



A new method of describing phytoplankton blooms: Examples from Helgoland Roads

S. Mieruch ^{a,*}, J.A. Freund ^a, U. Feudel ^a, M. Boersma ^b, S. Janisch ^b, K.H. Wiltshire ^b

^a ICBM, University of Oldenburg, P.O. Box 2503, 26129 Oldenburg, Germany

^b Biologische Anstalt Helgoland, Alfred Wegener Institute for Polar and Marine Research, P.O. Box 180, 27483 Helgoland, Germany

ARTICLE INFO

Article history:

Received 17 July 2008

Received in revised form 5 June 2009

Accepted 12 June 2009

Available online 21 June 2009

Keywords:

Microalgae
diatoms
Phaeocystis
bloom start
Helgoland Roads

ABSTRACT

Phytoplankton blooms, in their pivotal position in pelagic seasonal succession require precise classification criteria in order to evaluate such parameters as bloom start, bloom timing, bloom maximum and growth rates. Such bloom parameters are linked directly to species and bloom specific features. Currently the phytoplankton bloom concept, though intuitively clear, lacks operational criteria allowing the precise definition of bloom parameters. We present a semi-quantitative method of classification of marine phytoplankton blooms based on an algorithmic estimation of several bloom descriptors computed from densely recorded phytoplankton data, like the Helgoland Roads long-term time series. Combining these descriptors we propose a novel classification scheme which may serve useful in the discussion of species fitness, competition and succession of marine algae. Special emphasis is put on the detection of the bloom start, because of its crucial importance for many current research topics, including trigger mechanisms and climate-induced temporal shifts in the context of the match/mismatch hypothesis. Visual examination of scatter plots of these parameters leads us to propose three types of blooming algae.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Phytoplankton are commonly distinguished by their morphology to a species level or, with less specificity combined into taxonomic groups (diatoms, dinoflagellates, cyanobacteria, flagellates etc.). Commonly over a seasonal cycle algal species or groups develop blooms, sometimes also called 'phytoplankton events' (Beliaeff et al., 2001). Abundance and timing of these phytoplankton blooms are pivotal to a proper functioning of the marine food web and climate induced shifts are currently discussed in the context of the match/mismatch hypothesis (Beaugrand et al., 2003; Cushing, 1990; Platt et al., 2003; Wiltshire & Manly, 2004), i.e. the temporal and/or spatial decoupling of hitherto synchronised processes.

When discussing phytoplankton blooms, scientists base their assessments on classical batch culture studies whereby cell count records/biomass concentrations are segmented into different phases. Phytoplankton blooms, spring blooms as well as red tides, are considered to have a lag phase with a few initial cells (Spencer, 1954), an exponential growth phase where cells produced also produce more viable cells, a declining growth rate phase where cessation of the exponential growth occurs, a stationary phase where cell death and growth are balanced and, finally, a death phase where the population declines (Fogg and Thake, 1987; Spencer, 1954) caused by several factors e.g. sinking, grazing, nutrient depletion, viral

infection etc.. Indeed, in situ algae can actually bloom in roughly such a manner as has been shown in the works of Lund (1950) and Platt and Subba Rao (1970). In the context of detecting parameters like the bloom start, we see the development of a phytoplankton bloom as a change in the dynamics of cell counts switching from slow growth (lag phase) to pronounced exponential growth. As a major advantage, the dynamical aspect of this approach is universal and totally different from species specific threshold concepts.

Closer examination of oceanic and neritic algae shows, however, that marine species do not always bloom in such a straightforward manner and that even from year to year single species can show different bloom patterns. Thus, from long term observations we know that the blooms of phytoplankton species on a seasonal basis in marine systems are difficult to describe even in general terms. Only a few attempts of mathematically differentiating, for instance, start, maximum and duration of blooms, have been reported for coastal waters by e.g. Beliaeff et al. (2001) and Rolinski et al. (2007). A description of blooms in marine systems, and especially in shelf seas, is more difficult than in limnic systems since strong hydrodynamic forcing results in a highly fluctuating environment (dynamical noise). Sometimes these fluctuations hamper reproducible or well-controlled measurements (observational noise) and, on top of that, cell counting methods introduce statistical fluctuations (statistical noise). However, describing bloom dynamics is not straightforward, especially in high frequency (e.g. daily) data. Indeed, several different methods have been described (see above), all of which have their merits depending on the shape of the blooms and sampling frequency but in Wiltshire et al. (2008) we found that there is no easy way to describe a bloom data set with highly resolved and variable data.

