp	cp	c	c	cp	cp	cp	cp
Walcheren 20	C	p	p	p	c	cp	c	p	cp	p		cp		cp	p			cp	p	cp	p
Walcheren 70	Cp	cp	c	cp	c	c	c			c		cp	c	cp	c	c			cp		cp

3.1.2 The OSPAR assessment applied to OSPAR areas in the North Sea

The status of each of the OSPAR areas is presented for chlorophyll-a (mean and 90-percentile combined) in Table 3.2. For the assessment of chlorophyll all observations in the OSPAR area are used to calculate the mean and the 90-pecentile. The status of *Phaeocystis*, based on the maximum concentration of cells/l is presented in Table 3.3. An area is a problem area when chlorophyll-a (mean and/or 90-percentile concentration) and/or *Phaeocystis* is above the area-specific assessment level. In only one case (Oyster Grounds 1996) the mean chlorophyll-a is above the assessment level while the 90-percentile chlorophyll-a is below assessment level. In all other cases it is either chlorophyll-a 90-percentile or are both statistics above the assessment level. Table 3.4 shows the overall assessment. When the indirect effect oxygen is included as well two of

the assessments will be changed. Oxygen is only below assessment (6 mg/l) in the Oyster Grounds in the years 1995 and 2003. In those years chlorophyll-a and *Phaeocystis* are not above assessment. Table 3.5 shows the assessment averaged over a period of six years.

Table 3.2 OSPAR assessment for chlorophyll-a in the OSPAR areas. The colour of the cells indicates the status: red, with c = Problem Area; blue = Non Problem Area.

OSPAR areas	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Coastal Waters	c	c		c	c	c	c		c	c	c	c	c	c		c		c	c		
Southern Bight	c	c		c	c	c	c					c	c	c	c			c			c
Oyster Ground							c*														
Dogger Bank						c															

* only the mean chlorophyll -a is above the assessment level.

Table 3.3 OSPAR assessment for Phaeocystis in the OSPAR areas. The colour of the cells indicates the status: red, with p= Problem Area; blue = Non Problem Area.

OSPAR areas	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Coastal Waters	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p	p		p	p	p	p
Southern Bight	p	p		p			p	p	p	p		p		p	p	p		p	p	p	p
Oyster Ground																					
Dogger Bank																	p				

Table 3.4 OSPAR final assessment for phytoplankton for the OSPAR areas. The colour of the cells indicates the status: red, with p= Problem Area, caused by Phaeocystis; red, with c= Problem Area, caused by chlorophyll-a; red, with cp = Problem Area, caused by Phaeocystis and chlorophyll-a; blue = Non Problem Area.

OSPAR areas	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Coastal Waters	cp	cp	p	cp	cp	cp	cp	p	cp	cp	cp	cp	cp	cp	p	cp		cp	cp	p	p
Southern Bight	cp	cp		cp	c	c	cp	p	p	p		cp	c	cp	cp	P		cp	p	p	cp
Oyster Ground							c														
Dogger Bank						c											p				

Table 3.5 OSPAR final assessment for phytoplankton, averaged over six years for the OSPAR areas. The colour of the cells indicates the status: red = Problem Area, caused by Phaeocystis and/or chlorophyll-a; blue = Non Problem Area.

OSPAR areas	1990 - 1995	1991 - 1996	1992 - 1997	1993 - 1998	1994 - 1999	1995 - 2000	1996 - 2001	1997 - 2002	1998 - 2003	1999 - 2004	2000 - 2005	2001 - 2006	2002 - 2007	2003 – 2008	2004 - 2009	2005 - 2010
Coastal Waters																
Southern Bight																
Oyster Ground																
Dogger Bank																

3.2 The WFD assessment for the Dutch coastal water bodies

3.2.1 Chlorophyll-a

The assessment results for chlorophyll-a per coastal water body for the years between 1990 and 2010 are expressed in Ecological Quality Ratios (EQRs), according to Figure 2.4 and given in Table 3.6.

Table 3.6 WFD assessment for chlorophyll-a of the Dutch coastal water bodies for the years 1990-2010. The coloured classes are: high (blue), good (green), moderate (yellow), poor (orange) and bad (red; not occurring) status.

Water body	Monitoring station	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Ems Coast	**Huibertgat-oost**																					
Holland coast	**Noordwijk 2**																					
N. Delta coast	**Goeree 6/2**																					
Wadden coast	**Terslg 4/Boomkdp**																					
Zeeland coast	**Walcheren 2**																					

3.2.2 *Phaeocystis*

The assessment results for *Phaeocystis* per coastal water bodies for the years between 1990 and 2010 are expressed in Ecological Quality Ratios (EQRs), according to Figure 2.5 and given in Table 3.7.

Table 3.7 WFD assessment for Phaeocystis of the Dutch coastal water bodies for the years 1990-2010. The coloured classes are: high (blue), good (green), moderate (yellow), poor (orange) and bad (red; not occurring) status.

Water body	Monitoring station	199	199	199	199	199	199	199	199	199	199	200	200	200	200	200	200	200	200	200	200	201
Ems Coast	Huibertgat-oost																					
Holland coast	Noordwijk 2																					
N. Delta coast	Goeree 6/2																					
Wadden coast	Terslg 4/Boomkdp																					
Zeeland coast	Walcheren 2																					

3.2.3 Final assessment of phytoplankton

In Table 3.8 the final assessments for phytoplankton, which are the combined assessments for chlorophyll-a (Table 3.6) and *Phaeocystis* (Table 3.7), according to Chapter 2.2.3 are given. To illustrate the calculation of the final EQR two examples are given:

In year 2010 the assessment of *Phaeocystis* is lower than that of chlorophyll-a. This implies that the final assessment is the mean of both assessments, hence lower than the chlorophyll assessment.

EQR_{chl} = 0.93 and EQR_{phyto} = 0.51

Mean of EQR_{chl} and EQR_{phyto} = $(EQR_{chl}+EQR_{Phaeo})/2$ = 0.72

EQR_{phyto} = Minimum($(EQR_{chl}+EQR_{Phaeo})/2$, EQR_{chl}) = Minimum(0.72, 0.93)=0.72

In the year 2009 the assessment of *Phaeocystis* is higher than that of chlorophyll-a. This implies that the final assessment is equal to the chlorophyll-a assessment.

EQR_{chl} = 0.33 and EQR_{phyto} = 0.61

Mean of EQR_{chl} and EQR_{phyto} = 0.47

EQR_{phyto} = Minimum($(EQR_{chl}+EQR_{Phaeo})/2$, EQR_{chl}) = Minimum(0.47, 0.33)=0.33

Table 3.8 WFD assessment for phytoplankton of the Dutch coastal water bodies for the years 1990-2010. The coloured classes are: high (blue), good (green), moderate (yellow), poor (orange) and bad (red; not occurring).

Water body	Monitoring station	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005		2006	2007	2008	2009	2010
Ems Coast	Huibertgat-oost																						
Holland coast	Noordwijk 2																						
N. Delta coast	Goeree 6/2																						
Wadden coast	Terslg 4/Boomkdp																						
Zeeland coast	Walcheren 2																						

There is no clear pattern in the assessments expressed as classes. Figure 3.1 shows the final assessment as EQRs in a graph. This graph clearly shows the large interannual variation, which masks any trend. In the WFD the results are averaged over the growing seasons of 6 years (Figure 3.2). In the graph with EQRs averaged over six succeeding years any long-term change can easily be spotted.

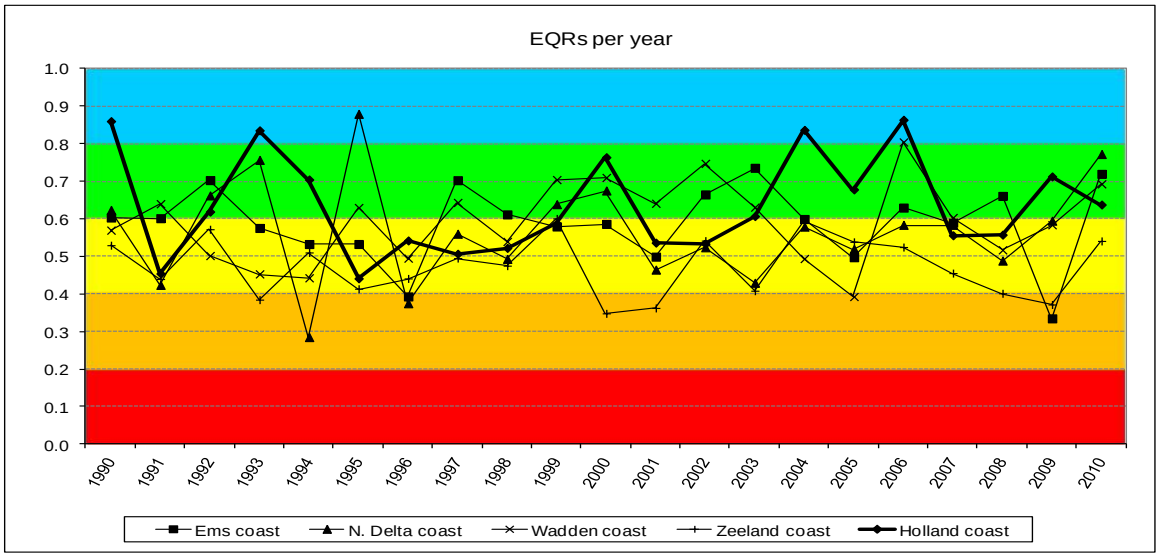

Figure 3.1 Graphical representation of the WFD results in EQRs for five Dutch coastal water bodies. The colours of the background represent the classes: high (blue), good (green), moderate (yellow), poor (orange) and bad (red).

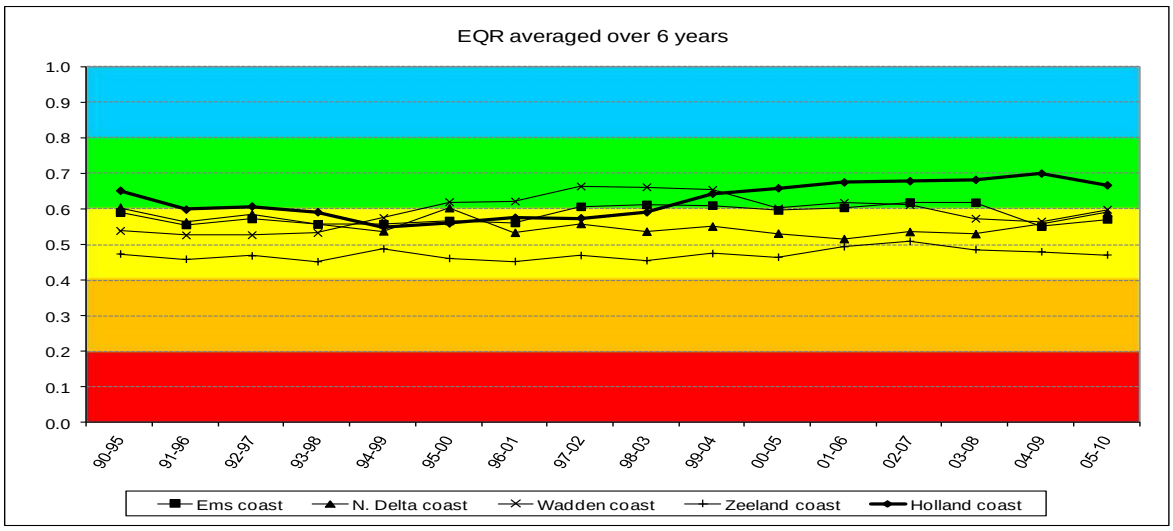

Figure 3.2 Graphical representation of WFD results, averaged over the growing seasons of six years, in EQRs for five Dutch coastal water bodies. The colours of the background represent the classes: high (blue), good (green), moderate (yellow), poor (orange) and bad (red).

4 OSPAR and WFD assessments: differences and similarities

4.1 Comparison of the assessment of the phytoplankton status according to OSPAR and WFD

In Table 4.1 the differences between the Dutch WFD and the OSPAR assessment methods are given.

Table 4.1 Definitions according to WFD and OSPAR for the Dutch marine waters

	WFD	**OSPAR**
Target areas	**marine water bodies:** Coastal waters within 1 nautical mile from the coast, divided into: Zeeland coast Northern Delta Coast Holland coast Wadden Coast Ems Coast	**marine areas:** whole Dutch Continental Shelf, divided into: Coastal waters (Sal. <34.5) Southern Bight (Sal. >34.5) Oyster Grounds (Sal. >34.5) Dogger Bank (Sal. >34.5)*
Data	Data of only one station per water body have been used over the growing season.	Data of all stations in an area have been used over the growing season.
Period	March to September (incl).	March to September (incl).
Chlorophyll-a	**Parameter:** 90-percentile **Calculation:** In general more samples have been taken during the summer months. To avoid overrepresentation of months with more than one sample, monthly means have been calculated per station/area. From these monthly means the 90-percentile value has been calculated. **Thresholds:** area-specific. The g/m boundary is 50% above the h/g boundary, (see Figure 2.4).	**Parameter:** mean and 90-percentile **Calculation:** calculated as mean value and 90-percentile of all samples in all relevant months in all stations in the target area. **Threshold:** area-specific, 50% above the "background value". The threshold between non-problem and problem area is equivalent to the WFD g/m boundary.
Phaeocystis	**Criterion:** The frequency of normal blooms per year. A normal bloom has been defined as a bloom with a concentration above 10^6 cells/l. **Threshold:** 2 months/year: boundary between good and moderate status.	**Criterion:** The number of cells/l of extreme blooms. An extreme bloom has been defined as a bloom with a concentration above 10^7 cells/l **Threshold:** 10^7 cells/l: boundary between no problem and problem area.
Other criteria	None; nutrients are used as supporting quality elements.	Other indicator species, oxygen; nutrients are used as supporting criteria.
Final judgement	The minimal value of the chlorophyll score and the mean value of the chlorophyll and *Phaeocystis* scores.	Problem area when there is at least one direct or indirect effect. Non-problem are when there are no direct or indirect effects. See for further details Table 2.3. Trends are indicated.

*Other OSPAR areas are the Wadden Sea, the western Scheldt and the Ems-Dollard estuary.

4.2 The OSPAR assessment applied to individual Dutch monitoring stations in the North Sea

The stations that WFD and OSPAR have in common are given in Table 4.2. Note that the OSPAR assessment is based on chlorophyll-a and *Phaeocystis* per area (*cf* Table 2.4), which implies that the assessment levels are valid for the whole area, which contains several WFD water bodies.

As WFD distinguishes five classes, while OSPAR has only two classes, the assessments of the WFD stations have been aggregated into two classes as well, with the boundary good/moderate between the two classes, which is equivalent to the OSPAR assessment level between the status Non Problem Area and Problem area.
Although the WFD boundary *good/moderate* for chlorophyll is identical to the OSPAR assessment level *NPA/PA* for the stations Walcheren 2, Terschelling 4/Boomkensdiep and Huibertgat oost the calculations of the 90-percentile are not identical. OSPAR uses all data available directly, while WFD first calculates the mean chlorophyll value per month. The reason for this is to avoid that months with more observations bias the 90-percentile. With respect to the marine target areas the following comparison can be made: OSPAR coastal waters with the combined score of the WFD coastal areas: Zeeland coast, Northern Delta Coast, Holland coast, Wadden Coast and Ems Coast.

Of course this comparison can be made for each of the years 1990 to 2010, but as has been shown in Figure 3.1 the interannual variation is large and more can be seen in an assessment over a period of six years. With a running mean this implies that there are still 17 comparisons to be made. To reduce this number we have chosen three successive periods of six years. To include the most recent year, 2010, these periods are: 1993 – 1998, 1999 – 2004 and 2005-2010.
In Table 4.3-4.5 the assessments by WFD (2^{e} column) and by OSPAR (4^{e} column) are given for the three periods, respectively. OSPAR distinguishes non-problem areas (blue) and problem areas (red). In the middle column the “translation” of the five WFD classes into two classes, with the OSPAR colours has been given.

Table 4.2 Dutch monitoring stations that WFD and OSPAR have in common with their WFD good/moderate class boundaries and the OSPAR assessment levels.

Transect	Monitoring station	OSPAR area	Chlorophyll-a 90 percentile (µg/l)	
			WFD boundary good/moderate	OSPAR assessment level
Walcheren	Walcheren 2	Coastal waters	15	15
Goeree	Goeree 6/Goeree2[6]	Coastal waters	21	15
Noordwijk	Noordwijk2	Coastal waters	21	15
Terschelling	Terschelling 4/ Boomkensdiep[7]	Coastal waters	15	15
Rottumerplaat	Huibertgat oost	Coastal waters	15	15

[6] *From 2007 on the new station Goeree 2 will be used for the WFD, because Goeree 6 is outside the 1-nautical mile zone .*
[7] *In 2007 Terschelling 4 has been replaced by Boomkensdiep.*

Table 4.3 Assessment by WFD and OSPAR for the six-year period from 1993-1998.

1993-1998	WFD Per station		WFD in OSPAR colours		OSPAR per station			OSPAR Coastal waters
Chlorophyll-a score								
ED coast - Huibertgat oost								
Wadden coast - Terschelling4								
Holland coast - Noordwijk2								
N Delta coast - Goeree 6								
Zeeland coast - Walcheren 2								
***Phaeocystis* score**								
ED coast - Huibertgat oost								
Wadden coast - Terschelling4								
Holland coast - Noordwijk2								
N Delta coast - Goeree 6								
Zeeland coast - Walcheren 2								
Final score*								
ED coast - Huibertgat oost								
Wadden coast - Terschelling4								
Holland coast - Noordwijk2								
N Delta coast - Goeree 6								
Zeeland coast - Walcheren 2								

*The final score for WFD is the lowest value of on the one hand the mean of the *Phaeocystis* and chlorophyll score and on the other hand the chlorophyll score. For OSPAR see Table 2.3. In this comparison only the chlorophyll concentrations and the *Phaeocystis* occurrence have been taken into account.

