



## UvA-DARE (Digital Academic Repository)

### Monitoring and prediction of phytoplankton dynamics in the North Sea

Blauw, A.N.

**Publication date**

2015

**Document Version**

Final published version

[Link to publication](#)

**Citation for published version (APA):**

Blauw, A. N. (2015). *Monitoring and prediction of phytoplankton dynamics in the North Sea*. [Thesis, fully internal, Universiteit van Amsterdam].

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Chapter 1

## Introduction

### 1.1 BACKGROUND

The North Sea is a marine area of high economic and ecological value. It is a relatively shallow shelf sea, receiving nutrients from a large part of northern Europe. These factors contribute to a relatively high primary production by phytoplankton (also called microalgae). This primary production forms the basis of the marine food chain and supports populations of many species of fish, shellfish, marine mammals and other valuable marine organisms. The area is intensively used for shipping, fisheries, aquaculture, recreation and energy production.

#### 1.1.1 Changes in the North Sea

The North Sea ecosystem is subject to several pressures: eutrophication, climate change, increasing fishing pressure, pollution and many developments in coastal and marine infrastructure. In response to changing pressures the ecosystem has changed during recent decades and will most likely continue to change in the future.

During the 20<sup>th</sup> century concentrations of nutrients in rivers have increased due to the use of fertilizers in agriculture and increasing populations in western Europe. Klein and van Buuren (1992) estimated that nitrate concentrations in the Rhine were 3 to 4 times higher in 1990 than in 1935 while phosphate concentrations were 7 times higher. The increasing input of nutrients to the North Sea led to eutrophication issues, such as oxygen depletion in seasonally stratified waters in the Oystergrounds area and the German Bight (Niermann et al., 1990). Also the duration of nuisance algal blooms of *Phaeocystis* in the Marsdiep area has increased from the mid-seventies to the nineties due to eutrophication and decreased thereafter (Cadée and Hegeman, 2002). *Phaeocystis* blooms form a large part of the seasonal spring phytoplankton bloom in the southern North Sea, and cause foam accumulation on beaches (Figure 1.1). In response to these eutrophication issues North Sea bordering countries have joined forces to reduce the input of nutrients into the North Sea. This resulted in a reduction in phosphorus inputs of circa 50% on average

in the years 2000 – 2005, compared to 1985. Also modest reductions in nitrogen inputs have been achieved (OSPAR Commission, 2008).

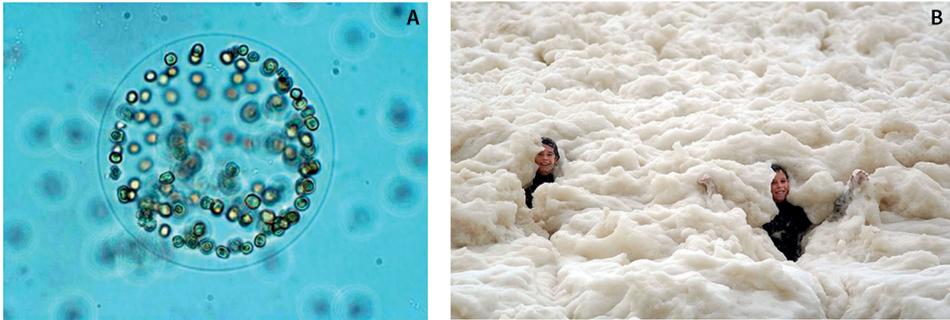


Figure 1.1: A) *Phaeocystis* colony, showing the spherical gel of mucus in which *Phaeocystis* cells are embedded (courtesy of Louis Peperzak) ; B) foam on beach attributed to *Phaeocystis* blooms ([http://cfb.unh.edu/phycokey/Choices/Prymnesiophyceae/PHAEOCYSTIS/Phaeocystis\\_Image\\_page.html](http://cfb.unh.edu/phycokey/Choices/Prymnesiophyceae/PHAEOCYSTIS/Phaeocystis_Image_page.html)).

More recently, global warming has led to increasing temperatures and other climate-related changes in the North Sea (Edwards and Richardson, 2004; Dulvy et al., 2008). Changes in river discharges and associated nutrients due to shifts in rainfall patterns have in the past affected phytoplankton dynamics, for example in Belgian coastal waters (Breton et al., 2006). Variation in currents and exchanges with Atlantic water, due to changes in wind patterns and ocean circulation, have also shown to have a significant impact on marine plankton (Beaugrand et al., 2002; Weijerman et al., 2005; McQuatters-Gollop et al., 2011).

Fishing activities have strongly intensified since modern technology was introduced in the early 1900s. Fish stocks have declined considerably since then, which is clearly expressed in the 94% decrease in the landings of demersal fish per unit of fishing power since 1880 (Thurstan et al., 2010, Figure 1.2). During the same period the abundance of benthic fauna in the North Sea has dramatically decreased, especially large slow-growing species, such as the ocean quahog (*Arctica islandica*) and common brittle star (*Ophiothrix fragilis*) (Callaway et al., 2007). Also many long-lived pelagic fish species including tuna, sharks and rays have experienced strong declines in abundance during the same period (Lindeboom, 1995; MacKenzie and Myers, 2007).

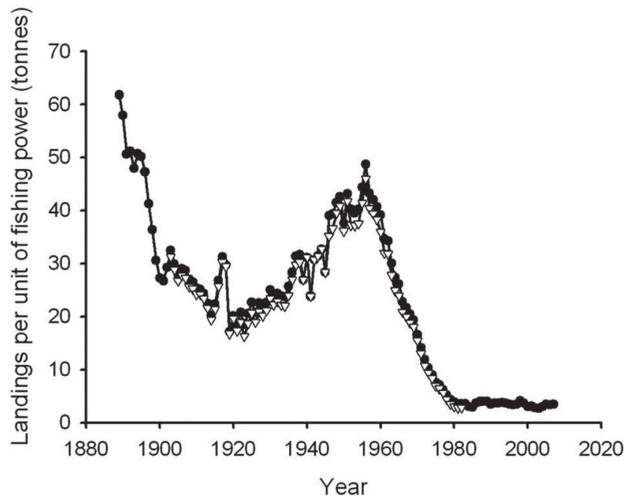


Figure 1.2: Tonnes of bottom-living fish landed per unit fishing power by bottom-trawl fisheries in the UK (from Thurstan et al., 2010).

Other pressures that may affect marine ecosystems include pollution with a wide range of chemical pollutants and large construction projects, such as the development of marine wind farms and harbours and deepening of shipping channels. These may influence transport and mixing processes in the water and thereby the growth conditions and spatial dispersal of planktonic organisms.

### 1.1.2 Need for monitoring

Coastal areas are often intensively used for different types of economic activity. The many different users, pressures and impacts on the North Sea described above are similar in many coastal areas worldwide. In such intensively used marine areas conflicts of interest may arise when different uses impact each other. For example, there is a long-standing debate between the conflicting interests of fisheries and nature conservation. Nature conservation organisations accuse fisheries of destroying populations and habitats of many marine species. Conversely, fisheries organisations blame the reduction of nutrient inputs, especially P, for declining fish stocks. Large infrastructure projects, such as the development of harbours or wind farms, need to perform environmental impact assessments to prove that nature will not be negatively impacted to get the required permits. And recreation along coasts benefits from attractive, healthy and safe waters, without bad odours, suspect discoloration or floating scums. Aquaculture requires good water quality as well, without bacterial and viral pollution from sewage or toxic algal blooms, but with high concentrations of phytoplankton as food source.