* Corresponding author. Present address: Institute of Environmental Physics, University of Bremen, Otto-Hahn-Allee 1, 28359 Bremen, Germany.

E-mail address: sebastian.mieruch@iup.physik.uni-bremen.de (S. Mieruch).

The most effective way to rectify this situation is to evaluate data sets on phytoplankton with a high taxonomic resolution and differentiate the blooming characteristics of different microalgae preferably over long periods of time. Luckily, high resolution long-term data exist, for example the Helgoland Roads (HR) time series. This is part of the Helgoland Roads monitoring programme of biological, chemical and physical parameters (e.g. Wiltshire 2004). In this paper we aim to use 16 key species of the HR data base to calculate descriptors, characterising phytoplankton species and blooms of the North Sea habitat at Helgoland. We choose these 16 species, because we have reliable unbiased and quality controlled data from 1962 onwards as described in Wiltshire (2004). Secondly these 16 species are very important and representative food species for the respective region. Throughout this paper we use the term “bloom” for the largest count of cells for each species during each year in the data set as an operational definition knowing that we neglect for algorithmic simplicity the sometimes observable multiple blooms for particular species. A special focus will be on the appropriate extraction of the bloom start as this information is relevant for many subsequent analyses, e.g. the method of bloom triggered averaging (Freund et al., 2006a), and also important in the context of timing shifts and the afore-mentioned match/mismatch hypothesis.

The paper is organised as follows: In Sect. 2 we elaborate on data analysis methods which we developed in order to compute the diverse bloom descriptors. In Sect. 3 we combine the computed quantifiers to yield a classification of blooms and/or algal species observed in the Helgoland pelagic environment. We end with a discussion of our results and their ecological relevance and give an outlook on potential applications and future research.

2. Methods

2.1. Data acquisition

Surface water samples were taken (usually before 9 am) on working days at the “Kabel Tonne” site (54°11.3'N, 7°54.0'E) between the two islands at Helgoland using a bucket. This sample was mixed well and subsampled into a glass bottle for future analyses of nutrients and phytoplankton. The phytoplankton samples taken from this were preserved in a brown glass bottle using Lugol's solution. The preserved sample was counted under an inverted microscope to species level using the Utermöhl method (25–50 mls of volume counted). Based on Poissonian statistics we require a minimum of 120 cells per litre for a phytoplankton bloom. The probability of observing no cell at this concentration amounts to less than 5% or 2.5% for sample volumes 25 ml or 50 ml, respectively.

The species lists used are described in Wiltshire and Dürselen (2004). The data is archived in the PANGAEA databank (URL: <http://www.pangaea.de>).

All of the microalgal data from the time series were quality controlled for counting errors, identification misnomers and occurrence conformity (Wiltshire and Dürselen, 2004). We chose 16 representative algae from this data set. We also include *Noctiluca scintillans* in our study, it being an important bloomer in the North Sea although it is not an autotrophic microalga, but rather a heterotrophic dinoflagellate. The range of species evaluated was chosen to represent the diverse array of organisms. They include organisms typical for all seasons, diatoms, dinoflagellates and heterotroph organisms, with low and high cell counts etc.. The raw data are unevenly sampled (work-daily) records. As can be seen from Fig. 1 cell counts were heavily affected by fluctuations that probably result from the small-scale patchy nature of the algal distribution in surface waters on a daily basis due to changes of environmental conditions (e.g. wind causing turbulence and hydrodynamic flow) or the patchy structure of algal populations. Such fluctuations might be especially important at the bloom start; as an illustration see the third panel of Fig. 1 which

exhibits the rising flank interspersed with dips down to zero counts. Both, data gaps and fluctuations prevented a reliable determination of bloom characteristics, hence, as a first step of data processing we had to interpolate and smooth the raw data.