Table 4.4 Assessment by WFD and OSPAR for the six-year period from 1999-2004.

1999-2004	WFD Per station		WFD in OSPAR colours		OSPAR per station		OSPAR Per area
Chlorophyll-a score							
ED coast - Huibertgat oost							
Wadden coast - Terschelling4							
Holland coast - Noordwijk2							
N Delta coast - Goeree 6							
Zeeland coast - Walcheren 2							
***Phaeocystis* score**							
ED coast - Huibertgat oost							
Wadden coast - Terschelling4							
Holland coast - Noordwijk2							
N Delta coast - Goeree 6							
Zeeland coast - Walcheren 2							
Final score*							
ED coast - Huibertgat oost							
Wadden coast - Terschelling4							
Holland coast - Noordwijk2							
N Delta coast - Goeree 6							
Zeeland coast - Walcheren 2							

*The final score for WFD is the lowest value of on the one hand the mean of the *Phaeocystis* and chlorophyll score and on the other hand the chlorophyll score. For OSPAR see Table 2.3. In this comparison only the chlorophyll concentrations and the *Phaeocystis* occurrence have been taken into account.

Table 4.5 Assessment by WFD and OSPAR for the six-year period from 2005-2010.

2005-2010	WFD Per station		WFD in OSPAR colours		OSPAR per station		OSPAR Per area
Chlorophyll-a score							
ED coast - Huibertgat oost							
Wadden coast - Boomkensdiep							
Holland coast - Noordwijk2							
N Delta coast - Goeree 2							
Zeeland coast - Walcheren 2							
***Phaeocystis* score**							
ED coast - Huibertgat oost							
Wadden coast - Boomkensdiep							
Holland coast - Noordwijk2							
N Delta coast - Goeree 2							
Zeeland coast - Walcheren 2							
Final score*							
ED coast - Huibertgat oost							
Wadden coast - Boomkensdiep							
Holland coast - Noordwijk2							
N Delta coast - Goeree 2							
Zeeland coast - Walcheren 2							

*The final score for WFD is the lowest value of on the one hand the mean of the *Phaeocystis* and chlorophyll score and on the other hand the chlorophyll score. For OSPAR see Table 2.3. In this comparison only the chlorophyll concentrations and the *Phaeocystis* occurrence have been taken into account.

5 Neighbouring countries

5.1 OSPAR

Within OSPAR the countries around the North-East Atlantic have been collaborating since 1998. Together they developed the Comprehensive Procedure, as described above. This report concerns the indicators of the direct effects of eutrophication (category II). These indicators are chlorophyll-a and *Phaeocystis.* The relevant parameters for the indicator chlorophyll-a are the mean and the 90-percentile concentration of chlorophyll during the growing season, which for the North Sea is set to March – September inclusive (NL, BE, UK, DE) or March to October inclusive (FR). For *Phaeocystis* the parameter is the maximum number of cells/l in the whole year. Analytical methods for measuring chlorophyll and counting *Phaeocystis* by the Netherlands and the other countries are given in Annex A.

5.1.1 OSPAR – chlorophyll-a

Chlorophyll-a is assessed for OSPAR by the Netherlands, Belgium, Germany, UK and France. In Table 5.1 the assessment levels used by the various countries are shown (OSPAR, 2009). Within OSPAR the HASEC (Hazardous substances and eutrophication Committee) is working on a harmonization of the assessment levels (OSPAR, 2011).

Table 5.1 OSPAR area-specific assessment levels for 90-percentile and mean chlorophyll-a concentrations (µg/l) in coastal and offshore waters.

Chlorophyll-a (µg/l)	The Netherlands	Germany	Belgium	United Kingdom	France
coast					
90-percentile	15	6*	15	15	15
mean	7.5	3	7.5		--
(max)		(14)			
Offshore					
90-percentile	4.5	4.6*	8.4	10	15
mean	2.25	2.3	4.2		--
(max)		(9)			
Growing season	Mar-Sep	Mar-Sep	Mar-Oct**	Mar-Sep	Mar-Oct

*To be verified by Germany

5.1.2 OSPAR – *Phaeocystis*

Phaeocystis is assessed by the Netherlands, Belgium and Germany, but not by UK and France. Table 5.2 shows the assessment levels used by these countries (OSPAR, 2009).

Table 5.2 OSPAR assessment levels for Phaeocystis (cells/l). In coastal and offshore waters the same levels are valid.

Phaeocystis (cells/l)	The Netherlands	Germany	Belgium	United Kingdom	France
coast and offshore					
maximum	10^7 and duration	10^6	10^7 and >30 days[8]	no indicator	no indicator

[8] *This value is in revision and will be replaced by 4.10^6 cells/l in Belgium and probably also in the Netherlands according to a study of Lancelot et al., 2009.*

5.1.3 OSPAR – Final assessment on the national borders

NL-BE The OSPAR areas on the border between the Dutch and the Belgian continental shelf are for the Netherlands the Dutch Coastal Waters and the Southern Bight (Baretta-Bekker et al., 2008) and for Belgium the coastal area and the offshore area (Anonymous, 2008). Table 5.3 shows the results.

Table 5.3 Comparison of the combined OSPAR assessments of chlorophyll-a and Phaeocystis of the Dutch and Belgian areas on the Dutch-Belgian border over the 6-years period: 2001-2005. PA: Problem Area; (PA): Problem Area, not based on quantitative monitoring data, but on scattered scientific information; PPA: Potential Problem Area, not enough data to perform an assessment. ***NB This combined assessment is not necessarily identical to the overall assessment, using all indicators!***

Country	Water body	Assessment chlorophyll	Assessment *Phaeocystis*	Final assessment
NL	NL-Coastal waters	PA	PA	PA
BE	BE-Coastal area	PA	(PA)	PA
NL	NL-Southern Bight	PA	PA	PA
BE	BE-Offshore area	PPA	--	PPA

BE-FR In contrary to the other countries France does not use the OSPAR criterion for the division between coastal and offshore waters (Sal < 34.5 and > 34.5, respectively), but divides the sea according to the WFD:

1. baseline – 1 nautical mile
2. 1-12 nautical miles and
3. beyond 12 nautical miles.

1 and 2 together could be compared with the OSPARs coastal areas, and 3 is comparable with OSPAR's offshore areas.

The OSPAR areas on the border between the Belgian and the French continental shelf are for Belgium the coastal area and the offshore area (Anonymous, 2008) and for France, the coastal zones (0-1 and 1-12 nm) and the area beyond 12 nm (Anonymous, 2008 - OSPAR COMP 2, 2008). Table 5.3a shows the results.

Table 5.3a Comparison of the combined OSPAR assessments of chlorophyll-a and Phaeocystis of the Belgian and French areas on the Belgian-French border over the 6-years period: 2001-2005. PA: Problem Area; (PA): Problem Area, not based on quantitative monitoring data, but on scattered scientific information; PPA: Potential Problem Area, not enough data to perform an assessment. ***NB This combined assessment is not necessarily identical to the overall assessment, using all indicators!***

Country	Water body	Assessment chlorophyll	Assessment *Phaeocystis*	Final assessment
BE	BE-Coastal area	PA	(PA)	PA
FR	FR 0-1 nm FR 1-12 nm	PA --	-- --	PPA NPA*
BE	BE-Offshore area	PPA	--	PPA
FR	FR- > 12 nm	--	--	NPA**

**: expert opinion; **: French water beyond 12 nm are considered NPA (screening procedure)*

NL-DE The OSPAR areas on the border between the Dutch and the German continental shelf are for the Netherlands the Dutch Coastal Waters, the Oyster Grounds and the Dogger Bank (Baretta-Bekker et al., 2008) and for Germany the coastal area and the offshore area (Brockmann et al., 2008). Table 5.4 shows the results.

Table 5.4 Comparison of the combined OSPAR assessments of chlorophyll-a and Phaeocystis of the Dutch and German areas on the Dutch-German border over the 6-years period: 2001-2005. PA: Problem Area; (PA): Problem Area, not based on quantitative monitoring data, but on scattered scientific information; NPA: Non Problem Area; PPA: Potential Problem Area, not enough data to perform an assessment. ***NB This combined assessment for phytoplankton is not necessarily identical to the overall assessment based on all indicators!***

Country	Water body	Assessment chlorophyll	Assessment *Phaeocystis*	Combined assessment
NL	NL-Coastal waters	PA	PA	PA
DE	DE-Coastal waters	PA	PA	PA
NL	NL-Oyster Grounds	NPA	NPA	NPA
DE	DE-Offshore water, inner part	NPA	?	?
NL	NL-Dogger Bank	NPA	NPA	NPA
DE	DE-Offshore water, outer part	NPA	?	?

NL-UK The OSPAR areas on the border between the Dutch and the British continental shelf are for the Netherlands the Southern Bight, the Oyster Grounds and the Dogger Bank (Baretta-Bekker et al., 2008) and for UK the offshore area (UK, 2008) Table 5.5.

Table 5.5 Comparison of the combined OSPAR assessments of chlorophyll-a and Phaeocystis of the Dutch and British areas on the Dutch-German border over the 6-years period: 2001-2005. PA: Problem Area; (PA): Problem Area, not based on quantitative monitoring data, but on scattered scientific information; PPA: Potential Problem Area, not enough data to perform an assessment; -- not used. ***NB This combined assessment for phytoplankton is not necessarily identical to the overall assessment based on all indicators!***

Country	Water body	Assessment chlorophyll	Assessment *Phaeocystis*	Combined assessment
NL	NL-Southern Bight	PA	PA	PA
UK	UK-Offshore water	NPA	--	NPA
NL	NL-Oyster Grounds	NPA	NPA	NPA
UK	UK-Offshore water	NPA	--	NPA
NL	NL-Dogger Bank	NPA	NPA	NPA
UK	UK-Offshore water	NPA	--	NPA

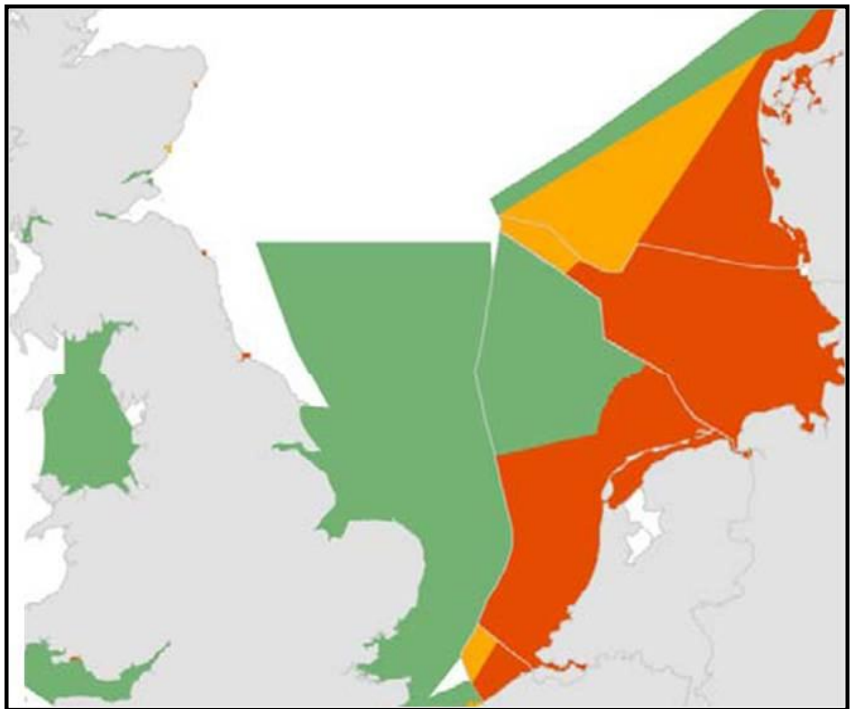

Figure 5.1 The OSPAR the final assessment, based on data of 2001-2005. Red: Problem Area; Orange: Potential Problem Area; Green: Non Problem Area. Source: OSPAR (2009). NB Only part of North of France is visible in this figure.

In Figure 5.1 The eutrophication status of the various area in the North Sea is shown. This status is based on all OSPAR eutrophication indicators and is not everywhere identical to the phytoplankton (combined chlorophyll – *Phaeocystis*) assessment. Differences are in the German areas.

5.2 Water Framework Directive

For the WFD all European countries have developed their own assessment systems for the quality elements as described in the directive. These assessment methods have been intercalibrated in the so-called Geographical Intercalibration Groups (GIGs). The West European countries discussed their methods in the North East Atlantic GIG. In general three metrics have been selected: chlorophyll, Indicator Taxa (Frequency of *Phaeocystis* cell counts) and Taxa Cell Counts (Frequency of phytoplankton taxa cells counts). The Netherlands and the neighbouring countries Belgium and Germany all use the indicators chlorophyll-a and *Phaeocystis*. Belgium has the composition and blooms of harmful species under development, UK takes the bloom frequency of any phytoplankton taxon into account as well as the seasonal succession of algal functional groups and France calculates, as UK does, the bloom frequency of any phytoplankton taxon. In the following comparison only the indicators chlorophyll-a and *Phaeocystis* are considered.

5.2.1 WFD – chlorophyll-a

The intercalibration process resulted in a harmonized tool and class boundaries for chlorophyll-a. The chlorophyll tool assumes that six years of chlorophyll data are available with monthly sampling during the growing season. For the North Sea the growing season has been defined as the period of seven months from March to September, inclusive (NL, DE, UK) or eight months from March to October, inclusive (FR). The assessment parameter is the 90-percentile of the chlorophyll-a concentration in µg/l. In the Netherlands all available data within one month are averaged, in order to ensure that the contribution of each month is equal, independent of the number of samples. It is not clear how the other countries handle this issue.

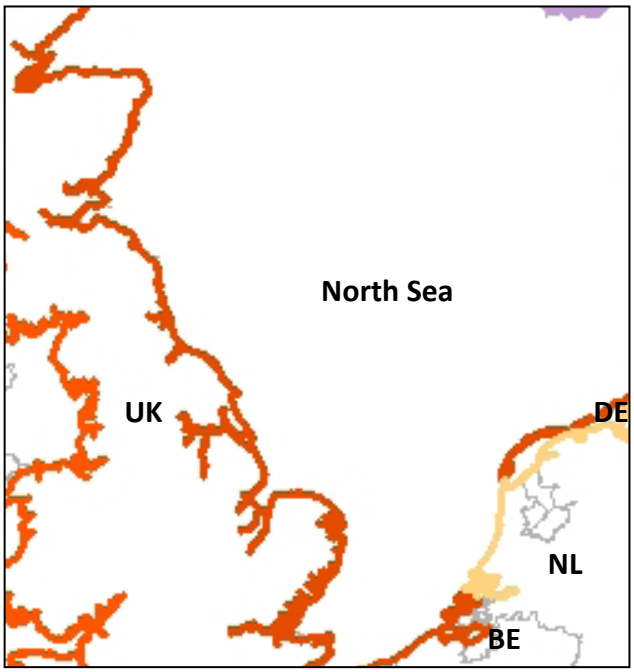

Figure 5.2 The coasts of the Netherlands, Belgium, Germany (partly) and UK. The orange coastal water bodies are from the type NEA 12b, the yellow are NEA 3/4.

The coastal water bodies have been divided in different types (Figure 5.2). The whole eastern coast of UK, the Belgium coast and the Dutch Zeeland coast and Wadden coast belong to the same type (NEA 1/26b - the euhaline type) and have the same class boundaries after the intercalibration process.
The Northern Delta coast, the Holland coast and the Ems coast belong to another type (NEA 3/4 - the polyhaline type). The Netherlands share this type with Germany. As the class boundaries Germany proposed for this type were much lower in the first phase of the intercalibration than those proposed by the Netherlands, the Dutch values were not accepted by the Commission. In the second phase Germany and the Netherlands came to a proposal, that has been accepted by the Netherlands, but not yet by the Commission (Davies, 2012): For the Ems coast the Netherlands reduced their threshold values, while Germany increased its thresholds for this area. The end result is shown in Figure 5.3. This was justified by the acknowledgement that the Ems area is a transition region between western and eastern waters.
For the intercalibration process the Netherlands and Germany agreed to recognise one water type NEA 3/4 from the Rhine Delta to the Eider with a gradient from west to east and to adjust the values in the Ems area, as this is the connection between the western and eastern part of the water type NEA 3/4.