To manage these conflicting interests in coastal waters, arguments on the changing status of the North Sea due to pressures or mitigating measures should be supported by sound data. To this end many monitoring programmes are in place to assess changes in nutrient and phytoplankton concentrations, oxygen depletion events and populations of economically or ecologically important marine species.

Preferably, these monitoring programmes should enable to answer not only the question *if* a key variable is changing, but also *why* it is changing, so appropriate measures can be taken. Therefore, additional variables need to be measured, that may provide a better understanding of natural and human-induced variation in the key variables of the monitoring programme. For example, in the Dutch eutrophication monitoring programme for coastal waters, not only nitrogen, phosphorus, oxygen and chlorophyll (as an approximation of total phytoplankton biomass) are measured, but also dissolved silicate, water transparency, temperature and salinity. Data from this monitoring programme show that nutrient concentrations in coastal waters have decreased from 1988 to 1995, following the decrease in nutrient inputs from rivers. However, chlorophyll concentrations measured by the same monitoring programme do not show a significant decline (De Vries et al., 1998). The Marsdiep station, which is sampled much more frequently, also does not show a decrease of yearly averaged chlorophyll concentrations or of the magnitude of the spring bloom (Philippart et al., 2010). However, here a significant decrease of chlorophyll concentrations in autumn is observed from the early nineties to 2007. This example illustrates that the detection of changes in phytoplankton abundance is affected by the monitoring location, monitoring frequency and by the method used for quantification of phytoplankton abundance (average concentration, peak concentration, bloom duration, timing of blooms etc.).

Traditionally, measurements of phytoplankton and supporting variables are made by taking water samples from ships at weekly to monthly intervals. Increasing the sampling frequency of such monitoring programmes to better capture the natural variability of phytoplankton, is expensive. In recent decades, alternative monitoring methods are being developed such as automated buoys and satellite remote sensing, that enable monitoring of phytoplankton abundance at a high temporal and/or spatial resolution, without additional cruises by dedicated monitoring vessels. However, these monitoring methods are so far only used for research and not for trend and compliance checking. Reasons for not using these high resolution data include limited access and experience with these data.

### 1.1.3 Objectives of this thesis

In this thesis phytoplankton variability in the North Sea is investigated with a range of traditional and novel monitoring methods. Phytoplankton variability is characterized in

space and time, and the drivers of this variability are analyzed at different time scales with various methods for data analysis and modeling. In this way experience is gained with the analysis and interpretation of data from automated monitoring methods. Furthermore, the results will show how traditional and sensor-based monitoring strategies affect the information that can be gained from the resulting data sets.

This chapter starts with a description of available monitoring methods for phytoplankton (section 1.2). Subsequently, I give an overview of phytoplankton in the North Sea and the processes affecting its dynamics (section 1.3), followed by an introduction to relevant methods for data analysis and modeling (section 1.4). I conclude with an outline of the thesis (section 1.5).

## 1.2 MONITORING PHYTOPLANKTON ABUNDANCE

Phytoplankton abundance can be defined and measured in many ways, for instance in terms of chlorophyll concentration, chlorophyll fluorescence, number of phytoplankton cells, biovolume and dryweight. In ecosystem models phytoplankton is usually defined in carbon units, so it is consistent with units of carbon dioxide, oxygen and primary production. However, there are no methods available to measure phytoplankton directly in carbon units under field conditions. All these measures give a different approximation of phytoplankton abundance. Available methods to measure phytoplankton abundance are described below (section 1.2.1). Subsequently, the different platforms (moorings, ships, satellites etc.) are described that are used to measure phytoplankton abundance in the North Sea (section 1.2.2). Measurement methods differ in their applicability from different platforms and thus the attainable spatial and temporal resolution.

### 1.2.1 Measurement methods for phytoplankton

The most commonly used measure for phytoplankton abundance is the chlorophyll *a* concentration. Chlorophyll *a* is a pigment that is present in all cyanobacteria and algae, and therefore it is commonly used as approximation of total phytoplankton abundance. However, the chlorophyll content per cell varies over time and between species, so total chlorophyll per sample can also vary through other processes than changes in phytoplankton abundance (Riemann et al., 1989). The approach can be generalized to the measurement of different photosynthetic pigments, which provides information on the concentrations of different taxonomic groups (Mackey et al., 1996). Measurement of the full range of photosynthetic pigments can be done with high-performance liquid chromatography (HPLC). Another measurement method that is commonly used is spectrophotometry. This method can measure chlorophyll *a*, in a 2-step procedure. But

often the second step is omitted and chlorophyll is measured, including by-products of chlorophyll a.

Through the microscope phytoplankton cells can be counted and the species composition can be identified. Generally samples are concentrated before analysis, by letting the phytoplankton cells settle in an Utermohl chamber. This classic approach can provide highly detailed information on the species composition. However, small picocyanobacteria and pico-eukaryotes may not settle and/or are too small to be identified, resulting in an underestimation of their numbers and total phytoplankton abundance. Another potential bias in microscope cell counts is due to differences in interpretation and identification of phytoplankton species by different specialists handling the microscopic analysis (Culverhouse et al., 2003; Peperzak, 2010).

Flow cytometers automatically count particles in water and measure their fluorescence, scatter and absorption characteristics at different wavelengths in order to cluster phytoplankton particles in groups with similar properties (Sosik and Olson, 2007; Thyssen et al., 2008; Pomati et al., 2011). The number of different groups that can be distinguished is limited, compared to microscopic cell counts. However, flow cytometers can count large numbers of cells and are very effective in distinguishing the smaller pico-cyanobacteria and pico-eukaryotes. Some make pictures of each cell, to aid the identification of different species. Flow cytometers are being tested for flow-through operation along ships of opportunity (Thyssen et al., 2009).

Fluorometers measure fluorescence: the light emitted by phytoplankton pigments at specific wavelengths to get rid of excess light. Fluorometers are commonly used on moorings, such as the Smartbuoy (Greenwood et al., 2010; Capuzzo et al., 2013) and moorings of the BATS program (Dickey et al., 2001), in CTD profiles and on ships of opportunity (Petersen et al., 2008). Fluorescence can also be measured at different wavelengths, characteristic for different pigment types, to get a rough estimate of the major taxonomic groups (Jakob et al., 2005; Smythe-Wright et al., 2014). Fluorescence measurements are susceptible to non-photochemical quenching, causing an underestimation of chlorophyll levels when light intensities are high (Kiefer, 1973).