2.2. Pre-processing: application of smoothing splines

The choice of an appropriate data smoothing method is one of the most decisive elementary steps for subsequent analyses. We chose the method of smoothing splines as it represents the optimal compromise between smoothing without deviating too much from the original data (Reinsch, 1967).

It should be noted that smoothing splines accomplish both tasks, smoothing and interpolation simultaneously. However, in contrast to common interpolating splines, which pass through all the data points, the method of smoothing splines accounts for measurement errors and, consequently, allows deviations of the smoothed curve from data points.

To be specific, let us denote the measurement points by pairs (x_i, y_i) and the smoothed curve by pairs $(x_i, f(x_i))$ both for $i = 1, \dots, N$. The smoothing spline algorithm is based on a standard method of variational calculus by minimising the functional

$$\tilde{S}(f, \lambda) = \int_{x_1}^{x_N} [f''(x)]^2 dx + \lambda \left(\sum_{i=1}^N \frac{1}{\sigma_i^2} [y_i - f(x_i)]^2 - N \right). \quad (1)$$

The first term integrates the square of local curvature; the smoother the curve $f(x)$ the smaller its contribution to the functional. The second term sums the residuals normalised to the local variance σ_i^2 . Given an appropriate choice of σ_i one would expect this sum to equal N . Deviations from this prescribed value N are multiplied by the Lagrange parameter λ (penalty parameter) and this product constitutes the second contribution to the functional. The minimisation procedure yields conditions for $f(x)$, in particular the form of piecewise cubic polynomials (Reinsch, 1967), that can be cast into the form of matrix equations which can be solved by an algebraic computer program. The values of $\lambda = 5 \cdot 10^6$ and $\sigma = 1 \cdot 10^4$ were adapted to the HR data (Mieruch and Freund, 2005).

An example of this treatment is given in Fig. 1 where we depict the measured counts of the species *Guinardia delicatula* (blue line and diamonds) together with the smoothing spline (red curve) and a smoothed (green) curve that was obtained by applying a moving average with a sliding rectangular window of width five days to linearly interpolated data. Both smoothing procedures result in similar traces although the spline is visibly smoother. A key issue with regard to algal bloom description is the correct extraction of the bloom start from measured data. When devising an algorithm for this seemingly simple task, one is confronted with several problems: As can be seen from Fig. 1, a determination of the bloom start by visual inspection heavily depends on whether one plots cell counts on a linear or logarithmic scale. To account for the detection limit we replaced zero counts by the detection limit of 20 cells per litre, which results from the fact, that the counting volume was 50 ml (cf. Sect. 2.1), thus finding one cell per 50 ml corresponds to 20 cells per litre. From now on zero counts should be understood as concentrations below the detection limit.

Selecting the bloom start as the first day of zero counts preceding the maximum from raw data is also problematic (as seen in the bottom panel of Fig. 1). Approaching the annual peak from left to right and to select the terminal day of leading zeroes might be suitable here. However, in other cases we observed unrealistic bloom durations larger than two hundred days. Hence, below we describe an algorithm that searches for the bloom start by moving in the left direction away from the annual maximum. For the example depicted in the second and third panel of Fig. 1 this algorithm, when applied to raw data,