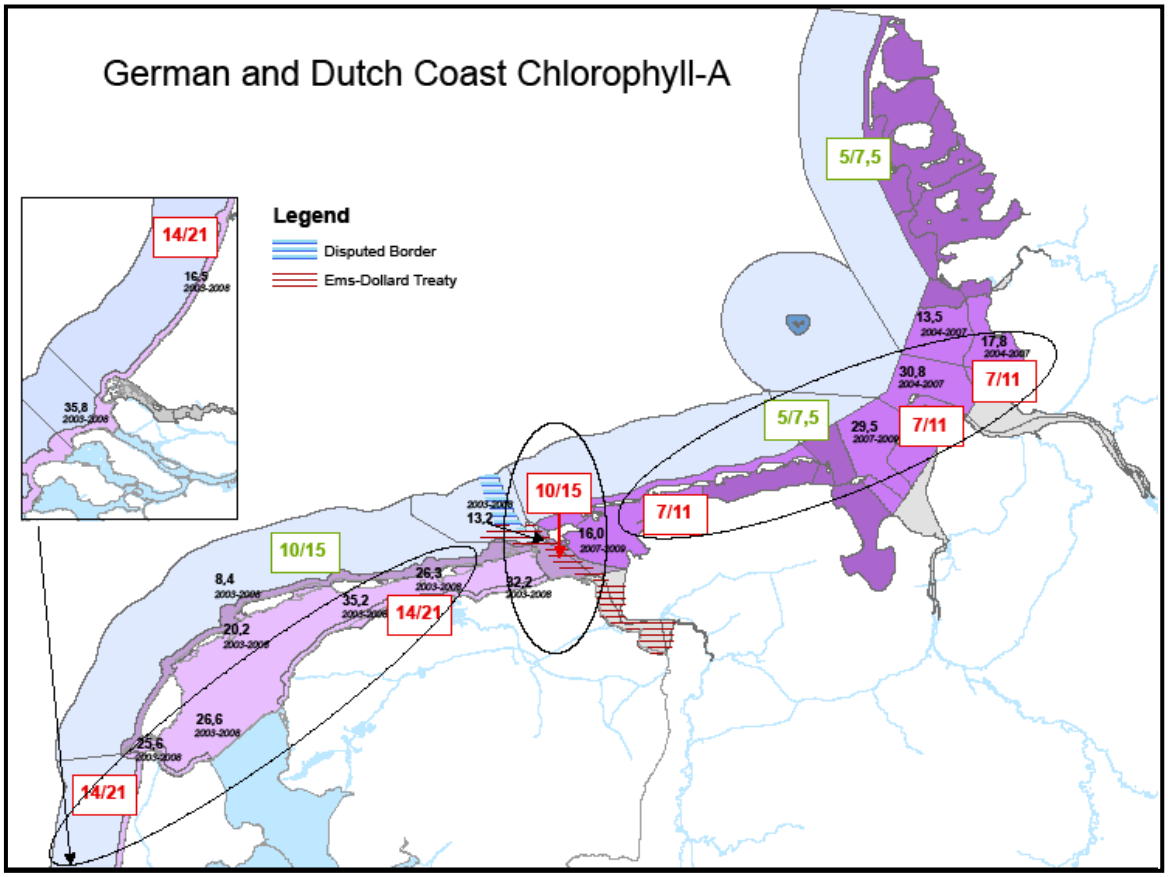

Figure 5.3 Intercalibration geographic area NEA 3/4 of the Netherlands and Germany with class boundaries high/good and good/moderate representing a gradient due to hydrological circumstances.

In Table 5.6 the class boundaries G/M for the Belgian, Dutch and German coastal waters are given. For NEA 3/4 there is a gradient as explained above. Note that the class boundaries of the Belgian and Dutch NEA 1/26b are identical to the UK class boundaries for the same type.

Table 5.6 All coastal water bodies in France, Belgium, the Netherlands and Germany, from south to north with the WFD class boundaries for 90-percentile chlorophyll (µg/l).

Country	Water body	type	G/M boundary
France	French northern coast	NEA 1/26b	15 µg/l
Belgium	Belgian coast	NEA 1/26b	15 µg/l
The Netherlands	Zeeland	NEA 1/26b	15 µg/l
	Northern Delta coast	NEA 3/4	21 µg/l
	Holland coast	NEA 3/4	21 µg/l
	Wadden coast	NEA 1/26b	15 µg/l
	Ems coast	NEA 3/4	15 µg/l
Germany	Weser coast	NEA 3/4	11 µg/l
	Elbe coast	NEA 3/4	11 µg/l
	Eider coast	NEA 3/4	11 µg/l

5.2.2 WFD - *Phaeocystis*

The classification process tool for *Phaeocystis,* harmonized in the intercalibration process, is based on the percentage of months with one or more occasions where the counts of *Phaeocystis* exceed a predetermined threshold.

The definition of the metric is according to Carletti & Heiskanen (2009)[9]:

- The number of months with samples where the *Phaeocystis* cell count is over a predefined threshold
- Six years of routine monitoring data should be used from the whole year
- A minimum of 12 sampling occasions (monthly) per year is recommended (but not essential).
- Classification is based on the number of months with one or more samples where the *Phaeocystis* counts exceed a threshold and is calculated as the percentage of the total number of months in the six year period (72).

The Netherlands, Germany and Belgium all use this tool with the same metric for *Phaeocystis*, as described for the Netherlands in Ch. 2.1.4. A difference between the Netherlands and Belgium is, that the Netherlands samples during the growing season (seven months), assuming that no blooms occur between October and March, while Belgium samples during the whole year, which makes their assessment slightly more accurate (see Ch. 6.5).

As a *Phaeocystis* spring bloom is a natural phenomenon, a bloom duration of one month (8,3%) is assumed to be normal and the class boundaries high/good and good/moderate are set at 10% and 17%, respectively.

UK uses a different metric which does not take into account *Phaeocystis* specifically, but uses the abundance of any one species. The metric is the frequency (%) of blooms of any single phytoplankton taxon above the bloom threshold.

France uses the bloom frequency of any phytoplankton taxon (bloom threshold : 100,000 cells/l) and bloom frequency of harmful phytoplankton taxa (bloom threshold : 1,000,000 cells/l).

N.B. Carletti & Heiskanen (2009) state in their technical report:

These boundaries equate to H/G and G/M EQRs of 0.83 and 0.49 calculated by dividing e.g. the reference percentage (8.3) by the boundary percentages, but this is not true for the Netherlands, because NL uses normalized EQRs, i.e. the boundary between high and good is normalized to 0.8, the boundary between good and moderate is normalized to 0.6 etc. This aspect is neglected in the whole intercalibration process, but may have consequences!!

[9] *As noted earlier the intercalibration results of Carletti & Heiskanen, 2009 are not finally valid, because they did not follow an accurate comparison protocol and need to be re-evaluated with deadline 2016.*

5.2.3 WFD – Final assessment around the national borders

The phytoplankton assessment of the water bodies on the border between the 1-nautical mile coastal zone of the Netherlands and Belgium , Belgium and France and the Netherlands and Germany are shown in Tables 5.7, 5.8 and 5.9, respectively.

Table 5.7 NL-BE Comparison of the WFD assessments of the Dutch Zeeland coast and the Belgian coast over the 6-years period: 2005-2010.

Country	**Water body**	**Period**	**Assessment chlorophyll**	**Assessment *Phaeocystis***	**Final assessment**
NL	Zeeland coast	2005-2010	moderate	good	moderate
BE	Belgian coast	2005-2010	moderate	moderate	moderate

Table 5.8 BE-FR Comparison of the WFD assessments of the Belgian coast and the French coastal area FRAC02 over different periods.

Country	**Water body**	**Period**	**Assessment chlorophyll**	**Assessment *Phaeocystis***	**Final assessment**
BE	Belgian coast	2005-2010	moderate	moderate	moderate
FR	FRAC02	2003-2008	good	moderate	moderate

Table 5.9 NL-DE Comparison of the WFD assessments of the Dutch Ems coast and the German Weser coast over the 6-years period: 2005-2010.

Country	**Water body**	**Period**	**Assessment chlorophyll**	**Assessment *Phaeocystis***	**Final assessment**
NL	Ems coast	2005-2010	good	moderate	moderate
DE	Weser coast	2005-2010	?	?	?

6 Sensitivity of indicator performance to methodological aspects

The indicators for chlorophyll and *Phaeocystis*, as used in the OSPAR COMP and in the WFD assessment of biological status are based on the same parameters (*viz.* chlorophyll-a and *Phaeocystis* concentrations), but differ in the metrics that are used to determine the indicator value (Table 1, Ch 5.1).
Moreover, the OSPAR and WFD assessments use different sampling frequencies to calculate the metric (Ch 5.1), sampling frequencies differ between sampling stations, and various other methodological issues (for example, the observation period) affect the results of the metrics.
The indicator values and, consequently, assessment results may to some extent be affected by these issues. Hence, the precision and confidence of the indicator, the risk of misclassification (EC, 2005) and the effects on the comparability of the OSPAR and WFD assessments have to be considered.
The next sections address the issues of sampling frequency and several other methodological issues, and compare the WFD and the OSPAR metrics for *Phaeocystis* and for chlorophyll-a.
In the following sections the concept of misclassification is mentioned. This concept is explained below.
Misclassification An indicator leads to a correct assessment when the observed indicator values accurately represent the status: In cases where the observed assessment values correctly indicate that the threshold is not exceeded and status is “good” (section A in Figure 6.1), or where the values show that the threshold value is exceeded and status is “not good” (section C in Figure 6.1). However, when the observed indicator value exceeds the threshold while the true value is below the threshold (section B in Figure 6.1), an incorrect status is assigned. Similarly, in the contrasting case (section D in Figure 6.1), the observed value incorrectly indicates that the ecosystem status is good and an incorrect status is assigned. The latter two cases (B, D) are considered misclassification.

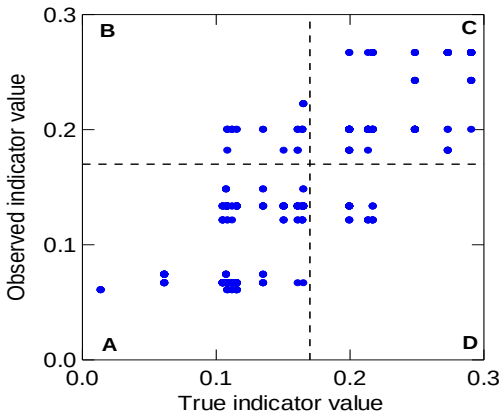

Figure 6.1 Example of observed and true indicator values in relation to an indicator threshold (dashed line), with observed values giving a correct classification (sections A and C) or a misclassification (sections B and D).

6.1 Metric values

The performance of the metrics for chlorophyll-a and *Phaeocystis* was analyzed using time series of chlorophyll-a and *Phaeocystis*, that were based on observations at station Noordwijk 10. Using an interpolation method, time series with daily values for chlorophyll-a and *Phaeocystis* were created. Details on the method to create this time series are given in Annex B. From this time series of interpolated values, data were sampled to simulate a sampling programme with samples taken at specific intervals (in this case, intervals of 1, 2, 3 and 4 weeks were used). These sampled data formed a subset (for example, with a value every 4 weeks) of the complete time series. On the

basis of these subsets, metric values were calculated that are considered “observed values” in this report, in contrast to the “true” values that are based on the complete time series (i.e. daily values). The maximum and average concentrations, and the 90-percentile value of chlorophyll-a during the growing season (March-September), calculated from the daily concentrations in the time series, are shown in Figure 6.2. The maximum and average cell numbers and the bloom frequency of *Phaeocystis*, calculated from the daily concentrations in the time series, are shown in Figure 6.3. The metric values presented in Figures 6.2 and 6.3 are assumed to represent the “true” values, and were used as a reference for the performance of the metric for other sampling frequencies.

The performance of the metrics at the various sampling intervals was evaluated by looking at the accuracy and the precision of the metric values. The accuracy is the degree of closeness of the observed values to the true value. If observed values do not show a bias, i.e. there is no consistent deviation from the true values and a regression of observed values against true values has a slope of 1, the metric can be considered accurate. The precision is the degree of reproducibility in the observed values. If a metric has a small random variability (error) it can be considered precise. We evaluated the accuracy and precision of the metrics by looking at the slope of the regression of observed values against true values, by looking at the coefficient of variation as a measure of the random error, and by looking at the percentage of observed values that were close (<25% difference) to the true values.

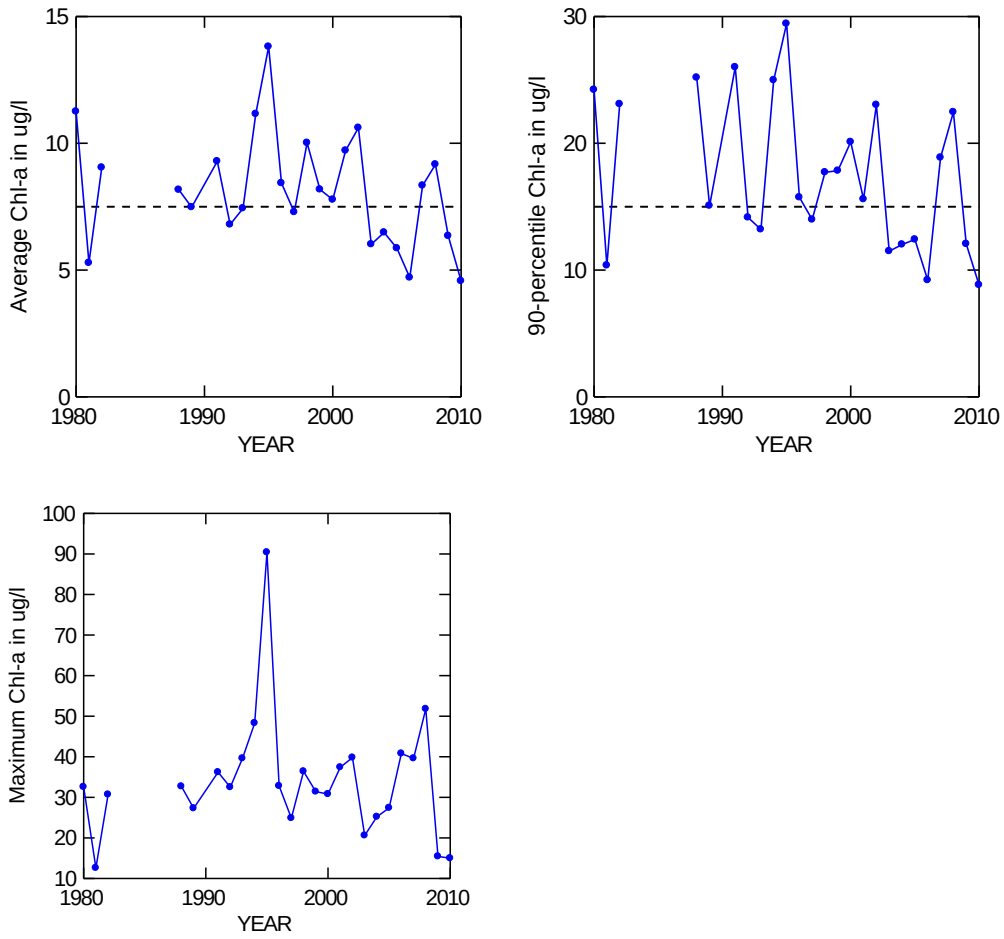

Figure 6.2 Average, 90-percentile and maximum chlorophyll-a concentrations, calculated from the interpolated daily concentrations. The dashed lines indicate the OSPAR/WFD thresholds for average and 90-percentile. For the years 1983-1987 no data were available.

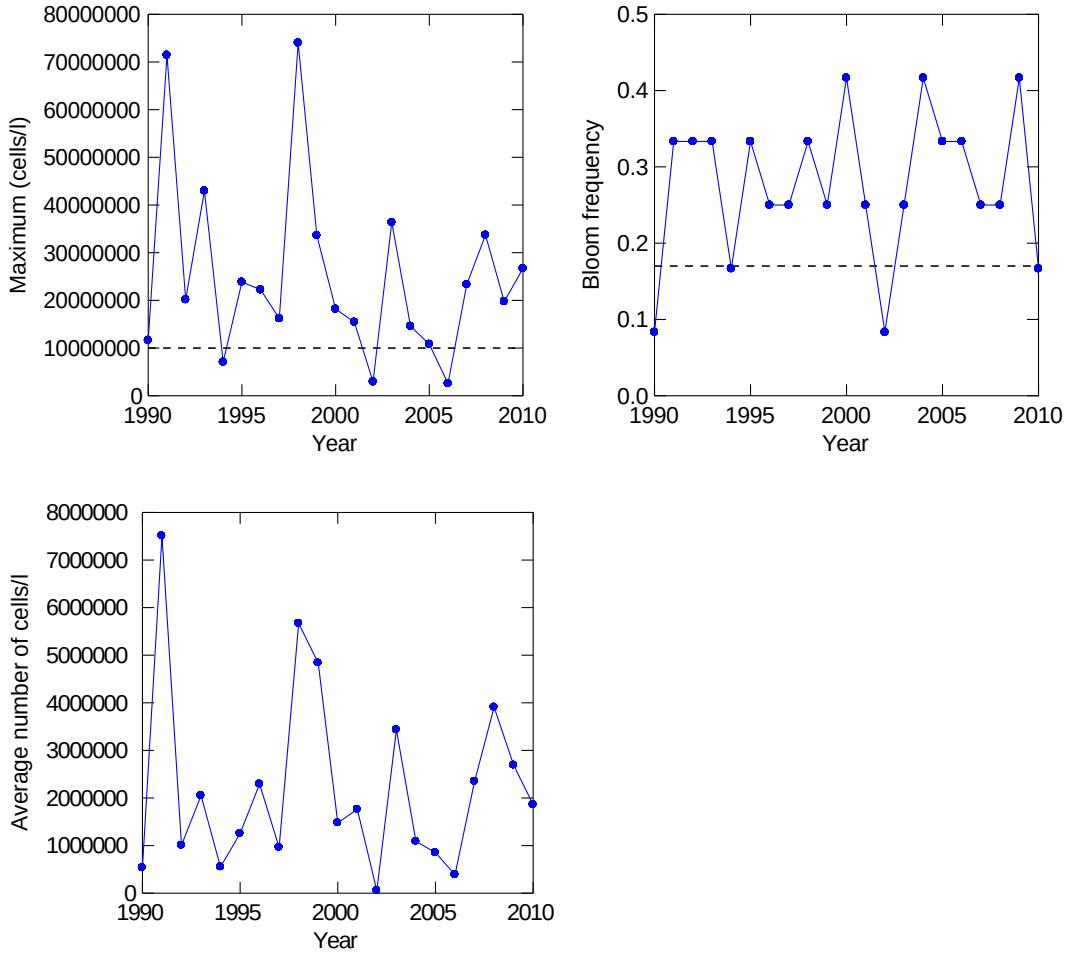

Figure 6.3. Maximum cell number, bloom frequency and average cell number of Phaeocystis, calculated from the interpolated daily concentrations. The dashed lines indicate the 10^7 cells/l level (maximum cell number) and the bloom frequency of 0.17.

6.2 Effects of sampling frequency

6.2.1 OSPAR *Phaeocystis* metric: Maximum cell numbers

The OSPAR indicator for *Phaeocystis* blooms uses the annual maximum cell number as metric, with maximum values exceeding the threshold of 10^7 cells/l taken as indicating a eutrophication problem. Maximum cell numbers showed considerable variation between years. Maximum cell numbers, determined with a sampling interval of 1, 2, 3 or 4 weeks were compared to the true maximum cell numbers (Figure 6.4). This comparison shows that, with increasing sampling intervals, the range in values becomes larger and the range also extends to increasingly lower values.