The Continuous Plankton Recorder (CPR) uses silk to filter water along transects of ships of opportunity since the 1940s (Beaugrand, 2004; McQuatters-Gollop and Vermaat, 2011; Edwards et al., 2013). The coloration of the silk (the phytoplankton colour index PCI) is classified in the lab and the species attached to the silk are identified. This method is biased towards the larger phytoplankton groups in the surface layer, that cannot pass through the silk. The resulting dataset has an exceptional spatial and temporal coverage, dating back to long before most other phytoplankton time series started and covering much of the North Atlantic.

Chlorophyll maps from satellite remote sensing are estimated from observed water-leaving radiance based on scattering and absorption characteristics of phytoplankton and other substances in the surface layer (Ruddick et al., 2001; Van der Woerd and Pasterkamp, 2008; Tilstone et al., 2012). Algal blooms that reach high biomass near the surface, such as *Phaeocystis* and *Karenia* blooms, can be observed with satellite remote sensing (Stumpf et al., 2003, 2009; Kurekin et al., 2014). However, many toxic dinoflagellates are already toxic at low concentrations, when they form only a small part of the total phytoplankton population. Also dinoflagellates often form subsurface blooms near the pycnocline, where they are too deep to be observed by satellite remote sensing.

Various molecular monitoring methods for specific harmful species are being developed. Rinta-Kanto et al. (2005), for example, used qPCR (quantitative polymerase chain reaction) for real-time monitoring of *Microcystis* in Lake Erie. In marine waters, Greenfield et al. (2008) could accurately measure time series of several potentially harmful algal species using two molecular probe techniques: a sandwich hybridization assay (SHA) and fluorescent in situ hybridization (FISH).

### 1.2.2 Platforms for phytoplankton monitoring

Based on spatial coverage, monitoring data can be clustered into “point” data at fixed stations, “line” data along transects and “map” data. Table 1.1 gives an overview of the applicability of the methods described above to measure phytoplankton abundance for different platforms. Methods that require extensive processing and analysis in a laboratory are most suitable for monitoring at fixed stations. Automated sampling methods that can measure phytoplankton directly in the water are suitable for application at fixed stations and along transects. For a map view of phytoplankton abundance, remote sensing methods are needed that estimate phytoplankton abundance from some distance above the water.

Table 1.1: Applicability of measurement methods for different monitoring platforms.

Method \ platform	Laboratory analysis (water samples from piers or ships)	Automated sampling (moorings, Ferrybox, gliders, vertical profilers)	Remote sensing (cameras on satellites, airplanes, watchtowers, ships)
Chlorophyll concentration	X		
Microscopy	X		
Flow cytometry	X	X	
Fluorescence	X	X	
Molecular techniques	X	depends on method	
Water colour: absorption and scatter	X	X	X

### *“Point” data at fixed stations*

Many monitoring programs use fixed monitoring locations, assuming that these stations are representative of the surrounding area. The temporal resolution of such “point” data can vary from several observations per hour from moorings, one to seven observations per week from piers to weekly, biweekly or monthly observations from ships. An example of a monitoring programme with moorings in the North Sea is provided by CEFAS. They have deployed so-called Smartbuoys (Figure 1.3) at several stations in the North Sea, measuring chlorophyll fluorescence, turbidity, temperature, salinity, light, oxygen and nitrate for many years at high resolution (Kröger et al., 2009; Greenwood et al., 2010). Data from four of these Smartbuoys are analyzed in chapters 6 and 7 of this thesis. Examples of pier monitoring stations in the North Sea include Helgoland (Wiltshire et al., 2010; Kraberg et al., 2011), Flødevigen (Dahl and Johannessen, 1998; Naustvoll et al., 2012) and Marsdiep (Cadée and Hegeman, 2002; Philippart et al., 2010). Time series of weekly phytoplankton measurements from ships are collected at station L4 along the English south coast since 1992 (Harris, 2010; Widdicombe et al., 2010) and at Stonehaven in Scottish coastal waters since 1997 (Bresnan et al., 2009). Furthermore, fortnightly to monthly phytoplankton measurements are collected by national governments bordering the North Sea at various locations: station 330 in Belgian coastal waters since 1989 (Breton et al., 2006) and 18 stations in Dutch coastal waters since 1991 (Baretta-Bekker et al., 2009; Blauw et al., 2010; Peperzak, 2010). Measurements are mostly collected at the water surface or integrated over the top few meters. Occasionally, water samples are taken at different depths or vertical profiles of fluorescence are measured.

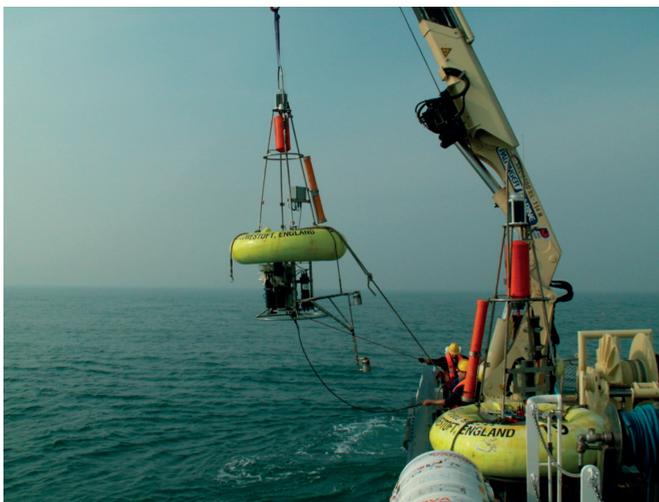


Figure 1.3: A Smartbuoy mooring being deployed from a ship (courtesy of CEFAS).

### *“Line” data along transects*

Ships that automatically collect data on phytoplankton on their way are called ships-of-opportunity or voluntary observing ships (VOS). Often ferries are equipped with automated sampling devices as they go back and forth along fixed transects. The temporal resolution of such “line” data depends on the frequency of the ferry service, typically between daily and monthly. The intake of water for sampling by these ships is typically at a fixed depth, a few meters below the surface. Well-known examples of phytoplankton monitoring along transects are the Continuous Plankton Recorder (CPR) (Edwards et al., 2013), the Alg@line (Ruokanen et al., 2003) and Ferrybox (Petersen et al., 2008). Figure 1.4 shows an example of measurement results from the Ferrybox.

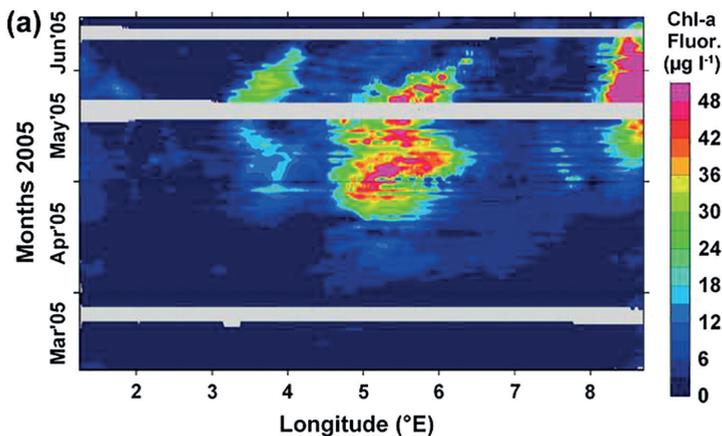


Figure 1.4: Example of results of a Ferrybox: chlorophyll fluorescence along a transect between Cuxhaven (Germany) and Harwich (Great Britain) from March 2005 through June 2005 (from Petersen et al., 2008).