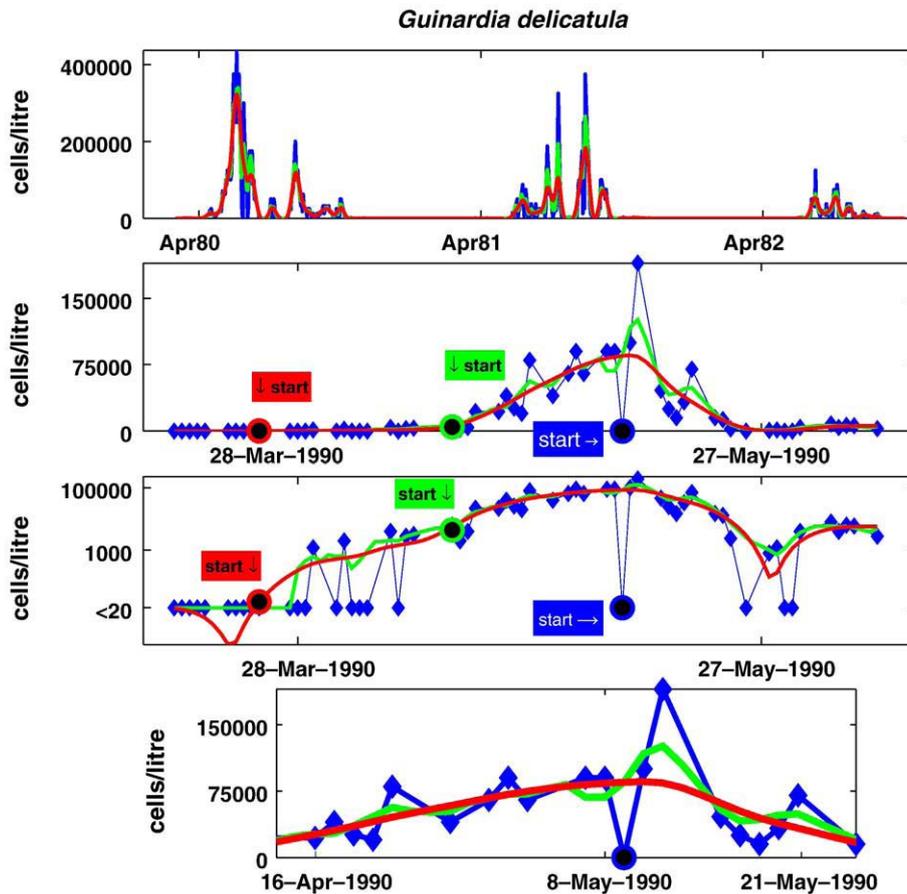


Fig. 1. An example of the Helgoland Roads data: raw data (blue) together with two smoothed curves (red: smoothing spline, green: moving five-day average). A representation of the same data on a linear (second panel) or a logarithmic scale (third panel) illustrates that a visual determination of the bloom start might result in drastic discrepancies. Notice the detection limit of 20 cells per litre in the logarithmic representation. Noise in the raw data causes difficulties in the algorithmic determination of the bloom start: Applying our algorithm (see text and cf. Fig. 2) to raw data would select an obviously erroneous start (blue). Depending on the smoothing method our algorithmic criterion selects either one (red) or the other (green) bloom start. From biological considerations the red starting point corresponding to the smoothing spline should be preferred. The last panel demonstrates the smoothness of the red spline in a blow-up.

would be trapped in a fluctuation (08 May 1990). To prevent this and since the common sense notion of an algal bloom implies continuous data we used smoothed curves.

2.3. Extraction of bloom descriptors from pre-processed data

Following the pre-processing described above we computed several bloom descriptors:

- Annual maximum counts (cells/litre)
- Bloom start (ordinal dates)
- Start counts (cells/litre)
- Bloom specific growth rate at bloom start (1/day)
- Start-to-peak distance (days)

In the following we detail these descriptors:

2.4. Annual maximum counts

Though precisely defined, the idea of the maximum annual peak is sometimes challenged by additional peaks or even plateaus of smaller but comparable height. Due to the high variability of values across years and species we preferred to use decadal logarithms of maximum values in graphical representations. By default, with the term “annual” we mean the 365 day-period starting from the first of January, although in our algorithm (Mieruch and Freund, 2005) this date can be shifted as desired, e.g. for cold water bloomers.