Figure 6.5 shows that, in this data set, in some cases misclassification may occur: when the true value for maximum cell number (as calculated from daily values, on the horizontal axis) is higher than 10^7 cells/l (and thus in terms of OSPAR are considered a bloom), the maximum cell number calculated from datasets with sampling intervals of 1-4 weeks occasionally are below 10^7 cells/l (not a bloom).

Table 6.1 gives the results of the comparison of observed values with the true values. These data show that the slope of regression of the observed values against true values is significantly lower than 1, and also decreases with lower sampling frequencies. This indicates that the accuracy of the metric is strongly affected by sampling frequency, and that even at a sampling interval of one week

there is a risk that observed values underestimate the maximum cell number. The precision of the metric also decreases with lower sampling frequencies, as illustrated by the increasing coefficient of variation. Overall, as illustrated by the percentage of observed values that deviate less than 25% from the true values, the performance of the metric is poor at sampling intervals of more than one week.
Table 6.2 gives the percentage of observations with misclassification, for the various sampling intervals. The risk of misclassification increases with lower sampling frequency.

Table 6.1 Results of the comparison of observed maximum cell numbers of Phaeocystis (OSPAR) at 1-, 2-, 3- or 4-week sampling interval values with true values. Slope and r2 of the linear regression of observed values against true values, mean coefficient of variation of the observed values and the percentage of observations with <25% difference from true values.

Sampling interval	Regression of observed versus true values		Slope significantly different from 1?	Coefficient of variation	observed values with <25% difference from true values
	slope	r^2			
1 week	0.82 ± 0.02	0.93***	p<0.001	0.14	69%
2 weeks	0.68 ± 0.02	0.80***	p<0.001	0.25	45%
3 weeks	0.58 ± 0.02	0.65***	p<0.001	0.37	34%
4 weeks	0.51 ± 0.02	0.56***	p<0.001	0.50	26%

Table 6.2 Percentage of cases of misclassification of the maximum cell number of Phaeocystis (OSPAR) at 1-, 2-, 3- or 4-week sampling interval.

Sampling interval	% misclassification
1 week	10%
2 weeks	19%
3 weeks	29%
4 weeks	37%

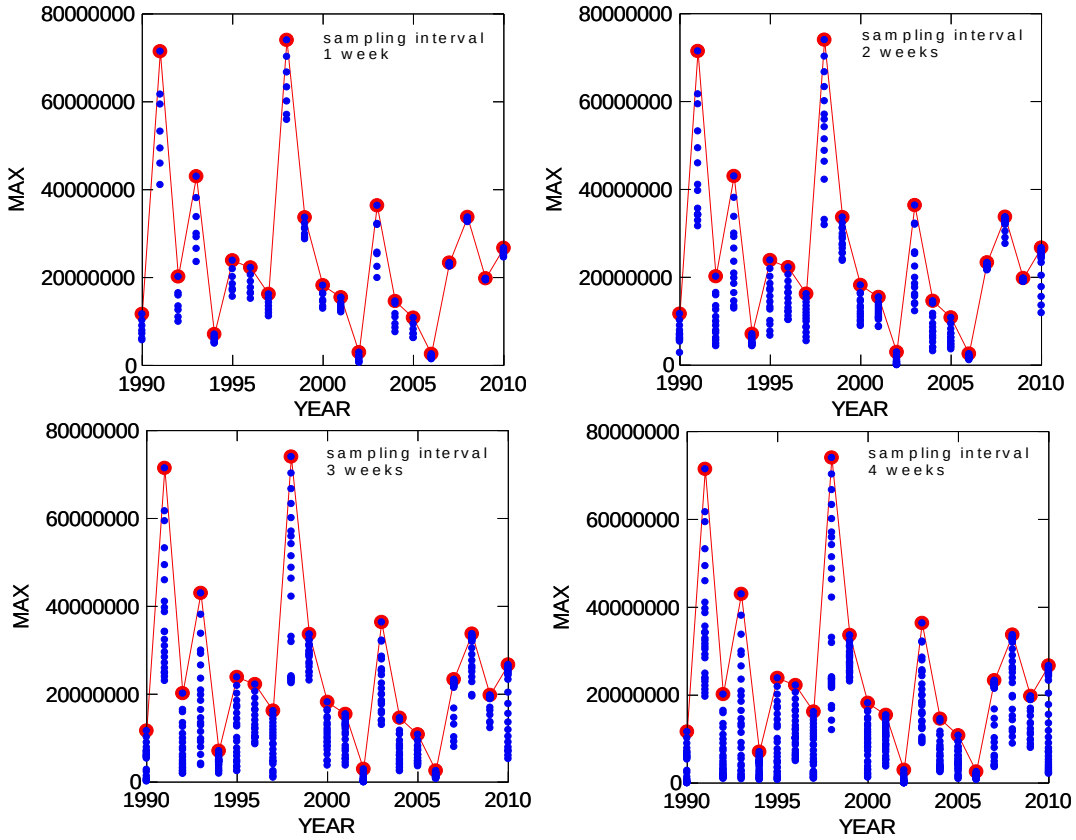

Figure 6.4. Annual maximum cell numbers of Phaeocystis at Noordwijk 10 determined from interpolated daily values (red line and red dots; also see Figure 6.3) and the annual maximum cell numbers of the different data sets that can be generated from the interpolated daily values with 1-, 2-, 3- or 4-week sampling intervals (blue dots).

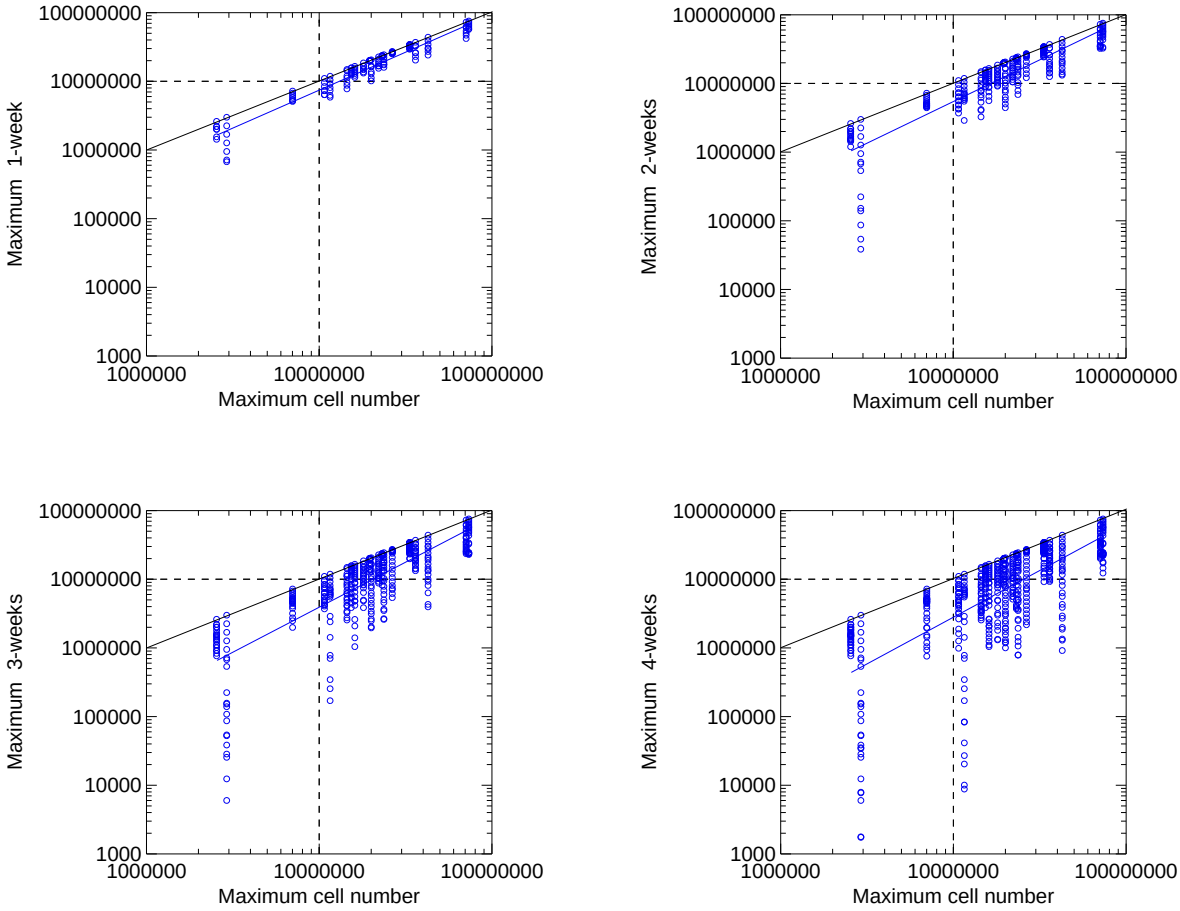

Figure 6.5. Maximum cell numbers of Phaeocystis in the different data sets that can be generated from the interpolated daily values with 1-, 2-, 3- or 4-week sampling intervals plotted against the "true" maximum value. The dashed lines indicate the level of 10^7 cells/l. The full blue line shows the linear regression line, the black line shows y=x.

6.2.2 WFD *Phaeocystis* metric: Bloom frequency

The WFD indicator for *Phaeocystis* blooms uses bloom frequency as a metric, with blooms exceeding the threshold 10^6 cells/l in more than 2 months (>17% of the year), assuming that one bloom per year is a normal phenomenon, indicating moderate status or worse. Similar to maximum cell numbers in the previous paragraph, bloom frequency showed considerable variation between years. In Figure 6.6 bloom frequencies, determined with a sampling frequency of 1, 2, 3 or 4 weeks were compared to the bloom frequency, determined from daily sampling (Figure 6.3). It must be realized, that bloom frequency determined from daily values may lead to overestimation of the real bloom frequency. As the metric values in this case are based on daily values, even if on only one day in a month a bloom (concentration $>10^6$ cells/l) is observed, this month is flagged as a bloom month. This means that the bloom frequency calculated on the basis of daily samples, may seriously overestimate the actual frequency of blooms of *Phaeocystis* (see Ch 6.5.1).

This comparison shows that, with increasing sampling intervals, the observed values tend to become lower (Figure 6.7).

Table 6.3 gives the results of the comparison of observed bloom frequencies with the true values. There is a bias in the observed values, as the slope of observed values against true values is lower than 1. This bias increases with increasing sampling interval, indicating that the observed frequencies change with sampling interval. The precision of the observed values is in the same range as the one for the metric maximum cell number (compare with Table 6.1).

Figure 6.7 shows that in some cases in this dataset misclassification occurs. Table 6.4 gives the percentage of observations with misclassification, for the various sampling frequencies. The risk of misclassification increases with longer sampling intervals.

Table 6.3 Results of the comparison of observed bloom frequencies of Phaeocystis (WFD) at 1-, 2-, 3- or 4-week sampling interval values with true values. Slope and r2 of the linear regression of observed values against true values, mean coefficient of variation **of the observed values and the percentage of observations with <25% difference from true values.**

Sampling interval	Regression of observed versus true values		Slope significantly different from 1?	Coefficient of variation	observed values with <25% difference from true values
	slope	r^2			
1 week	0.58 ± 0.05	0.48***	p<0.001	0.21	61%
2 weeks	0.36 ± 0.04	0.21***	p<0.001	0.31	31%
3 weeks	0.31 ± 0.03	0.14***	p<0.001	0.45	20%
4 weeks	0.25 ± 0.03	0.12***	p<0.001	0.46	12%

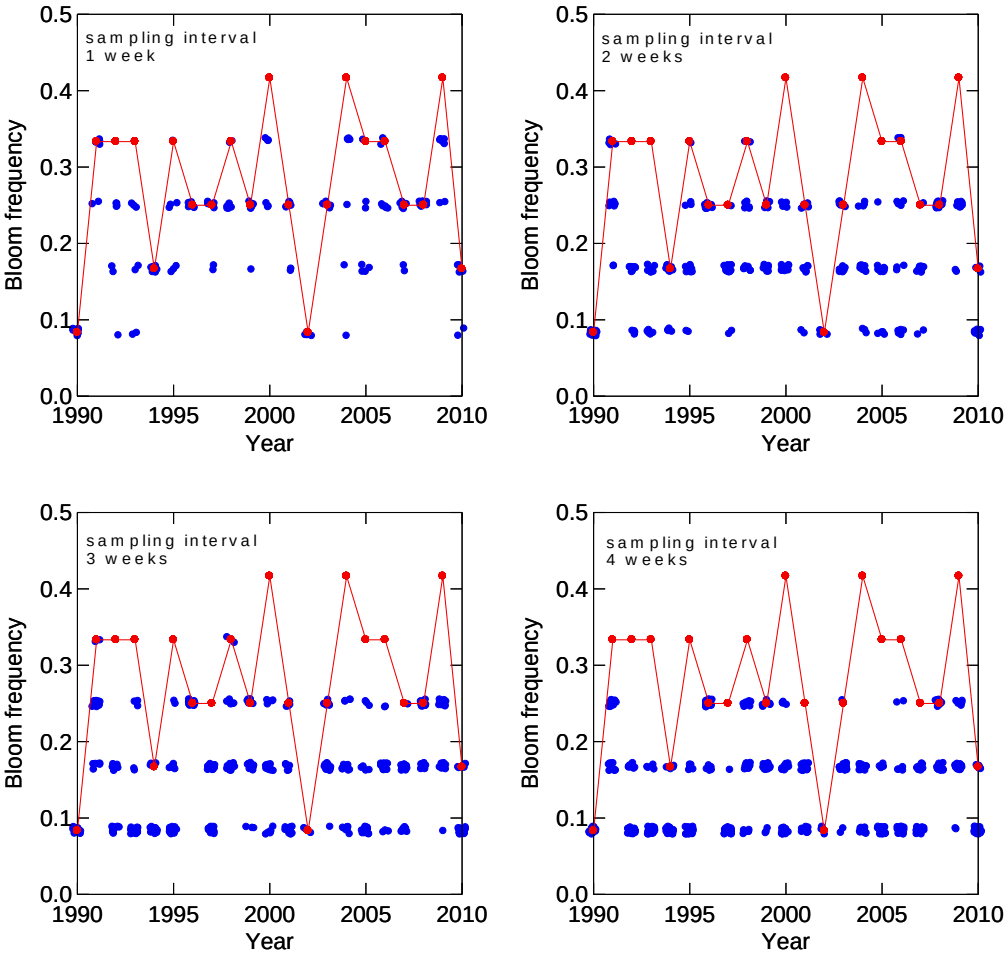

Figure 6.6 Bloom frequency determined from interpolated daily values (red line and red dots; also see Figure 6.3) and frequencies determined with 1-, 2-, 3- or 4-week sampling intervals (blue).

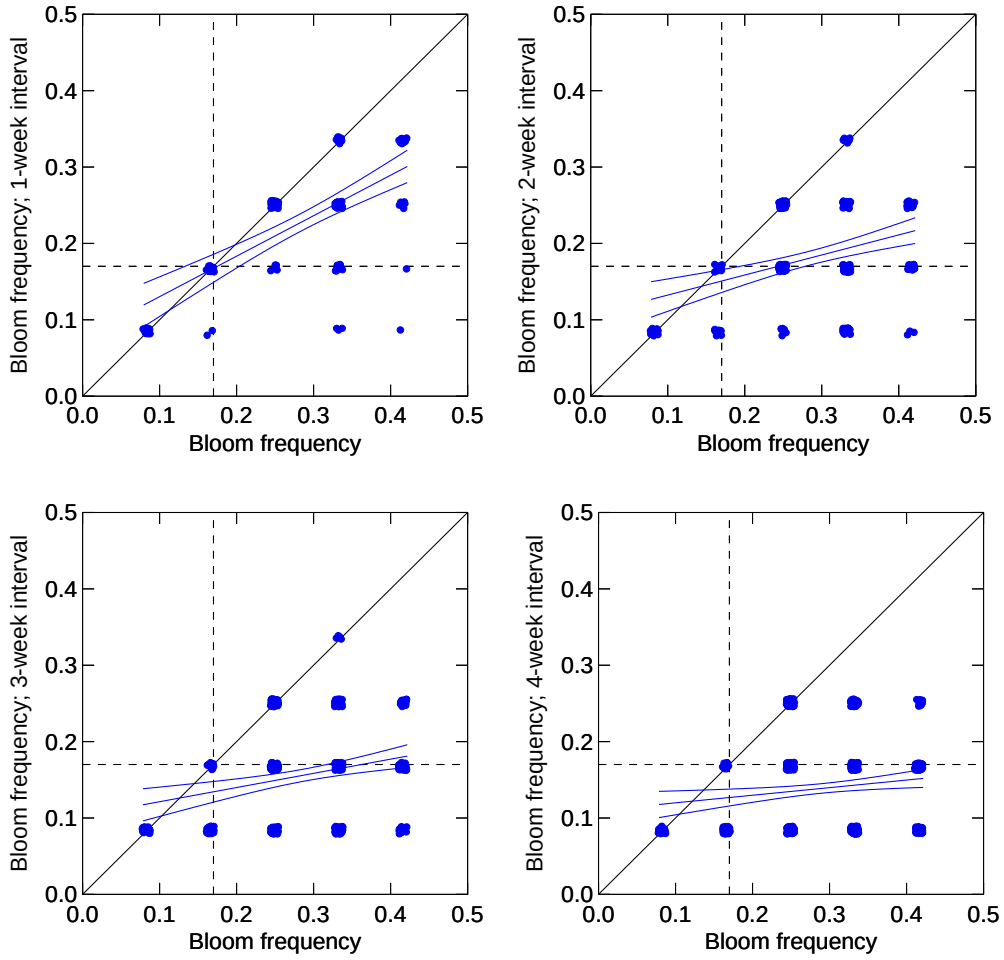

Figure 6.7 Bloom frequency determined with 1-, 2-, 3- or 4-week sampling intervals plotted against the true bloom frequency. The dashed lines indicate the level of 10^7 cells/l. The blue line shows the linear regression line, the black line shows y=x.