### *Map data*

Satellite remote sensing data can provide “map” data of various variables, including chlorophyll, total suspended matter and temperature in the surface layer, with a temporal resolution depending on the satellite frequency (typically daily) and cloud cover (Figure 1.5). Cloud cover varies between regions of the world and with seasons. The depth of the surface layer for which the satellite remote sensing data are representative varies with the transparency of the water. Capuzzo et al. (2013) estimated these “optical depths” at three sites in the North Sea: satellites can see up to 2 m depth in the English turbidity plume, 4.5 m in the seasonally stratified central North Sea (Oystergrounds) and 10 m in the clear waters north of the Dogger bank.

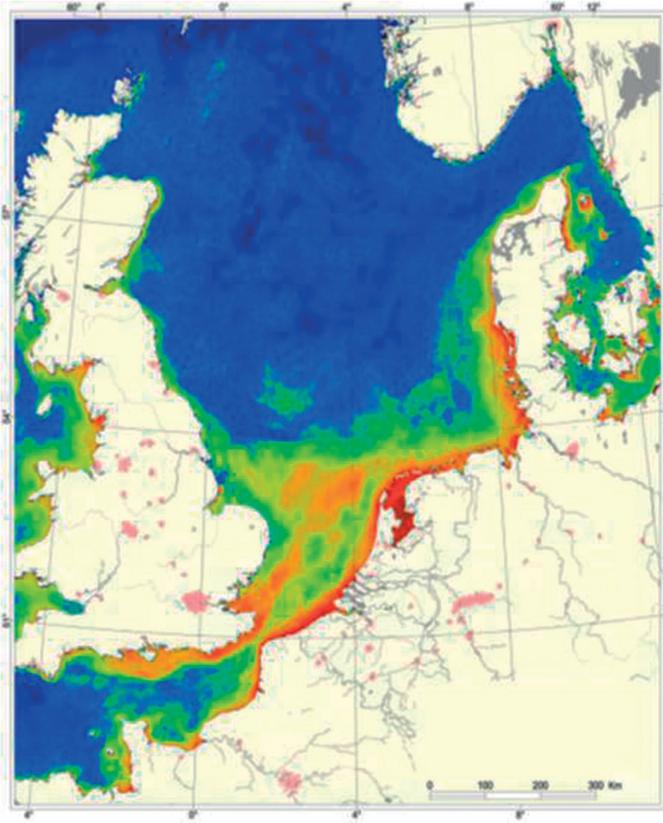


Figure 1.5: Median chlorophyll concentrations in spring 2003 based on MERIS satellite data (from REVAMP atlas, Peters et al., 2005).

### 1.3 PHYTOPLANKTON IN THE NORTH SEA

Important issues that are directly linked with phytoplankton variability in the North Sea include eutrophication, carrying capacity and toxic blooms. Too high phytoplankton biomass, due to eutrophication, may negatively affect ecosystem health through mortality of benthic organisms after oxygen depletion or insufficient light near the bottom to support sea grass meadows. Too low phytoplankton biomass may negatively affect the carrying capacity of an area for shellfish and fish, if the primary production by phytoplankton limits the growth of organisms feeding (indirectly) on phytoplankton. For these issues the main interest is in the variables total phytoplankton abundance and primary production. Additionally, *Phaeocystis* is considered an indicator of eutrophication,

and blooms of this species are associated with nuisance foam accumulation on beaches (Lancelot, 1995).

Toxic blooms of *Dinophysis* or *Alexandrium* species occasionally cause damage in the aquaculture areas in the Oosterschelde estuary in the south of the Netherlands (Van der Fels-Klerx et al., 2012; Burson et al., 2014), and in coastal waters of the UK (Brown et al., 2001; Bresnan et al., 2009), Denmark (Hansen et al., 2003) and Norway (Naustvoll et al., 2012).

Main factors that are known to influence temporal and spatial variability of marine phytoplankton are physical transport and mixing patterns, the availability of nutrients and light, and food web interactions.

### 1.3.1 Physical transport

The North Sea is a shelf sea between the United Kingdom and continental Europe. Atlantic waters enter through the Channel on the southern end and along the Scottish coast on the northern end. The Coriolis effect, caused by the rotation of the Earth, creates a counter-clockwise residual current. Continental rivers entering the sea tend to stay close to the coast and flow to the right. Atlantic waters entering the North Sea along the Scottish coast flow south along the UK coast until they meet waters flowing north from the Channel. There both currents join, cross the North Sea and flow north along the Dutch, German and Danish coasts. Here, fresh water discharges from continental Europe enter the North Sea through major rivers, such as the Meuse, Rhine, Ems and Elbe. This river input forms a band of relatively fresh and nutrient-rich water along the eastern shores of the North Sea. North of Denmark brackish water from the Baltic Sea enters the North Sea through the Kattegat and Skagerrak. Most water leaves the North Sea with the residual current along the Norwegian Coast (Figure 1.6).

The residual current pattern described above is perturbed by tidal and wind-driven currents. Tidal currents are driven by ebb and flood, moving waters back and forth with a 12-hour periodicity. Wind-driven currents may cause strong deviations from the residual current patterns described above, for example high concentrations of Rhine water may occasionally be observed in Belgian coastal waters (Lacroix et al., 2004).

High-resolution time series of chlorophyll tend to show a distinct 12-hour periodicity caused by waters with different phytoplankton concentrations moving back and forth with the tidal ebb and flood currents (Brunet and Lizon, 2003; Chen et al., 2010). Also phytoplankton blooms are transported with the current as they grow. This means that large stretches of coast may be affected by the same bloom, as the currents carry a growing phytoplankton population downstream. Hence, local conditions sometimes cannot explain the development of blooms, if these blooms originated elsewhere.

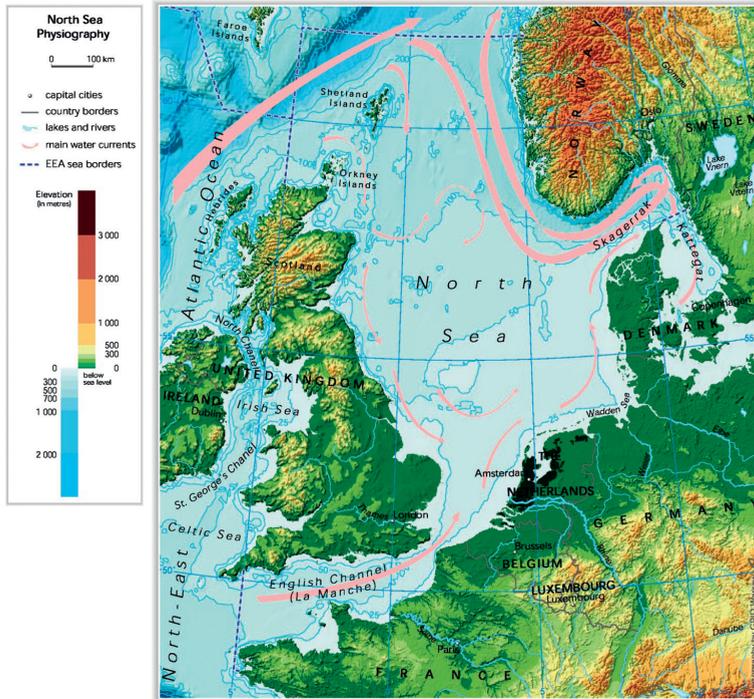


Figure 1.6: Map of the North Sea: bathymetry and residual currents (from <http://www.eea.europa.eu/data-and-maps/figures/>).

### 1.3.2 Mixing processes

The southern part of the North Sea is relatively shallow and has relatively large tidal ranges. Towards the north the sea becomes gradually deeper, with lower tidal ranges (Figure 1.7). Therefore, the southern part of the North Sea is vertically mixed year-round, whereas the northern part is seasonally stratified.

Stratification is caused by density differences between the upper and lower layers of the water column. As these density differences increase, for example due to heating of the surface layer by the sun, increasingly more energy is needed to mix the two water layers. In seasonally stratified areas, such as the Oystergrounds area in the central North Sea, a distinct warm surface layer of several meters deep is observed during the summer months (Sharples et al., 2006). Since vertical mixing is strongly reduced in such stratified systems, particles that have sunken to deeper water layers do not return to the upper layer until the system is mixed again by autumn storms (Greenwood et al., 2010). Therefore, the surface layer is very clear but also poor in nutrients. Nutrients that have been taken up for

growth by phytoplankton sink as organic particles and the nutrients remain in the lower layer after remineralisation. Since the upper layer is depleted of nutrients, phytoplankton abundance in such stratified systems is generally highest at the interface between the upper and lower layer. Here, a so-called “deep chlorophyll maximum” is formed (Weston et al., 2005).

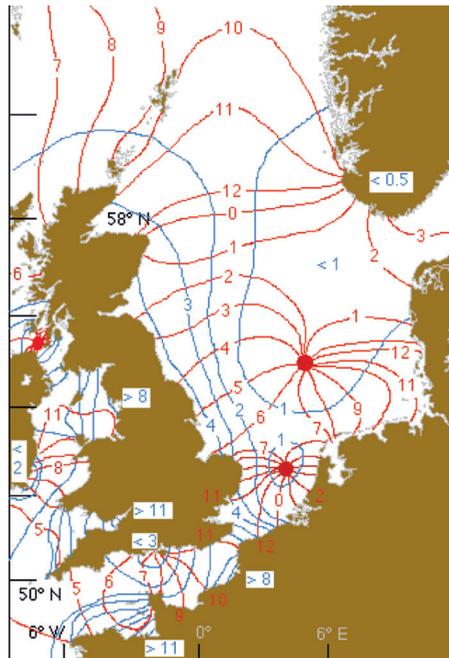


Figure 1.7: Tides in the North Sea as derived from observations. Red lines are co-phase lines of the M2 tide, labelled in hours after the moon's transit through the meridian of Greenwich. Blue lines give the mean tidal range at spring tide (from: <http://gyre.umeoce.maine.edu/physicalocean/Tomczak/IntroOc/notes/figures/fig11a5.html>).

In addition to temperature stratification, salinity stratification due to inflows of fresh and brackish waters also occurs in the North Sea. Due to the large discharge of brackish water from the Baltic Sea into the deep Skagerrak area the Norwegian coastal current is permanently stratified (Lekve et al., 2006). Also in the Rhine plume along the shallow Dutch coast salinity stratification occurs, but only temporarily. In these shallow waters spring tides and strong winds eventually mix up the low-salinity top layer with the higher salinity sea water (Simpson et al., 1993; De Boer et al., 2009).

### 1.3.3 Nutrients

Phytoplankton can continue to grow until one of the resources for growth gets depleted. Therefore, phytoplankton can reach higher concentrations in areas where nutrient concentrations are high. In coastal waters nitrogen, phosphorus or light are the resources that are most likely to limit phytoplankton growth. For diatom species silicate availability may also limit growth. Nutrient concentrations are generally higher in river water than in sea water. Therefore, nutrient availability is positively correlated with the fraction of fresh water and inversely correlated with salinity (de Vries et al., 1998). This is reflected by relatively high chlorophyll concentrations in coastal waters compared to off-shore waters (Radach and Gekeler, 1996; De Vries et al., 1998; Figure 1.5).

Although the positive correlation between nutrient availability and phytoplankton abundance is obvious at large temporal and spatial scales, it is much less clear at smaller scales. In time series, variation in phytoplankton abundance within the year is often negatively correlated with nutrient availability, since nutrients that have been taken up by phytoplankton are not available anymore in the water column (e.g. Lekve et al., 2006). Correlations of individual nutrients (nitrogen, phosphorus and silicate) with phytoplankton abundance are often confounded by the effects of the other limiting resources, including light. If phytoplankton growth is limited by phosphorus but nitrogen is available in excess, variation in nitrogen concentrations has only minor effects on phytoplankton growth rates. Furthermore, in addition to nutrient concentrations, nutrient ratios may also affect the phytoplankton species composition, since some species are better competitors for nitrogen while others are better competitors for phosphorus (Tilman, 1977; Brauer et al., 2012).

Several phytoplankton species are able to use both inorganic and organic nutrients (Antia et al., 1991). For instance, toxic dinoflagellates of the *Alexandrium* genus (Figure 1.8A) can take up dissolved organic nutrients, such as urea (Dyhrman and Anderson, 2003). Other dinoflagellates, such as *Dinophysis* (Figure 1.8B), can even eat other algae and use their biomass as nutrient and energy source. *Dinophysis* is even able to preserve the photosynthetic pigments of its prey to use it for its own energy production. These “stolen pigments” are called kleptoplastids (Park et al., 2006; Carvalho et al., 2008).

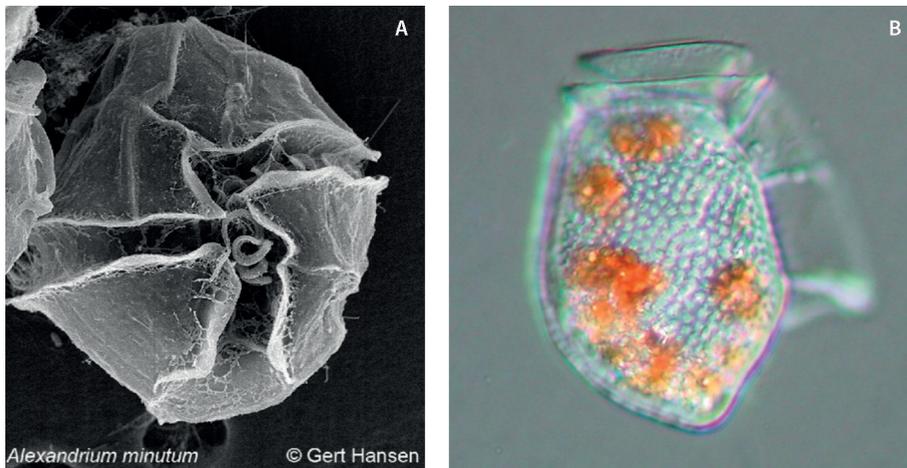


Figure 1.8: Pictures of toxic phytoplankton species in the north Sea: **A)** electron microscope image of *Alexandrium minutum* (from [www.marinespecies.org](http://www.marinespecies.org)) and **B)** light microscope picture of *Dinophysis acuta* (from <http://nordicmicroalgae.org>).