Table 6.4 Percentage of cases of misclassification of the bloom frequency of Phaeocystis (WFD) at 1, 2, 3 and 4 week sampling interval.

Sampling interval	% misclassification
1 week	18%
2 weeks	51%
3 weeks	62%
4 weeks	74%

6.2.3 Chlorophyll-a metric

6.2.3.1 Average concentrations

Average concentrations showed considerable variation between years (Fig. 6.2). When comparing the average concentration, determined with a sampling interval of 1, 2, 3 or 4 weeks, to the true averages it becomes clear that, with increasing sampling intervals, the range in values becomes larger but the values are scattered around the true value (Figure 6.8). This is illustrated in Figure

6.9 that shows the distribution of observed values in relation to the true average value. With a lower sampling frequency, larger deviations from the true values are observed.
Table 6.5 shows that the observed values agree well with the true values , indicating the high accuracy of the metric at all sampling intervals. The precision decreases slightly with longer sampling intervals. Overall, the metric has a good performance, even at relatively long sampling intervals.
The risk of misclassification, in this data set, increases slightly when the sampling interval becomes longer (Table 6.6).

Table 6.5 Results of the comparison of observed average chlorophyll-a concentrations (OSPAR) at 1-, 2-, 3- or 4-week sampling interval values with true values. Slope and r2 of the linear regression of observed values against true values, mean coefficient of variation of the observed values and the percentage of observations with <25% difference from true values.

Sampling interval	Regression of observed versus true values		Slope significantly different from 1?	Coefficient of variation	observed values with <25% difference from true values
	slope	r^2			
1 week	1.00 ± 0.01	0.99***	Not significant	0.02	96%
2 weeks	1.00 ± 0.02	0.93***	Not significant	0.07	96%
3 weeks	1.00 ± 0.02	0.86***	Not significant	0.09	92%
4 weeks	1.00 ± 0.02	0.72***	Not significant	0.17	83%

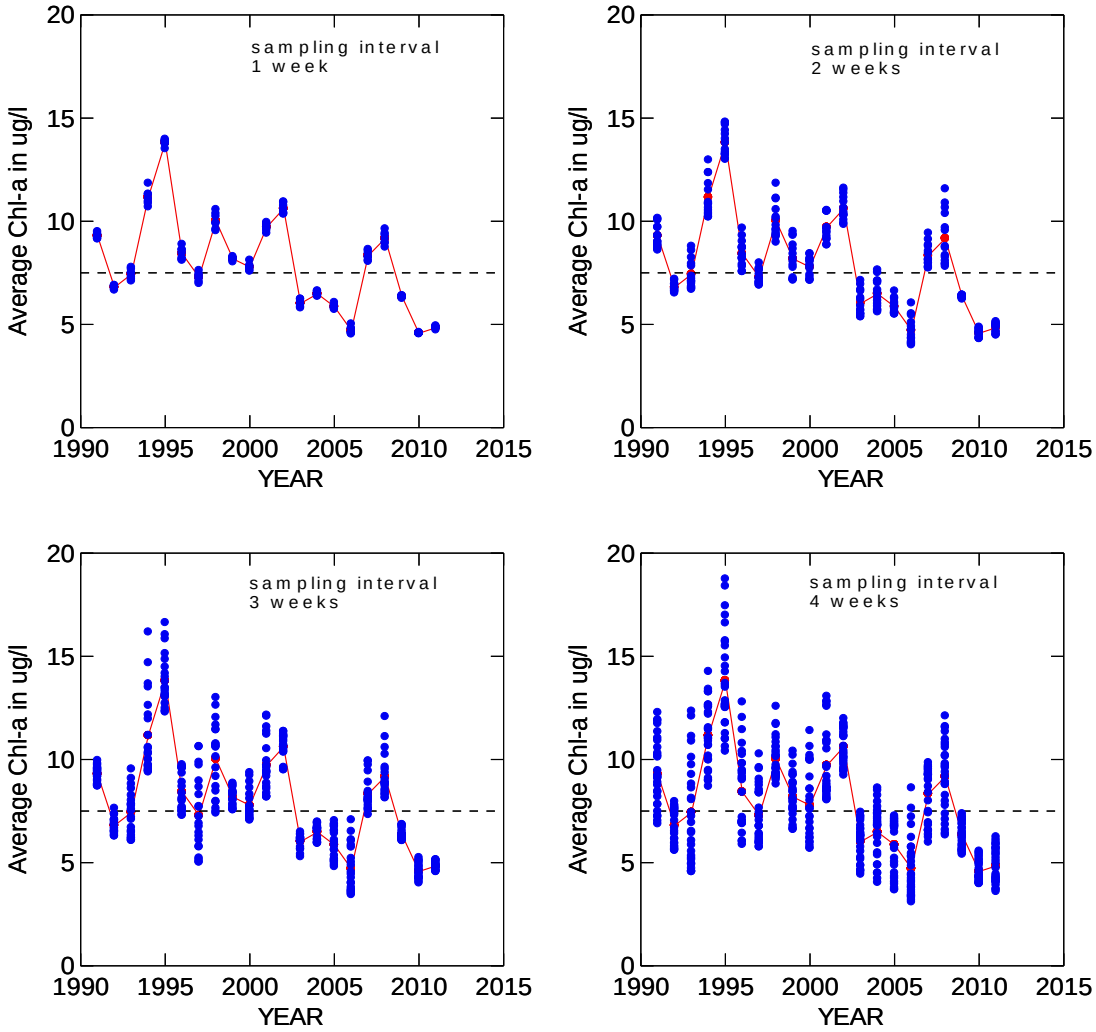

Figure 6.8. Average Chl-a concentration determined from daily values (red; also see Figure 6.2) and determined with 1-, 2-, 3- or 4-week sampling intervals (blue).

Table 6.6 Percentage of cases of misclassification of the average chlorophyll-a concentration (OSPAR) at 1, 2, 3 and 4 week sampling interval.

Sampling interval	% misclassification
1 week	4%
2 weeks	8%
3 weeks	11%
4 weeks	16%

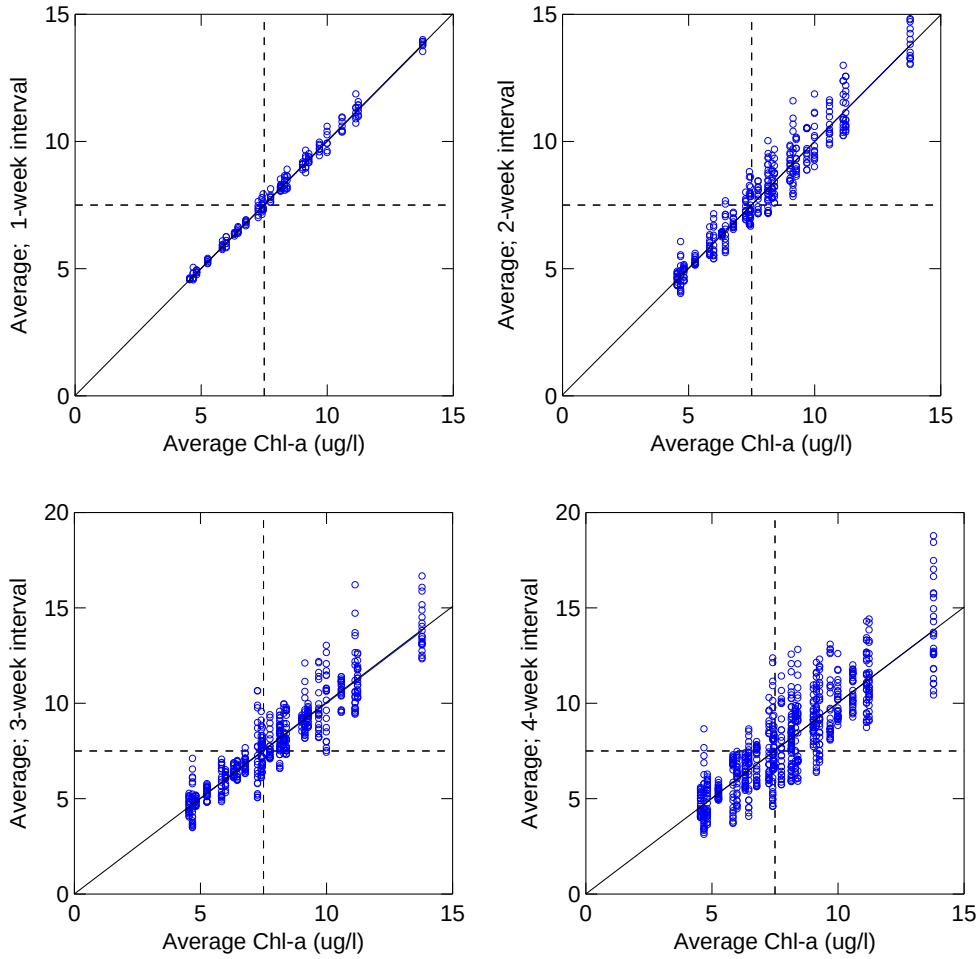

Figure 6.9 The observed average concentration of chlorophyll-a at 4 different sampling frequencies as a function of the true value. The dashed lines indicate the level of 7.5 µg/l. The regression line is equal to y=x (black line).

6.2.3.2 *90-percentile concentrations*

The 90-percentile values, determined with a sampling interval of 1, 2, 3 or 4 weeks, were compared to the true 90-percentiles (Figure 6.10). With increasing sampling intervals, the range in values becomes larger. Figure 6.11 shows the distribution of observed values in relation to the true 90-percentile value. This figure shows that not only the range in values increases with a lower sampling frequency, but also that the observed 90-percentile values tend to deviate stronger from the true values.

Both the accuracy of the observed averages as well as the precision decrease at low sampling frequencies (Table 6.7). The risk of misclassification in this data set increases when the sampling interval becomes longer (Table 6.8).

Table 6.7 Results of the comparison of observed 90-percntile of chlorophyll-a concentrations (OSPAR) at 1-, 2-, 3- or 4-week sampling interval values with true values. Slope and r2 of the linear regression of observed values against true values, mean coefficient of variation of the observed values and the percentage of observations with <25% difference from true values.

Sampling interval	Regression of observed versus true values		Slope significantly different from 1?	Coefficient of variation	observed values with <25% difference from true values
	slope	r^2			
1 week	0.93 ± 0.02	0.93***	p<0.001	0.08	97%
2 weeks	0.86 ± 0.02	0.79***	p<0.001	0.14	90%
3 weeks	0.79 ± 0.03	0.61***	p<0.001	0.18	74%
4 weeks	0.78 ± 0.03	0.56***	p<0.001	0.24	68%

Table 6.8 Percentage of cases of misclassification of the 90-percentile chlorophyll-a concentration (OSPAR/WFD) at 1, 2, 3 and 4 week sampling interval.

Sampling interval	% misclassification
1 week	8%
2 weeks	13%
3 weeks	19%
4 weeks	20%

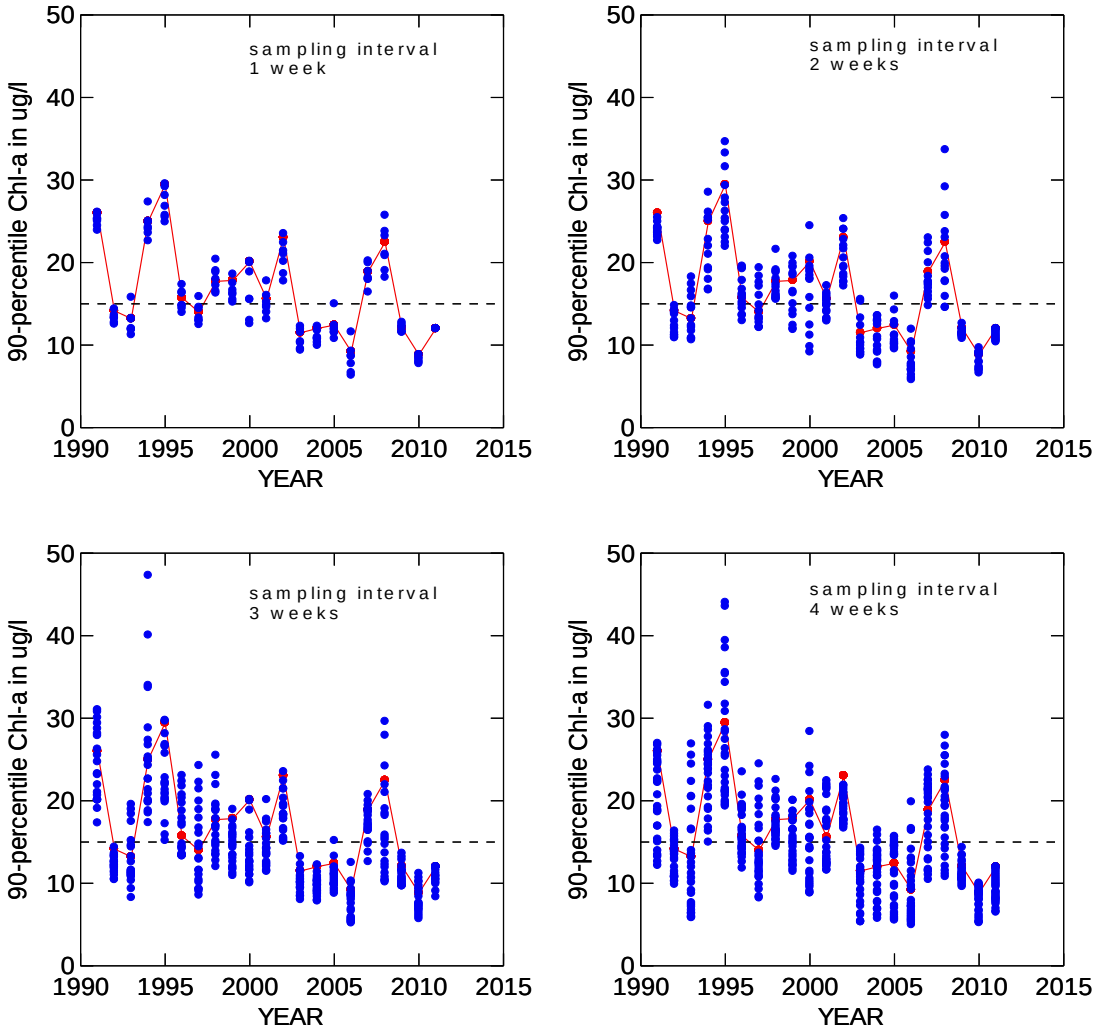

Figure 6.10. 90-percentile of Chl-a concentration determined from daily values (red; also see Figure 6.2) and determined with 1-, 2-, 3- or 4-week sampling intervals (blue).

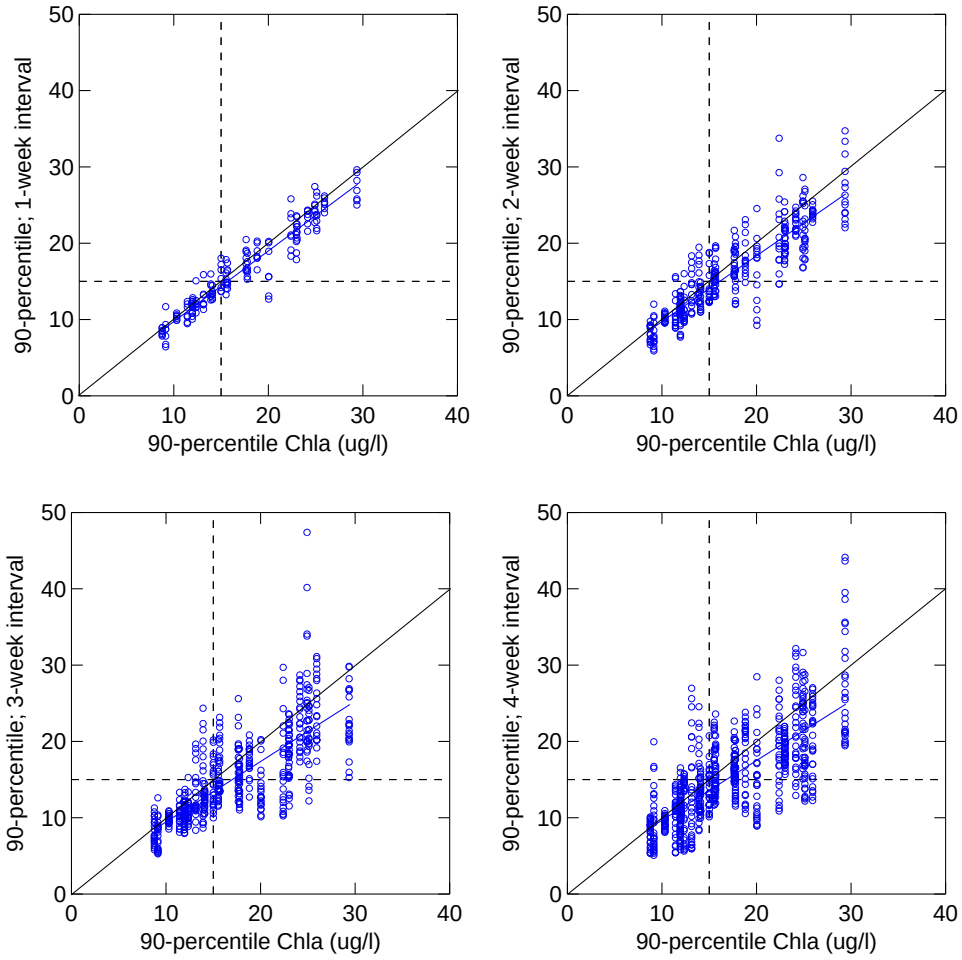

Figure 6.11 The observed 90-percentile concentration of chlorophyll-a at 4 different sampling frequencies as a function of the true value. The dashed lines indicate the level of 15 µg/l. The blue line shows the linear regression line, the black line shows y=x.