### 1.3.4 Light and temperature

Phytoplankton need light to perform photosynthesis: the production of biomass from light energy and inorganic substances. The availability of light for phytoplankton is affected by the intensity of solar irradiance, the extinction of light in water and the mixing depth. Extinction of light with depth is caused by absorption and scattering by suspended particles, including phytoplankton, dissolved substances and water itself (Kirk, 1994; Stomp et al., 2007). Mixing depth has a large impact on the light climate experienced by phytoplankton. Since light intensity decreases exponentially with depth, phytoplankton in shallow waters or shallow surface mixed layers receive higher light intensities than those in deep waters or deeply mixed waters, where they are forced to stay in dim light much of the time.

The availability of light affects both the temporal and spatial variability of phytoplankton. Spatial variability in light availability is caused by a combination of mixing depth and turbidity. Near-shore waters tend to be both shallower and more turbid than waters further offshore. So these effects counteract each other and, contrary to nutrient availability, there is no clear cross-shore gradient in light availability. Still in very turbid estuaries primary production is often found to be limited by light availability (Cloern, 1987). Also in various turbid coastal areas of the North Sea light availability limits phytoplankton growth throughout the year (Colijn and Cadée, 2003; Loebel et al., 2009).

Temporal variability of phytoplankton is strongly affected by periodic fluctuations in the intensity of solar irradiance: the seasonal cycle and the day-night cycle. In the North Sea, the seasonal cycle of solar irradiance leads to relatively low phytoplankton concentrations but high nutrient availability in winter. During the onset of stratification in spring, mixing depth quickly decreases and light intensity experienced by phytoplankton increases. Therefore, the onset of stratification is often the trigger for spring blooms in seasonally stratified waters (Sverdrup, 1953; Huisman et al., 1999). The spring bloom generally starts later in turbid near-shore waters than in clear offshore waters. After the spring bloom, phytoplankton concentrations may be reduced to lower levels throughout the summer when nutrient limitation limits phytoplankton growth and grazers on phytoplankton are abundant. In some areas also an autumn phytoplankton bloom is observed (Sommer et al., 2012).

The day-night cycle in solar irradiance creates strongly fluctuating growth conditions for phytoplankton. Part of the day there is no light available for growth and part of the day there can be too much light available for optimal photosynthesis. Too much light can damage the photosynthetic pigments if phytoplankton do not have developed defense mechanisms. One of these defense mechanisms is non-photochemical quenching, where the absorbed light energy is not used for photosynthesis but released as heat. Another effect of the day-night cycle in solar irradiance is that the surface layer quickly heats up during the day, which strongly reduces vertical mixing in the upper layers and thereby induces diurnal variation in stratification. In the central North Sea, Van Haren et al. (1998) found that water column stabilization during the day led to sinking of diatoms from the surface to deeper layers. Convective mixing at night by cooling of the water surface brought the diatoms back into the surface mixed layer. Other phytoplankton species can swim so they can vertically migrate to a position where light and nutrient conditions are optimal (Olsson and Granéli, 1991). Vertical migration often shows a day-night cycle as well, with phytoplankton moving up to the light during the day and down to the nutrients during the night. If vertical mixing in the water column is too strong, such vertical migration strategies are less successful.

### 1.3.5 Food web interactions and chaos

Phytoplankton forms the base of the marine food web. It is the main food source for many species of shellfish and zooplankton. If these grazers are abundant they can cause a high phytoplankton mortality. For example in the Eastern Scheldt estuary a high abundance of mussels and oysters leads to low chlorophyll concentrations in clear and nutrient rich waters (Prins et al., 1997). In coastal waters a peak abundance of zooplankton is often observed following the spring phytoplankton bloom (Sommer et al., 2012). Due to this high zooplankton abundance the spring bloom ends more abruptly than it would have

without zooplankton (Calbet et al., 2003). Also virus infections can cause an abrupt end of phytoplankton blooms (Brussaard et al., 2005)

Phytoplankton species differ in their competitive abilities for nutrients and light, their optimum temperature and salinity ranges, their vulnerability to turbulence and high light stress and their resistance to grazing. If all these environmental conditions would be constant over a long period, one might expect that the best adapted species for these conditions would outcompete all other species. However, Hutchinson (1961) already concluded that this does not happen in nature. In fact, a single milliliter of water may contain tens, if not hundreds, of coexisting phytoplankton species. This apparent paradox became known as “the paradox of the plankton”.

Huisman and Weissing (1999) showed that chaos theory can provide an explanation for this paradox. In models with 3 or more phytoplankton species competing for 3 nutrients at constant external conditions none of the phytoplankton species outcompeted the others. Instead, each species bloomed and decayed, following the other in a never-ending series of blooms. The system never reached equilibrium. Chaos theory was proposed by Lorenz based on his experience with weather models (Lorenz, 1963). It is defined as sensitive dependence on initial conditions. This means that a small difference in the state of the system, such as a minor variation in nutrient concentration or phytoplankton abundance, tends to grow exponentially over time. So after a while the bloom development can be dramatically different between two systems that started almost the same. This effect is also known as the butterfly effect. It is typically caused by several non-linear interacting processes and it is the main reason why weather predictions of more than a few days ahead tend to be rather unreliable. Benincà et al. (2008) demonstrated the presence of chaos in a plankton community kept under constant laboratory conditions for 9 years. Despite the constant conditions, the plankton community never reached a stable equilibrium, but the species abundances continued to fluctuate over several orders of magnitude. They estimated that sensible predictions could be made no more than 15-30 days ahead. These findings suggest that chaotic dynamics are likely to limit the predictability of phytoplankton abundance, even if the environmental conditions and their effects on phytoplankton growth and mortality rates would all be well known.

## 1.4 METHODS FOR DATA ANALYSIS AND MODELLING

Data analysis and modelling are closely related activities. Data analysis aims to identify patterns in data and relations between explanatory and response variables. Models aim to quantify these patterns and relations, to enable predictions for new similar situations. A large variety of modelling approaches is available ranging from empirical models to

mechanistic models. Empirical models are based on relations between observed data and explanatory variables, as derived with statistical methods. Mechanistic models describe relations between explanatory and response variables based on understanding of the underlying processes. In practice most model applications combine mechanistic and empirical components. Empirical models include options to control the shape of the model, based on understanding of the underlying mechanisms. And mechanistic models often include empirical relations between variables that have been identified in earlier studies. Furthermore, mechanistic models can be automatically calibrated to observed data, using inverse modelling methods. In this case the parameter values in the mechanistic model are estimated empirically.

Often the availability of observed data is limiting the complexity of (both empirical and mechanistic) models, in terms of the number of parameters that can be estimated reliably from the data. Empirical models are therefore vulnerable to overfitting if the chosen model is too complex for the supporting data. This leads to bad model performance when the model is applied to other similar data sets. Mechanistic models are vulnerable to overparameterisation if the chosen model is too complex for the supporting data. This leads to similar model performance with many different combinations of model settings, such that the “real” model parameters cannot be identified through calibration. Below some common approaches for data analysis and modelling are described, including those used in this thesis. Table 1.2 compares the model structure of the different modelling approaches described in this chapter.

**Table 1.2:** Overview of the model structure of different modelling approaches.

Model name	Model structure
Linear regression	$Y = a + bx_1 + cx_2 + dx_3$
Generalized linear model	$Y = f(a + bx_1 + cx_2 + dx_3)$
Generalized additive model	$Y = a + f(x_1) + g(x_2) + h(x_3)$
Neural network	$Y = f(x_1, x_2, x_3)$
Decision tree	<i>If <math>x_1 &gt; a</math>: <math>Y = bloom</math> If <math>x_1 &lt; a</math> AND <math>x_2 &gt; b</math>: <math>Y = bloom</math> If <math>x_1 &lt; a</math> AND <math>x_2 &lt; b</math>: <math>Y = no bloom</math></i>
Mechanistic model	$dY/dt = growth - mortality + influx$ $growth = f(x_1, x_2, x_3) * Y$ $mortality = g(x_4) * Y$ $influx = h(x_5)$

### 1.4.1 Data analysis methods

A first step in the identification of relations between variables is usually the calculation of correlations. Sometimes an explanatory variable affects a response variable with a time

lag. Such delayed responses can be identified with cross correlations. These show the correlation between variables over a range of time lags in a time series. Cross correlation analysis is performed as part of chapter 7. Autocorrelation is a special kind of cross correlation where a time series is correlated with itself at different time lags. Many time series are to some extent autocorrelated, because it takes time for a variable to change. Therefore the phytoplankton concentration at a specific moment in time is likely to be close to the concentration shortly before. Autocorrelation is an important issue in the analysis of time series. Most statistical methods, for example for significance testing, assume that the underlying data are independent. Time series data are generally not independent, so statistical procedures need to be corrected for autocorrelation within the data set (Brown et al., 2011).

Time series often show variability at different time scales, such as long-term trends, seasonal variability and fluctuations at shorter time scales. Time scale decomposition is commonly applied to separately study phenomena at different time scales. Time series analysis methods are generally designed to study fluctuations around a mean value. If the mean changes, due to a long-term trend, the time series is detrended. First, the local mean is estimated, for example by a moving average or a linear trend line, and then this local mean is subtracted from the time series (Chatfield, 2013). In this way the variability of a time series is split into a long-term trend and short-term fluctuations around that trend. This type of time scale decomposition has been applied in chapter 7 of this thesis. Alternatively the effects of trends and seasonal patterns on the analysis of short term fluctuations can be reduced by differencing the time series, i.e., by looking at the rate of change of phytoplankton concentrations instead of the actual phytoplankton concentrations. This approach is applied in chapter 6, and also in chapter 7 for the seasonal time scale.

If cyclic patterns, such as seasonal cycles, tidal cycles or predator-prey cycles, are the object of study spectral analysis is often applied. Traditionally, Fourier analysis is a common method, for example for the analysis of tidal components in time series of water levels. For the analysis of cyclic patterns in ecological time series Fourier analysis is less suitable, since ecological series are often not stationary. Wavelet analysis is a type of spectral analysis that is more suitable for ecological time series, because it does not require time series to be stationary (Cazelles et al., 2008). Wavelet analysis approximates fluctuations of a variable as a combination of localized periodic functions (wavelets) of different frequency (Torrence and Compo, 1998). Wavelet analysis is used to analyze periodicities in fluorescence time series in relation to the tidal cycle and day-night cycle in chapters 6 and 7 of this thesis. As an extension of wavelet analysis, wavelet coherence analysis identifies correlations and phase delays between oscillations of two time series. Wavelet coherence analysis has been applied in chapter 6 of this thesis.

### 1.4.2 Empirical models

The most commonly used empirical model is linear regression. In its simplest form the response variable is regressed against one explanatory variable as a linear relation. In multiple linear regression the response variable is fitted as the sum of linear relations with several explanatory variables. Non-linear relations can sometimes be fitted as well with linear models: by transforming the data of the explanatory variables or by using a transfer function (as in generalized linear models). Logistic regression models are a type of generalized linear model that is commonly used to simulate the frequency or probability of phytoplankton bloom events as a function of environmental variables (Lane et al., 2009).

Generalized Additive Models (GAMs) fit smooth functions of any shape between an explanatory variable and a response variable (Hastie and Tibshirani, 1990; Wood, 2006). Effects of individual explanatory variables are assumed to be independent, hence the name “additive”. An advantage of GAMs is that the flexibility of the fit can be controlled, to reduce overfitting. And the shape of the fitted “smooth functions”, that represent the effect of each explanatory variables on the response variable can be easily visualized. This is helpful to create understanding of the underlying mechanisms. This method is used in the North Sea for example by Llope et al. (2009) and in chapter 7 of this thesis.

Neural network models can flexibly fit relations between multiple variables of any shape. Benincà et al. (2008) fitted a neural network model to time series of several interacting functional groups to investigate the relations in the food web of a mesocosm kept at constant conditions for many years. Velo-Suárez and Gutiérrez-Estrada (2007) used neural networks to predict the development and transport of *Dinophysis* blooms in Spanish coastal waters.

Decision trees and fuzzy logic models use rule-based representations of interactions among parameters. They can be derived from expert knowledge, data-mining, or a combination of both. The most important difference between the two approaches is that in decision trees, the model parameters are classified in crisp classes, whereas in fuzzy logic models, they are classified in ‘fuzzy’, overlapping classes (Zadeh, 1965). Within fuzzy logic, the extent to which an element is considered to fulfil a certain criterion is not restricted to ‘true’ or ‘false’, but can be characterised by a value between zero (completely false) and one (completely true). An advantage of fuzzy logic is the possibility to combine qualitative and (partially) quantitative knowledge of physiological and physical processes with a certain degree of uncertainty (Droesen, 1996). Ibelings et al. (2003) applied fuzzy logic to successfully predict the occurrence of surface scums of the potentially toxic cyanobacterium *Microcystis* in Lake IJsselmeer, the Netherlands. In Blauw et al. (2006, 2010), we modeled algal blooms of five different species in five European marine

waters with fuzzy logic and found that this method is very useful to synthesize present knowledge and use it for predictions (chapters 4 and 5 of this thesis).