6.2.3.3 *Maximum concentrations*

The maximum concentrations of chlorophyll-a, determined with a sampling interval of 1, 2, 3 or 4 weeks, show a similar pattern as observed for the maximum concentrations of *Phaeocystis* (§6.2.1) when compared to the true maximum concentrations. With increasing sampling intervals, the range in values becomes larger and the range extends to increasingly lower values. This is illustrated with Figure 6.12, that shows that the observed maximum values tend to deviate strongly from the true values, also at a relatively high sampling frequency. With low sampling frequencies, the range in values that can be observed increases as well. As shown by the comparison with true values (Table 6.9), both the accuracy and the precision of this metric are low, in particular at sampling intervals of more than 2 week.

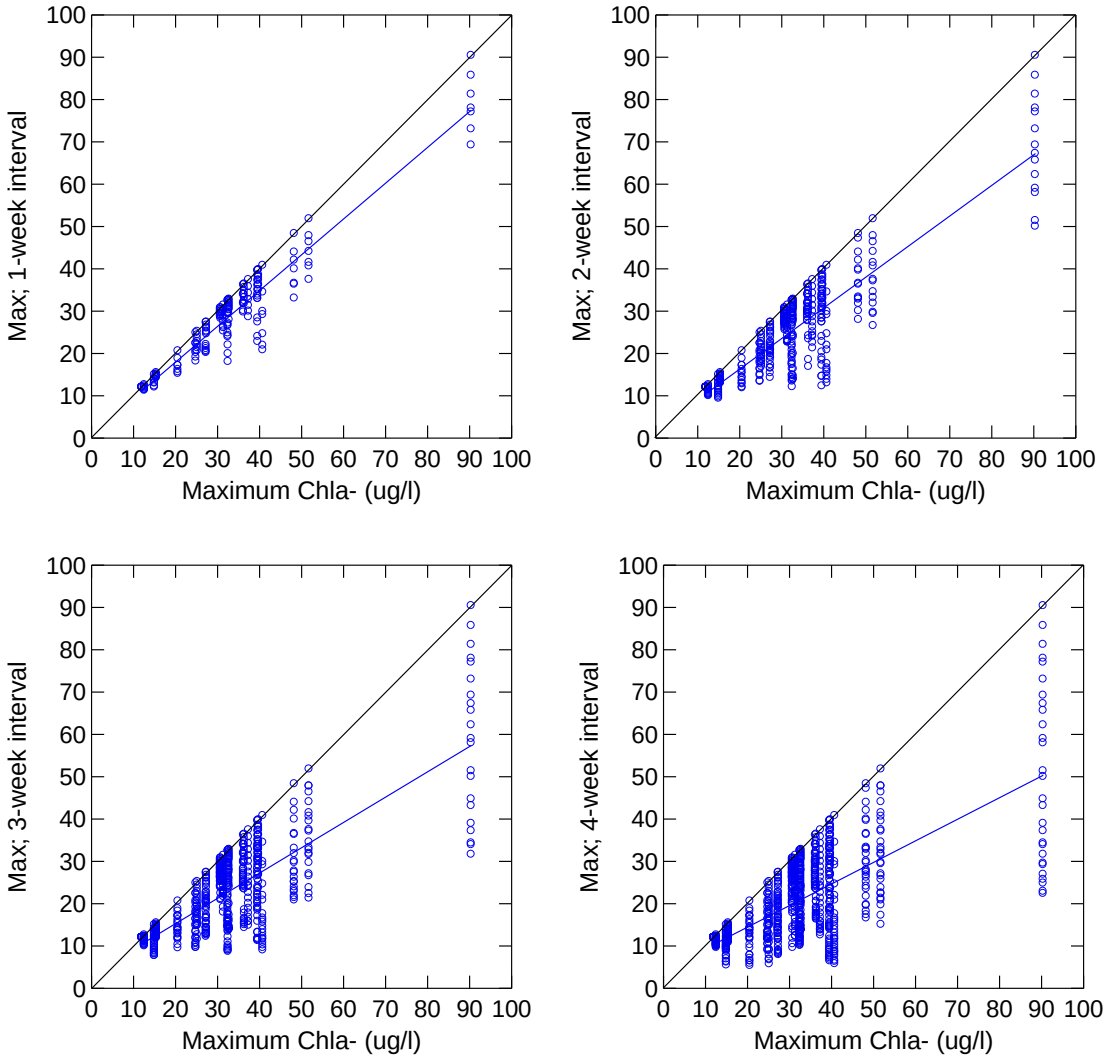

Figure 6.12 The observed maximum concentration of chlorophyll-a at 4 different sampling frequencies as a function of the true value. The blue line shows the linear regression line, the black line shows y=x.

Table 6.9 Results of the comparison of observed maximum of chlorophyll-a concentrations (OSPAR) at 1-, 2-, 3- or 4-week sampling interval values with true values. Slope and r2 of the linear regression of observed values against true values, mean coefficient of variation of the observed values and the percentage of observations with <25% difference from true values.

Sampling interval	Regression of observed versus true values		Slope significantly different from 1?	Coefficient of variation	observed values with <25% difference from true values
	slope	r^2			
1 week	0.85 ± 0.02	0.92***	p<0.001	0.08	89%
2 weeks	0.72 ± 0.02	0.77***	p<0.001	0.16	65%
3 weeks	0.60 ± 0.02	0.61***	p<0.001	0.23	51%
4 weeks	0.51 ± 0.02	0.47***	p<0.001	0.28	41%

6.3 Comparison of the OSPAR and WFD Phaeocystis assessment values

The WFD and OSPAR indicators for *Phaeocystis* use different metrics. In addition to the effects of differences in sampling frequencies and other methodological issues discussed above, the differences between the metrics and the thresholds used for status assessment, may result in differences in assessment results. Using the values for maximum cell number and bloom frequency shown in Figure 6.3, a comparison was made between the metric values and status assessment for WFD and OSPAR (Figure 6.13).

This comparison shows that there is a weak, non-significant correlation between the value of the WFD metric Bloom frequency and the OSPAR metric Maximum cell number (r^2=0.07, p>0.05, n=21).

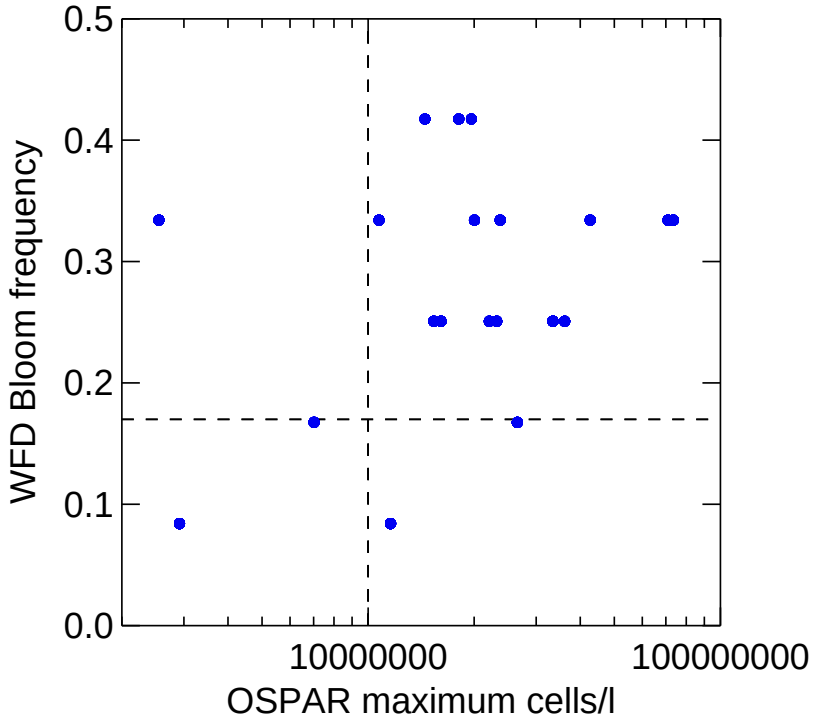

Figure 6.13 Comparison of the WFD Phaeocystis metric (bloom frequency) with the OSPAR metric (maximum concentration). The dashed lines represent the thresholds of 0.17 for bloom frequency (WFD) and the concentration of 10^7 cells/l (OSPAR). Note the log-scale on the x-axis.

6.4 Comparison of the chlorophyll-a average and 90-percentile assessment values

The true values for average and 90-percentile values of chlorophyll-a shown in Figure 6.2 were compared (Figure 6. 14). This comparison shows that there is a significant correlation between the two values. The comparison shows that the rule of thumb that has been applied in the WFD implementation (90-percentile is 2 times the average) is correct.
A threshold value of 7.5 µg/l for the average concentration agrees with a threshold of 15 µg/l for the 90-percentile value.

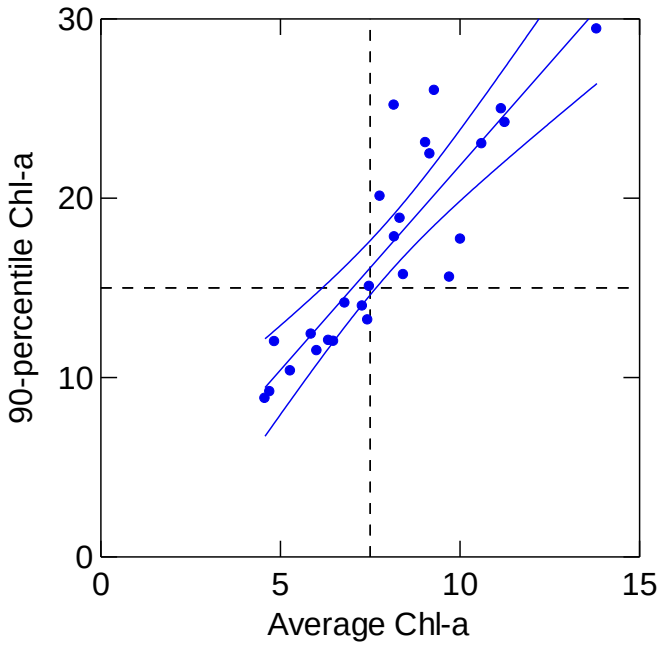

Figure 6.14 Comparison of the 90-percentile values with the average Chl-a concentrations. The dashed lines represent the thresholds of 7.5 (average) and 15 (90-percentile).

6.5 Other issues

6.5.1 WFD *Phaeocystis*: Calculation method for bloom frequency

The WFD metric bloom frequency is based on the assumption that *Phaeocystis* blooms should not occur in more than two months per year. If the metric is based on one sample per month, the calculation is straightforward (although there is the issue of accuracy and precision with such a low monitoring frequency, as discussed in the previous paragraph).
However, when more than one sample per month is taken, there is a risk of overestimating bloom frequency: as even with one sample in a month exceeding the threshold of 10^6 cells/l the whole month is flagged as bloom month, several short peaks in different months, or longer continuous blooms of more than one month, may result in a high bloom frequency. If several samples are taken within a month, the chance of detecting such peaks increases and thus the chance of flagging a month increased. This is illustrated with the data from the time series for the year 2005 (Figure 6.15). In April, May and August the cell concentrations are $>10^6$ cells/l, resulting in a bloom frequency of 0.25 (3 months). However, the total duration of the blooms is 31 days only. The consequence of this phenomenon is that bloom frequency determined from a high number of samples (like the frequency presented in Figure 6.3) bears little relation with the actual number of days where blooms occurred (Figure 6.16).
It is not very likely in practice that bloom frequencies will be based on daily samples. However, the current method leads to a significant bias if more than one sample per month is taken. For those cases, a relatively simple alternative method could be to average the number of blooms per month

before calculating bloom frequency. As an example, a comparison of the results of the two methods for a 2 and 4-week sampling interval is shown in Figure 6.17.
With a 4-week sampling interval, both methods obviously give similar results. However, with a 2-week sampling interval averaging of the blooms per month give a more accurate and precise value for bloom frequency. The correlation with the number of days with a bloom is stronger (r^2=0.81) using the average of blooms, than when using the current method (r^2=0.49).

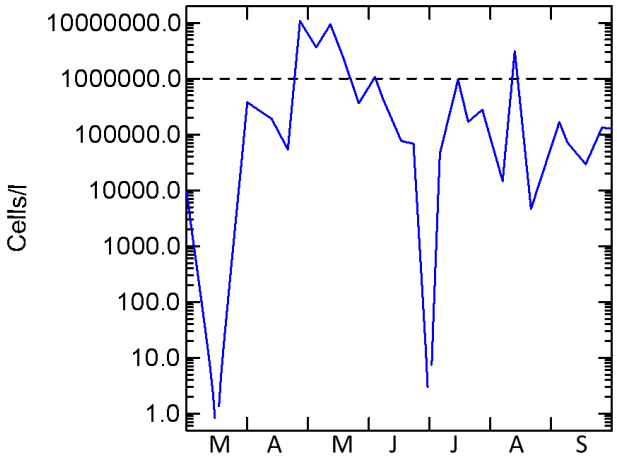

Figure 6.15 Time series of Phaeocystis concentrations in the year 2005 (interpolated daily concentrations), during the months March-September. The dashed line shows the threshold of 10^6 cells/l.

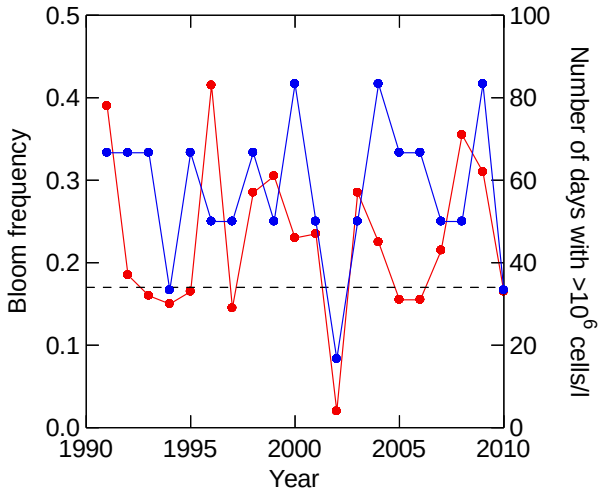

Figure 6.16 Bloom frequency of Phaeocystis (blue circles and line; data from Fig 6.3) and the number of days with Phaeocystis concentrations >10^6 cells/l (red circles and line) in the years 1990-2010, calculated from the interpolated daily concentrations.

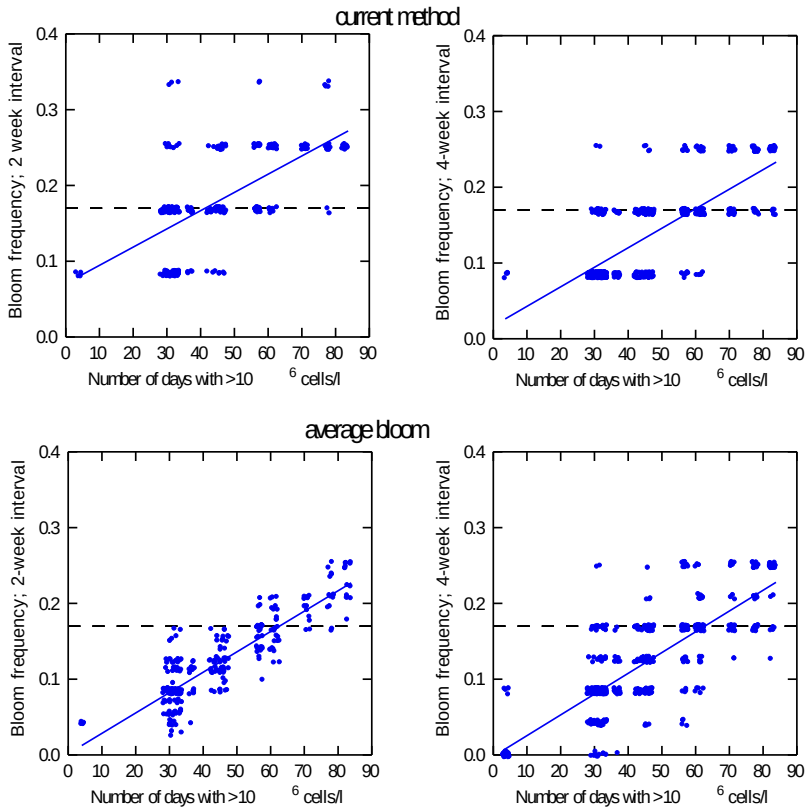

Figure 6.17 Bloom frequency based on the maximum concentration per month (current method: upper panel) and based on the average of blooms per month (lower panel) in relation to the number of days with cell concentrations >10^6 cells/l.

6.5.2 WFD *Phaeocystis*: definition of observation period

Phaeocystis blooms generally occur in spring and sometimes in summer. For the WFD metric preferably 12 monthly samples have to be used. In the Netherlands, however, the period March-September, inclusive (7 months) is used to calculate bloom frequency. As most blooms of *Phaeocystis* occur in this period, the error caused by not sampling the year round is small. The following comparison shows the difference between the assessment results of samples taken during the growing season only and from year-round sampling.

To have a considerable number of observations not only the Dutch WFD monitoring stations in the coastal water bodies have been used, but also all other stations in the North Sea. This data series cover 20 stations sampled in the years 1990-2010. As not all stations were sampled in all years, a total of 382 year series was available (Table 6.10).

In two cases blooms >10^6 cells/l were observed in February in the North Sea (Noordwijk 2 in 2005, Noordwijk 70 in 1992). In 11 cases blooms >10^6 cells/l were observed in October or November in the North Sea (Boomkensdiep in 2008; Goeree 6 in 2003; Huibertsgat oost in 2009; Noordwijk 10 in 2003; Noordwijk 20 in 1992 and 2003; Noordwijk 70 in 1994; Terschelling 4 and Terschelling 10 in 1997 and 2003). In total 13 cases with blooms > 10^6 cells/l outside the period from March to September inclusive, but no change in number of “years” for individual stations with concentrations > 10^6 cells/l (Table 6.10).