### 1.4.3 Mechanistic models

Mechanistic models describe the processes affecting phytoplankton concentrations step by step to calculate the changes in phytoplankton concentrations over time. Time is divided into finite time steps and for each time step the phytoplankton concentration is calculated based on the values of relevant variables in the previous time step. These models can describe phytoplankton concentrations in small well-mixed water bodies, such as laboratory cultures (Huisman and Weissing, 1994). For the simulation of larger water bodies the spatial domain is divided into a one- to three-dimensional grid that consists of finite segments. Many models have been developed that simulate physical, chemical and biological processes in the North Sea, for example ERSEM (Beukema et al., 1995; Baretta et al., 1995; Edwards et al., 2012), NORWECOM (Skogen et al., 1995), MIRO (Lancelot et al., 2005; Lacroix et al., 2007) and GEM (Blauw et al., 2009 [chapter 2 of this thesis]; Los and Blaas, 2010). These coupled physical-biological models simulate total phytoplankton biomass, expressed as chlorophyll, for comparison with field observations. Algal species are generally grouped together in two to four functional groups (Moll and Radach, 2003). Some models in the North Sea simulate *Phaeocystis* as a separate species together with groups of other marine species (Blauw et al., 2009; Lancelot et al., 2005). Stock et al. (2005, 2007) developed a model for *Alexandrium fundyense* in the Gulf of Maine, which was supported and validated by an extensive monitoring campaign, including cyst bed mapping.

### 1.4.4 Predicting the unpredictable

Due to non-linear relations between many interacting variables, chaos theory predicts that even the best possible model will not be able to predict day-to-day development of phytoplankton abundance far ahead. How can we cope with this inherent unpredictability? One approach is to predict ahead from operational monitoring data. With this approach you only need to make short-term predictions. This approach is adopted for example by Stumpf et al. (2009) in the Gulf of Mexico. When satellite remote sensing images show algal blooms in off-shore waters, on-shore transport of these blooms is predicted by numerical models. Other approaches, which are commonly used in weather forecasts, are ensemble forecasts and data assimilation. Ensemble forecasts show the increase of uncertainty of predictions with lead time, by starting several model predictions from slightly different initial conditions. This results in a plume of model forecasts, with trajectories close together a few days ahead and further apart at longer lead times, which is typical for chaotic dynamics. This approach has been adopted for algal bloom

predictions by Roiha et al. (2010). With data assimilation, models are constantly updated with recent field observations, to reduce any deviations from reality in an objective and consistent way (Gregg, 2008).

## 1.5 OUTLINE OF THIS THESIS

An important objective of this thesis is to characterize the variability of phytoplankton in time and space. Based on this knowledge the optimal sampling interval and locations can be selected. Phytoplankton concentrations show variability at different scales. At small scales, turbulent mixing and transport cause local fluctuations in concentrations. At larger scales, environmental conditions affect growth and mortality rates of phytoplankton. This results in seasonal variability and spatial gradients. Spatial gradients do not only occur horizontally. Also vertical profiles of phytoplankton concentrations are not always uniform. For many variables, such as water temperature, seasonal variability is a rather smooth pattern with limited inter-annual variability. Seasonal variability of phytoplankton, on the contrary, is characterized by bloom events lasting a few weeks. The timing, intensity and duration of these blooms differ between years and regions. Time series from mooring stations, spanning multiple years, provide an excellent opportunity to quantify the temporal variability of phytoplankton concentrations at these different time scales.

Another important question is: "What is driving this variability?" Moorings can only measure at high temporal resolution at a limited number of locations. Hence, a better understanding of the drivers of phytoplankton variability at the mooring sites may enable improved estimation of phytoplankton variability and optimal sampling intervals in the rest of the North Sea. Understanding the environmental drivers of phytoplankton variability also helps to predict how phytoplankton concentrations will change in response to changes in the environment. This can be very helpful, for example to estimate how phytoplankton concentrations are expected to change in response to global climate change or reduction of nutrient inputs. At shorter time scales this understanding can help to predict harmful algal blooms.

The optimal monitoring strategy depends not only on the variability of phytoplankton concentrations, but also on the type of questions that need to be answered and the level of statistical significance that is required for the answers. High-resolution data sets allow for a higher level of significance than low-resolution data sets. Furthermore, high-resolution data can address phytoplankton variability at shorter time scales, including bloom events. However, the analysis of high-resolution data poses some additional challenges. One of these challenges is the correction for autocorrelation in high-resolution data. Also data gaps complicate the analysis of high-resolution data sets. And automated measurements

need to be calibrated and validated with data from traditional monitoring methods. Hence, before we can benefit from the added value of high-resolution data, suitable methods for data analysis need to be developed and tested.

In this thesis, I have analyzed the response of phytoplankton abundance to environmental conditions with several types of traditional and novel monitoring data and various methods for data analysis and modeling. Our results demonstrate the potential power of these novel techniques, providing a much improved understanding of population fluctuations at several different time scales. The chapters in this thesis can be grouped into three approaches: in chapters 2 and 3 the mechanistic model “GEM” is validated with different types of observed data, chapters 4 and 5 make use of fuzzy logic models, and in chapters 6 and 7 time series analysis methods are applied to high-resolution mooring data.

Chapter 2 describes the set-up of GEM and validates the model results for spatial and seasonal patterns of chlorophyll and *Phaeocystis* across the North Sea. This approach supports analyses of long term trends, for example on eutrophication. In chapter 3 a combination of GEM results, remote sensing data, ship-based samples and pier-based *Phaeocystis* cell counts is used to reconstruct the spring *Phaeocystis* bloom of 2003 in Dutch coastal waters. This forms the basis for an early warning system based on near real-time monitoring data and harmful algal bloom prediction.

In chapters 4 and 5 fuzzy logic models are applied. In chapter 4 data from a Smartbuoy mooring, ship-based water samples and ARGUS video monitoring are combined to construct and validate a fuzzy logic model of *Phaeocystis* blooms and associated nuisance foam events along the Dutch coast. This study clearly shows the added value of data series of high temporal resolution for understanding mechanisms of harmful algal bloom events. Chapter 5 describes the development and validation of fuzzy logic models of harmful algal blooms of five different algal species in five European marine areas. This paper summarizes the results of the European research project HABES (Harmful Algal Blooms Expert System). It shows that fuzzy logic is suitable for predicting harmful algal blooms for a wide range of marine areas and phytoplankton species, provided that sufficient data and understanding of underlying mechanisms are available.

Chapters 6 and 7 describe the analysis and modelling of chlorophyll fluorescence observed by “Smartbuoy” moorings in the North Sea. These moorings provide high-resolution data measured at intervals of 12 to 30 minutes for many years. This offers an outstanding opportunity to obtain new insights on short-term variability. In chapter 6 wavelet analysis and wavelet coherence analysis are used to analyze drivers of fluctuations in chlorophyll fluorescence from a mooring in the Thames outer estuary. This shows strong coherence between all observed environmental variables, with the tidal cycle as dominant driver of their variability. In chapter 7 the drivers of fluctuations in chlorophyll

fluorescence are compared between four mooring sites in the North Sea across different time scales. In this chapter wavelet analysis, cross correlations and Generalized Additive Models (GAM) are combined to characterize phytoplankton variability and quantify its predictability from environmental conditions.

Finally, in chapter 8 the results from the different chapters in this thesis are synthesized and recommendations are given for a modernized monitoring strategy of the North Sea and further research.