In all of these cases bloom frequencies for those years and stations already exceeded the threshold. In 3 out of these 13 cases (this is in total 3 out of 382 cases) the assessment would have been different by including observations from months before or after March-September: Huibertsgat oost in 2009, Noordwijk 2 in 2005 and Terschelling 10 in 1997. Two of these stations are WFD stations (Table 6.10).

Table 6.10 Total number of years (sum of years for all stations), number of observations with Phaeocystis concentrations exceeding 10^6 cells/l, and number of observations of bloom frequencies above the WFD threshold, for March-September and for the entire year.

	nr *years* (all stations)	nr *years* with concentration>10^6 cells/l		nr *years* with bloom frequency >0.17	
		March-September	Entire year	March-September	Entire year
Coastal waters	256	213	213	73	76
Southern Bight	42	39	39	12	12
Oyster grounds	63	4	4	0	0
Dogger bank	21	13	13	0	0
Total	382	269	269	85	88

6.5.3 WFD *Phaeocystis*: threshold for bloom definition

The Netherlands use 10^6 cells/l as the threshold for the definition of a normal bloom in the WFD metric and count the annual frequency of these blooms. Belgium, however, proposes to use a threshold of $4*10^6$ cells/l (Denayer et al., 2010; Lancelot et al., 2009) and applies the threshold to the annual maximum abundance of *Phaeocystis*, and not to bloom frequency. A comparison of the two methods, carried out by Prins & Baretta-Bekker (2010) shows that a bloom frequency of >2 months per year (moderate status in the Dutch WFD metric) coincides with a maximum abundance above the "Lancelot" threshold of 4 million cells/l in all cases. At a bloom frequency of 1 or 2 months (good status in the Dutch WFD metric), the 4 million cells/l threshold is exceeded in less than 50% of the cases (Figure 6.18). In conclusion, using the maximum abundance threshold of 4 million cells/l for the boundary between good/moderate status would result in a more stringent classification than the Dutch WFD metric.

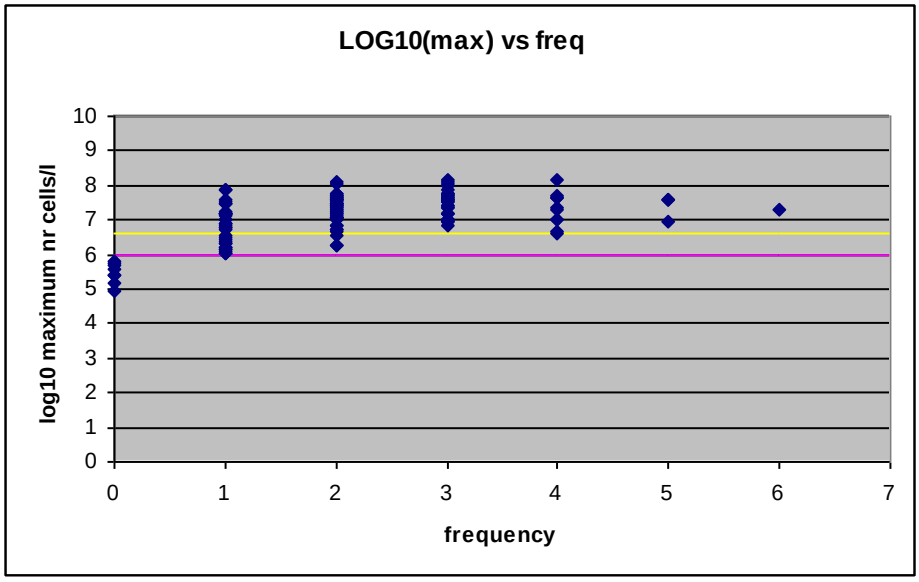

Figure 6.18 The log of the annual maximum number of cells against the frequency of months with blooms of more than 10^6 cells/l. The pink line indicates a concentration of 10^6 cells/l, the yellow line a concentration of 4.10^6 cells/l (Source Prins & Baretta-Bekker, 2010).

6.6 Conclusions of sensitivity analysis

The metrics show different responses to changes in sampling frequency.

Maximum concentrations cannot be used as a reliable metric, neither for *Phaeocystis* (currently used in OSPAR COMP) nor for chlorophyll-a. The metric is inaccurate and not precise. The chance of detecting the maximum concentration is low, even with relatively short sampling intervals (e.g.

weekly sampling). Due to the fact that both *Phaeocystis* cell concentrations and chlorophyll-a concentrations show short-lived high peaks during the growing season, the probability of not detecting the maximum but encountering much lower values is relatively high, and this probability increases when sampling intervals are longer. Consequently, the risk of seriously underestimating the maximum concentration as well as the risk of misclassification are high and strongly dependent of sampling interval.

The current metric for *Phaeocystis*, used for the WFD, has its limitations as well. With the current method, the calculation of bloom frequency is not independent of sampling interval. This is in particular due to the fact that only one observation per month above the threshold of 10^6 cells/l is sufficient to flag a month as bloom month. If the sampling frequency is higher than 1 sample per month, calculation of bloom frequency from the average number of blooms per month is more accurate.

In a small number of cases (3 out of 382) the WFD indicator for *Phaeocystis* gives a different status assessment when a longer observation period than the currently used period of March-September is used.

For chlorophyll-a, the 90-percentile of the concentrations during the growing season is used as a metric. The accuracy and precision of this metric decrease when the sampling interval increases. With sampling intervals of more than 2 weeks, the 90-percentiles are not a very reliable metric.

Average concentrations of chlorophyll-a are currently used in the OSPAR COMP, in addition to 90-percentiles. With long sampling intervals, averages are more accurate and precise than 90-percentiles. With sampling intervals of 2 weeks or less, the difference between these two metrics is small.

In general, it should be realized that at long sampling intervals, all metrics have a relatively low precision. The risk of failure to detect changes in status and the risk of misclassification is high in those cases.

7 Discussion

The OSPAR Comprehensive procedure (COMP) has been developed to specifically assess the eutrophication status of marine and estuarine waters in an integrated way, taking into account causative factors (nutrient enrichment) as well as direct and indirect effects. The WFD has been developed to assess the ecological and chemical status of the water system, including hydromorphological characteristics, contaminants and a number of biological quality elements and supporting physico-chemical elements. Eutrophication is addressed in the WFD through the supporting physico-chemical elements (nutrients), through the biological quality element phytoplankton, and where applicable through the other biological quality elements. Assessments of all elements are integrated only in the final assessment of ecological status.

The OSPAR COMP assesses the causes, direct effects and indirect effects of nutrient enrichment in relatively large areas, in the Dutch application averaging over a number of monitoring stations. WFD measures only the direct and the indirect effects of nutrient enrichment in relatively small water bodies, in the Dutch implementation represented by one monitoring station only.

The main part of this report focuses on the Dutch applications, only one element of the WFD assessment (the biological quality element phytoplankton) has been compared with the analogous part of the OSPAR COMP (phytoplankton as main indicator of eutrophication). Both assessment methods have been developed for different areas in the North Sea: OSPAR for both the coastal waters of the North Sea area with a salinity < 34.5 and for the marine waters of the North Sea, whereas the WFD is applied only to the coastal water bodies (1st nautical-mile zone of the Dutch North Sea).

As both assessment methods apply to different geographical regions, with different levels of pressures (for eutrophication), the assessment levels are different. The OSPAR assessment levels, valid for a large area of the North Sea, cannot be identical to the WFD assessment levels which are valid only for the small water bodies of the I nm coastal zone. This is illustrated in Figure 2.4, where the influence of the freshwater discharges can be seen in the spatial pattern of salinity. The salinity in the Dutch 1-nm coastal zone is between 28 and 32 (see Figure 2.4), much lower than the average salinity in the OSPAR coastal waters (see Figure 2.3).

Based on salinity the Dutch 1-nm coastal zone has been subdivided in polyhaline water bodies with salinities between 28 and 30 (the Holland coast and the Northern Delta coast) and euhaline water bodies with salinities between 30 and 32 (other water bodies). As a consequence of the difference in the salinity range the assessment level for chlorophyll-a in the WFD water bodies Holland coast and Northern Delta coast has been set higher than the levels of the euhaline coastal water bodies (Table 4.2). Whether this difference in assessment levels is proportional to the difference in nutrient loading of both water types is uncertain.. The differences in assessment levels explain why the results of the OSPAR AND WFD assessments (Tables 4.3 – 4.5) are not necessarily identical when applied to individual stations.

Other differences between the two assessments are due to the way of assessing *Phaeocystis*. In chapter 6 both methods are carried out with an artificial time series, based on measurements at Noordwijk 10. OSPAR takes the maximum number of *Phaeocystis* cells/l of all observations in the assessment area during the period of the assessment. WFD, on the other hand, calculates the number of months with a maximum concentration higher than 10^6cells/l, which is defined as a normal bloom. As *Phaeocystis* blooms are a normal phenomenon in spring, and as eutrophication causes higher and longer-lasting blooms, this method not only takes into account the number of peaks, but also the duration of the peaks, because when a peak lasts longer than one month, it is counted more than once.

8 Conclusions

General

- The metrics show different responses to changes in sampling frequency.
- Maximum concentrations cannot be used as a reliable metric, neither for *Phaeocystis* nor for chlorophyll-a.
- With long sampling intervals (>= 4 weeks) all metrics have a relatively low precision.

Chlorophyll-a – OSPAR

- With long sampling intervals, averages are more accurate than 90-percentiles. With sampling intervals of 2 weeks or less, the difference between these two metrics is small.

Chlorophyll-a – WFD

- As described under OSPAR – chlorophyll-a: average concentrations are more accurate than the 90-percentiles.

***Phaeocystis* – OSPAR**

- The value for the assessment level for *Phaeocystis* of Lancelot et al., 2009 (4×10^6 cells/l) is more appropriate than the value (10^6 cells/l) used in the OSPAR COMP.
- In general maximum concentrations cannot be used as a reliable metric with sampling intervals of 2 weeks or more.

***Phaeocystis* a – WFD**

- **Observation period** Monitoring during the growing season in contrast to monitoring during the whole year gives different assessment results only in a small number of cases (3 out of 382).
- **Threshold** Raising the bloom threshold concentration from the value currently used in the Netherlands (10^6 cells/l), to the concentration proposed by Belgium (4×10^6 cells/l), is not necessary, because in the Dutch method the frequency of normal blooms is calculated, while the Belgian threshold concentration is the concentration where colony formation starts.
- **Calculation of bloom frequency**
  With the current method, the calculation of bloom frequency is not independent of sampling interval. This is in particular due to the fact that only one observation per month above the threshold of 10^6 cells/l is sufficient to flag a month as bloom month. To be sure of flagging *Phaeocystis* blooms correctly, we either have to be sure that a *Phaeocystis* bloom duration is always longer than one month (then a monthly sampling frequency always will signal a bloom).
  Sampling frequency
  If the sampling frequency is higher than one sample per month, calculation of bloom frequency from the average number of blooms per month is more accurate.

9 References

Anonymous, 2008. OSPAR report on the second application of the Comprehensive Procedure for the Belgian marine waters. In: OSPAR, 2010. Quality Status Report 2010. OSPAR Commission. London. http://www.ospar.org/documents/07-08/icgcomp2/docse/00105rev1_be_revised%20comp2%20 report .pdf

Anonymous, 2011. Memo DE/NL Intercalibration of phytoplankton NEA 3/4 - class boundaries for chlorophyll-a. 7 pp.

Baptist, H.J.M. & E. Jagtman, 1997. De AMOEBES van de zoute wateren. WSV werkgroep van de zoute wateren. Rapport RIKZ-97.027: 149 pp.

Baretta-Bekker, J.G. P. Bot, T.C. Prins, W. Zevenboom, 2008. Report on the second application of the OSPAR Comprehensive Procedure to the Dutch marine waters. 79 pp. http://www.ospar.org/documents/07-08/icgcomp2/docs-/00104rev5_nl%20comp2%20report.pdf

Brockmann U., D. Topcu & M. Schütt,2008. Assessment of the eutrophication status of the German Bight according to the OSPAR Comprehensive Procedure. Assessed period: 2001 – 2005. In: OSPAR, 2010. Quality Status Report 2010. OSPAR Commission. London. 145 pp. http://www.ospar.org/documents/07-08/icgcomp2/docs-e/00108rev1_de%20revised%20comp2%20report.pdf

Cadée, G.C., J. Hegeman, 2002. Phytoplankton in the Marsdiep at the end of the 20th century; 30 years monitoring biomass, primary production, and *Phaeocystis* blooms. J. Sea Res. 48: 97–110.

Carletti Alessandro & Anna-Stiina Heiskanen, 2009. Water Framework Directive intercalibration technical report Part 3: Coastal and Transitional waters. JRC Scientific and Technical Reports EUR 23838 EN/3 - 2009. 224 pp.

Davies, Susan P. (ed.), 2012. Peer review of the intercalibration exercise Phase ii European water framework directive. Final draft report. 5 oct. 2012. 261pp.

Denayer, Sara, Jeroen Van Wichelen, Koen Sabbe and Wim Vyverman, 2010. Phytoplankton biomonitoring of the BCZ in the context of the EU Water Framework Directive. Report of November 2010. Project MM/WB/KRW-009 –Phytoplankton biomonitoring of the BCZ. 37 pp.

European Commission (EC), 2000. Directive 2000/60/EC of the European Parliament and of the council of 23 October 2000 establishing a framework for Community action in the field of water policy (Water Framework Directive). Official J. L 327/1..72 pp.

European Commission (EC-WFD), 2005. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No 13. Overall Approach to the Classification of Ecological Status and Ecological Potential.

European Commission (EC-MSFD), 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). 64 pp.

Lancelot, C., V. Rousseau, N. Gypens, 2009. Ecologically based indicators for *Phaeocystis* disturbance in eutrophied Belgian coastal waters (Southern North Sea) based on field observations and ecological modelling. J. Sea Res. 61: 44-49.

OSPAR, 2005. Common Procedure for the Identification of the Eutrophication Status of the OSPAR Maritime Area. OSPAR Reference number: 2005-3): 34 pp.

OSPAR, 2009 Eutrophication Status of the OSPAR Maritime Area. Second integrated Report. OSPAR report. 108 pp.

OSPAR, 2010. Quality Status Report 2010. OSPAR Commission. London. 176 pp.

OSPAR, 2011. Preliminary outcome of Task 2: draft revised assessment levels. Annex 7 of the Summary Record of the Meeting of the Intersessional Correspondence Group on Eutrophication in Dessau (Germany) on 15-18 November 2011. 4pp.

OSPAR, 2012a. SPAR MSFD Advice Document on Eutrophication. WG GES meeting, 17 February 2012, Brussels. 22 pp.

OSPAR, 2012b. Review and revision of the Common Procedure for the Identification of the Eutrophication Status of the Maritime Area. Meeting of the OSPAR Commission, Agenda Item 7. Bonn: 25-29 June 2012. 46 p.

Prins, T.C., 2010. Memo: Comparison of methods for the derivation of reference values for chlorophyll-a. 10 pp.

Prins, T.C. & J.G. Baretta-Bekker, 2010. WFD phytoplankton metric *Phaeocystis*; Application in the Netherlands. Deltares report 1203385-000. 47 pp.

Van den Berg, M. (ed.). 2004. Achtergrond referenties en maatlatten. Rapportage van de expertgroep fytoplankton. 29 pp.

Annex A - Measuring methods

Introduction

In the technical annex (Carletti & Heiskanen, 2009) no attention has been paid to the methods used to determine the chlorophyll-a concentration and the *Phaeocystis* cell counts.
To measure chlorophyll there are various techniques, including spectrophotometry, high performance liquid chromatography (HPLC), and fluorometry. Because monitoring programs demand that the numbers generated are accurate as well as precise, the choice of a technique is important.

Methods to measure chlorophyll-a

The Netherlands Surface water samples are filtered using Whatmann GF/C and either stored at 4 ± 2ºC and subsequently analysed within 24 hours, or samples are stored for a longer period between -35 and -20ºC. For chlorophyl-a analysis samples are filtered and filters are stored at < -65ºC (non-filtered samples should be stored in dark bottles at 4ºC and analysed within 30 hours). For chlorophyll-a the frozen filters are extracted prior to analysis with acetone, followed by separation with HPLC and fluorescence detection.

Belgium (Anonymous, 2008): Chlorophyll-a concentration is measured by MUMM's Chemistry Laboratory as detailed in Steendijk and Schreurs (2002), based on the High Performance Liquid Chromatography method (HPLC). Water samples are filtered on-board with GF/C filters, and are then frozen in liquid nitrogen and stored longterm at –80°C. Pigments are extracted in 90% acetone with the use of a cell-homogenizer, followed by centrifugation. The chlorophyll pigments are separated with reversed phase HPLC. Quality control procedures include first line control (instrument calibration and performance), second line control (use of control charts for batches) and third line control (twice-yearly intercomparison within QUASIMEME). From 2005 the laboratory was accredited for analysis of this parameter as conforming to norms NBN and ISO/IEC 17025:2005.

Germany In Germany there are at least two laboratories that measure chlorophyll-a. Both labs extract with acetone, but one of the labs uses the fluorometry methods, while the other uses spectrophotometry.

UK Up to 200 cm^3 of seawater is passed through Whatman GF/F glass fibre filters. These filters are extracted overnight in 90% acetone neutralised with $NaHCO_3$ in darkness in a refrigerator as described by Tett (1987). After centrifugation, extracted fluorescence is measured using a Tuner Designs Model 10 flourometer before and after acidification with 8% HCl, and used to calculate concentrations of chlorophyll. The fluorometer is been calibrated using pure solutions of chlorophyll-a (Sigma Chemical Co, ORIGIN) with concentration determined spectrophotometrically. This method through the acidification step will correct for degradation products i.e. phaeopigments, however, it will not correct for presence of chlorophyllide-a, a photosynthetically inactive precursor to chlorophyll-a. Only through chromatography will it be possible to separate chlorophyllide-a. Therefore all results from this method are reported as measurements of *extracted chlorophyll* (extracted-chl) and not chlorophyll-a (chl-a). The result is a likely over-estimation of chl-a, which may be considered a more precautionary approach.

France Using membrane filtration Whatman/C 47 mm filter GF), the particulate material is concentrated. The pigments are then extracted in a solvent (90% acetone). After centrifugation for

20 minutes at 6000 tr.mn^{-1}, the absorbance of the supernatant was measured by spectrophotometry at 665 and 750 nm, which provides the rate of chlorophyll a (the absorbance peak is at 665 nm). The concentration is expressed in mg/l or mg/m3 (accuracy ± 0.01). The detection limit is 0.04 mg/l.

Methods to count *Phaeocystis* cells or cells of other phytoplankton species

The Netherlands

Sampling Water samples are taken at 1 m below the water surface. If the water column is stratified two additional samples are taken with a rosette sampler and Niskin bottles, at the pynocline and at 1 - 3.5 m above the bottom.

The phytoplankton monitoring study comprises live samples on some of the stations (WFD station Noordwijk 2 and three other stations), for a qualitative inventory of phytoplankton composition. This is done especially for those species that are difficult to recognize after fixation with Lugol, and to follow developments in the phytoplankton in near-real time. In addition, formalin-fixed samples (for determining the density of coccolithophores) and Lugol-fixed samples for a quantitative description of the species composition and density of phytoplankton throughout the year, are taken. Only the last group of samples is relevant for the number of *Phaeocystis* cells per liter.

Fixation All Lugol-fixed samples have an approximate volume of 1 liter of seawater fixed with 4 ml of acetic acid Lugol in brown glass bottles. They are checked on receipt for correct fixation, labelling and registration and stored in the dark at 4 °C until analysis.

Just before processing the samples, the fixation is checked again. The contents of the bottles have to have a cognac-like colour. Samples, which have not been fixed correctly, are discarded. After analysis, the remainder of the sample is kept indefinitely in the dark and at 4°C.

Concentrating the samples The analysis focuses on determination of the number of phytoplankton cells per liter in the original, fixed, non-concentrated samples. In the case of turbid samples and samples which are suspected to contain species at densities of >10000 cells/l, two sub samples are prepared: one non-concentrated straight from the bottle and an approximately 10 times concentrated sub sample.

The procedure of concentrating the sample is as follows: Samples are left undisturbed for at least one week in a low-vibration, cold and dark environment. During this period all solids settle out. Without disturbing the settled material the samples are brought into the treatment room, and left at least a day for acclimatisation.

The largest part of supernatant is then removed from the sample bottle and collected in a calibrated measuring cylinder of 1000 ml. The settled residue (volume 40-200 ml) is transferred to a calibrated measuring cylinder of 250 ml after homogenization.

Lugol-fixed subsamples A sub sample of of 0.2, 1.3 or 2.3 ml of the unconcentrated and / or 1 to 2.3 ml of the concentrated sample is examined. This subsample is taken from the homogenized sample with a calibrated pipette (Eppendorf pipette or automatic) and pipetted into one or more sedimentation chambers with a surface area of 1.25 cm^2 and a height of 1 to 2 cm. In the sedimentation chambers 0.2 to 1 ml of diluted seawater algae-free Lugol is pipetted in advance. The sedimentation chambers were covered with a cover slip and left alone in a dark area for at least four hours to allow the cells to settle, assuming a sedimentation rate of 0.25 cm / hour for nano-plankton.

If there are less than 15 observations of an “important” taxon in a sample, additional observations are done in the next sub sample or, if necessary in a 10x concentrated subsample.

Plankton-poor samples must be concentrated to obtain a sub sample with a sufficient number of organisms. This is further complicated by the presence of inorganic sediment and detritus particles, especially in the Dutch coastal waters (due to turbid river outflows).

Analysis The samples are analyzed using an inverted microscope (Olympus IMT-2), with a long working distance condenser (numerical aperture 0.55), 10 × WHK eyepieces, one with a calibrated

ocular micrometer and the following lenses: Olympus SPlan-Apo 20x / 0.70 and 60x/1.40. Analyses are performed in bright field. To correct for any edge effects image fields are studied in sectors of the sedimentation cuvet. The analysis is supported by epi-fluorescence microscopy.
Determination *Phaeocystis* sp. appears as three distinct types of cells, single cells, flagellate cells and cells from colonies. In samples fixed in Lugol *Phaeocystis* colonies disintegrate easily into single cells. Therefore single cells may be natural, but also may originate from disintegrated colonies. The number of cells in small colonies are counted and in large colonies (> 100 cells) they are usually estimated. For analysis, the three types of cells are summed.
Although *Phaeocystis globosa* is the most common species of the genus *Phaeocystis* taxonomically it is difficult to distinguish this species from *P. pouchetii*. The single cells and flagellate cells of both species cannot be identified down to species level with a light microscope. In live samples colonies of both species can be distinguished, based on differences in morphology. Although the observed colonies of *Phaeocystis* in the live samples from the coastal zone station Noordwijk 2 all belonged to the species *Phaeocystis globosa*, it is not sure that all single and flagellate *Phaeocystis* cells belong to this species and therefore they all are identified as *Phaeocystis* sp.

Belgium Source: Anonymous (2008)
Coastal waters Phytoplankton indicator species have not been measured in Belgian waters as part of the standard monitoring programme before 2006. Cell counts for *Phaeocystis* spp. were not available for the assessment period 2001-2005 at the monitoring stations in order to evaluate if *Phaeocystis* spp. reaches the elevated levels (> 10^7 cells/l during more than 30 days). According to model simulations, this threshold is exceeded for parts of the Belgian marine waters (Lacroix *et al*., 2007). This is also confirmed by one-off scientific surveys in 2003 and 2004, which reveals that *Phaeocystis* spp. do occur in elevated levels in the Belgian part of the North Sea (AMORE project data; Lancelot *et al*.2007). This sparse information has been used for the assessment of the period 2001-2005. From 2006 *Phaeocystis* cell concentrations are counted.
Offshore waters No data are available to identify the extent of the *Phaeocystis* blooms until the offshore waters. Therefore this parameter is considered to remain inconclusive for the offshore waters.

Germany
No information available.

UK
Instead of a set of individual 'phytoplankton indicator species' UK deployed an index, which includes measures of *Phaeocystis* spp. and any phytoplankton taxa with abundance over a defined threshold. This is to provide a better assessment of disturbance of the phytoplankton as a result of nutrient enrichment.

France The procedure used is the one recommended by Belin (2001) (Method Utermöhl 322/89-CSRU/BCM) in Ifremer's phytoplankton monitoring networks. Observations are made as soon as possible, at least for the target species, and on the two types of samples: live and fixed. Fixing methods will differ depending on whether the observation is performed immediately after collection or offset in time.
- Fixing for immediate observation (in the day): application of lugol (2.5 ml/l). The stock solution has to be kept in the dark, and a conservation in brown bottle is limited to a few months.
- Fixing observation offset in time (overnight) : after application of Lugol, formaldehyde addition for conservation of fragile forms.
- Fixing for sample storage: after observation, the prolonged storage of the sample is possible by adding formaldehyde.
The observation is performed in a 10 ml settling tank at a suitable magnification. The minimum settling time is 6 hours (after about 2 hours, 80% of species are already on the bottom of the tank).

For the observation, when no species seems to be truly dominant, the tank is completely covered. In cases where one or more species appear to be very dominant, sub-sampling is performed. It is a numeration on the diameters of the tank or randomly selected areas in the tank. A multiplier to estimate the number of observed phytoplanktonic cells depending on the magnification, on the number of diameters, or the number of areas observed. The result is extrapolated to a volume of 1 liter (result number of cells/l).

Summary

Table A.1 below gives an overview of the major aspects of the Chlorophyll-a and *Phaeocystis* monitoring and measuring/counting procedure. Information on the methods used by Belgium was provided by S. Denayer (Gent University), and on the methods used by the UK by S. Milligan (CEFAS).

Table A.1 Overview of the Chlorophyll-a and Phaeocystis sampling and measuring/counting procedure in the national monitoring programmes, used for both OSPAR and WFD assessments.

Sampling	**BE**	**NL**	**DE**	**UK**	**FR**
No of sampling sites	3	1 for each WFD water body; five WFD sites in the North Sea	?	?	11 (regional monitoring network SRN for the eastern English Channel, from Dunkerque to the Bay of Somme) + national monitoring REPHY.
Sampling frequency	Monthly. But in in March-May 3 times a month	For WFD: once per month; For OSPAR 2 to 4 weekly in the growing season (March-Sept.)	?	?	Monthly. But twice a month between March and June for the SRN monitoring network.
Sampling depth	3 m below surface	1 m below surface	?	?	1 m below surface
Chlorophyll	**BE**	**NL**	**DE**	**UK**	**FR**
filter	GF/C	Whatmann GF/C		Whatmann GF/C glass fibre	Whatmann GF/C 47 mm
Storage of living samples		4 ± 2 ºC			4°C (refrigerator)
Storage for longer period	at -80 ºC	between –20 and -35 ºC			for few weeks: between -18 and -20°C, at best -25°C. Beyond a few weeks : liquid nitrogen (-196°C).
extraction	90% aceton	aceton	aceton	90% aceton, neutralised with NaHCO3	90% aceton
detection	Reversed phase HPLC	HPLC and fluorescence	1: fluorometry 2. Spectrophotometry	fluorometry (Tuner Designs Model 10) after acidification with 8% HCl	Spectrophotometry 665 and 750 mm

Table A.2 Overview of the Phaeocystis sampling and counting procedure in the national monitoring programmes an used for both OSPAR and WFD assessments. Part of the information has been copied from: Prins & Baretta-Bekker (2010). NB Phaeocystis colonies disintegrate due to fixation

Phytoplankton species					
Sampling	**BE**	**NL**	**DE**	**UK**	**FR**
No of sampling sites	3	WFD: 1 for each water body; five WFD sites in the North Sea OSPAR: more stations in each of the areas	?	?	11 (regional monitoring network SRN for the eastern English Channel, from Dunkerque to the Bay of Somme) + national monitoring REPHY.
Sampling frequency	Monthly. But in in March-May 3 times a month	WFD: one sample per month; OSPAR 2 to 4 weekly in the growing season (March-September)	?	?	Monthly. But twice a month between March and June for the SRN monitoring network.
Sampling depth	3 m below surface	1 m below surface	?	?	1 m below surface
Procedure	**BE**	**NL**	**DE**	**UK**	**FR**
Live samples collected	Yes	At one WFD station	xxx	?	Yes
Live samples counted	No	No	xxx	?	No
Sample fixation	Lugol	Lugol	xxx	Lugol	Lugol
Counting	**BE**	**NL**	**DE**	**UK**	**FR**
Counting procedure	Utermöhl chambers, inverted microscope	Utermöhl chambers, inverted microscope	xxx	Utermöhl chambers, inverted microscope	Utermöhl chambers, inverted microscopel
Separate counting of flagellate, non-flagellate cells and cells in colonies	Yes, colonies separated through inverse filtration	Yes, in Lugol-fixed samples (NB colonies partly disintegrate)	xxx	No, *Phaeocystis* not distinguished from other small flagellates	No
Counting of colony size	Yes	No	xxx	No	No
Calculation of total density	Sum of colonial cells, free-living flagellated cells and non-flagellated cells	Sum of colonial cells, free-living flagellated cells and non-flagellated cells	xxx	Few obs. of *Phaeocystis* Counted as “small flagellates possibly *Phaeocystis*”	No

Indicator	BE	NL	DE	UK	FR
WFD Threshold for bloom definition	Until now: 10^6 cells/l, but for future proposed to use 4 x 10^6 cells/l	10^7 cells/l	xxx		No theshold for *Phaeocystis* bloom.
Calculation of bloom frequency (WFD)	Step 1: calculation of samples exceeding threshold Step 2: Frequency based on number of samples exceeding threshold	Step 1: calculation of monthly maximum Step 2: Frequency based on number of months exceeding threshold	xxx		No theshold for *Phaeocystis* calculation of bloom frequency bloom
OSPAR Threshold for bloom definition	Until now: 10^6 cells/l, but for future proposed to use 4 x 10^6 cells/l	Until now: 10^7 cells/l, but for future proposed to use 4 x 10^6 cells/l			

The main differences in methods are:

- The UK and France do not distinguish *Phaeocystis* from other small flagellates
- WFD: Belgium uses a threshold of 4 million cells/l as bloom definition, the Netherlands use a threshold of 10^6 cells/l to calculate the frequency of blooms
- OSPAR: Belgium uses a threshold of 4 x 10^6 cells/l as maximum number of cells per year, the Netherlands use a threshold of 10^7 cells/l
- Belgium calculates the frequency of blooms as the frequency of samples exceeding the threshold. The Netherlands calculate the frequency of blooms as the frequency of months exceeding the threshold.
- Belgium uses a higher sampling frequency during the months with (generally) the highest *Phaeocystis* concentration
- The calculation method of Belgium gives more weight to the main bloom period March-May.

References

Lacroix, G., Ruddick, K., Gypens, N., Lancelot, C., 2007. Modelling the relative impact of rivers (Scheldt/Rhine/Seine) and Western Channel waters on the nutrient and diatom/*Phaeocystis* distributions in Belgian waters (Southern North Sea). Cont..Shelf Res. 27), 1422-1446.

Lancelot, C.; Rousseau, V.; Gypens, N.; Parent, J.-Y.; Bissar, A.; Lemaire, J.; Breton, E.; Daro, M.H.; Lacroix, G.; Ruddick, K.; Ozer, J.; Spitz, Y.; Soetaert, K.; Chrétiennot-Dinet, M.-J.; Lantoine, F.;Rodriguez, F., 2007. Advanced modeling and research on eutrophication AMORE II: final report. Belgian Science Policy: Brussel, Belgium. 84 pp.

Annex B - Creation of a time series of *Phaeocystis* data

A high-frequency time series, consisting of daily values of *Phaeocystis globosa* cell counts and chlorophyll-a concentrations, was created based on an existing time series. For this purpose, monitoring data from station Noordwijk 10 were used. This sampling station lies 10 km off the Dutch coast in the plume of the river Rhine outflow, and had a high monitoring frequency (weekly or bi-weekly in the period March-September) in the years 1980-1983 (Chl-a) and 1990-2010 (Chl-a, *Phaeocystis*). Based on the chlorophyll-a concentrations or cell counts of *Phaeocystis globosa* at each sampling date, daily values were calculated by interpolation using an exponential growth function.

The growth rate *g* between two sampling dates t_0 and t_1 (time interval Δt) was calculated from:

$$g = \ln(C_{t1}/C_{t0})/\Delta t$$

The concentration at any date t_x between two sampling dates was then calculated from

$$C_{tx}=e^{(g*\Delta t`)}*C_{t0}$$

(where $\Delta t`$ is the time interval between sampling date t_0 and date t_x)

With this interpolation method, the interpolated values between two sampling dates exponentially increase or decrease from the first sampling date to the next. A time series of daily values in the period March-September was created. Figure B.1 shows examples of the interpolated and the measured *Phaeocystis* cell concentrations in two years.

Calculation of metric values with various sampling frequencies

The time series with interpolated daily data was used to “sample” values with varying sampling frequencies, from daily to weekly to 2-, 3- and 4-weekly, in the period March-September.

For example, for the one-week interval sampling frequency, data were sampled at a 7-day interval. As the sampling can start at any of the 7 days in a week, this resulted in 7 different subsets of data per year. Similarly, with a 2-week interval this resulted in 14 different subsets per year (Table B.1). These subsets were used to calculate annual maximum and average concentrations (chlorophyll-a and *Phaeocystis*), 90-percentiles (chlorophyll-a) and annual bloom frequencies (*Phaeocystis*).

The bloom frequency was calculated using the method applied in the Netherlands: any month where a bloom is observed during sampling (*Phaeocystis* concentration $>10^6$ cells/l) is flagged as a bloom month; the bloom frequency is the percentage of bloom months in a year.

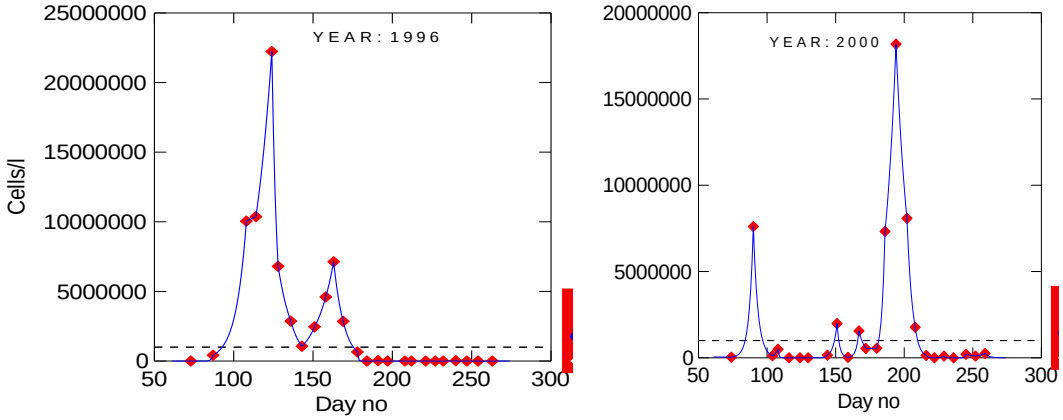

Figure B.1 Observed and interpolated cell concentrations of *Phaeocystis* at Noordwijk 10 in the years 1996 and 2000. The dashed line indicates the 10^7 cells/l level.

Table B.1. Number of subsets of data available per year.

Sampling interval	**N subsets of data per year**
Day	1
1 week	7
2 weeks	14
3 weeks	21
4 weeks	28