2.5. Bloom start

A naive approach would determine the bloom start as the instant when cell counts cross a species-specific threshold from below. However, even within single species a large variance of annual maximum counts makes such a simplistic threshold mechanism unlikely. We chose the following method to extract the bloom start from pre-processed data:

- First the algorithm defines an interval within which the bloom start will be searched. To this end, the algorithm starts at the bloom peak and moves to the left collecting all days with decreasing cell numbers. A stagnation point or a local minimum will be accepted as the search interval start if and only if related cell counts are below 1% of the annual peak (see also Greve et al., 2004). If cell counts exceed the 1% threshold the algorithm continues moving to the left until it hits another stagnation point or local minimum. Continuing as before, the onset of the search interval will eventually be found.
- Once the search interval has been determined the algorithm computes instantaneous growth rates (local slopes of three-day linear regressions of logarithmic cell counts).
- From the list of instantaneous growth rates we computed their increments. We define the bloom start as the point of maximum increment, i.e. of maximum acceleration of phytoplankton growth, given that both rates are non-negative.

Figure 2 illustrates the detection of the bloom start using the algorithm described above. There we have explicitly indicated the lag

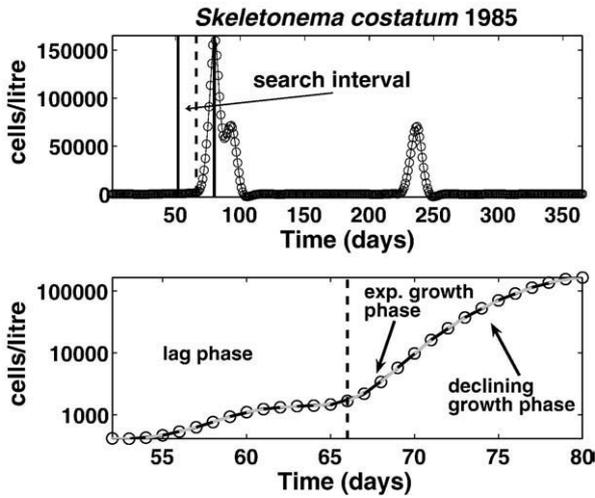


Fig. 2. Our algorithm to determine the bloom start is applied to pre-processed data (via a smoothing spline) shown in the upper panel. From the total range of times (upper panel) a search interval (indicated by the vertical lines) is segmented automatically. Alternating grey and black line segments in the lower panel (search interval) represent instantaneous growth rates (fits over three days). The dashed line indicates the bloom start which separates the lag phase from the exponential growth phase.

phase, the exponential growth phase and the declining growth phase as obtained from the smoothed data.

2.6. Start counts

The start counts are the cell numbers at the bloom start. This is neither a constant value nor a year-specific threshold.

2.7. Growth rate at bloom start

Instantaneous growth rates were already defined and used in Sect. 2. It should be noted that this value does not necessarily coincide with the maximal growth rate, which is experienced at an inflexion point, which needs additionally the second and third derivatives

($f' > 0, f'' = 0$ and $f''' < 0$) of the concentration profile ($f(t)$). The units of the growth rates at bloom start are given in 1/day.

2.8. Start-to-peak distance

This is the temporal distance (in days) between the bloom start (as defined in Sect. 2.3.2) and the day of annual maximum counts.

For the interested user we refer to a website (Mieruch and Freund, 2005) where own data can be analysed. This page offers a suite of our Matlab scripts which can be readily applied to existing data series of algal cell counts and which will compute the bloom descriptors.

3. Results

Following the computation of bloom descriptors we now proceed to the second stage of our semi-quantitative classification. We combine the above calculated bloom descriptors in a series of scatter plots and classify observed structures in two different ways: species-based and event-based.

3.1. Annual maximum counts vs. growth rate at bloom start

First of all we investigated the relationship between the initial growth rate at bloom start and the maximum density reached by the selected 16 key species (Fig. 3). One could reason that this relationship is a trivial one, as it could be expected that algae which start growing at a higher speed should reach higher densities than those with slower growth rates. However, this is not necessarily the case for three reasons: Firstly, the bloom start can be temporally separated from the time of maximum cell counts by as many as 30 days, time for random fluctuations to erase the correlation between initial growth rate at bloom start and maximum cell counts. Secondly, the time needed to reach the maximal cell count in the different blooms is widely variable, and could also influence the height of the maximum densities. Indeed the correlation between annual maximum counts vs. start-to-peak duration was very small (see below). Thirdly, the maximum occurs, when growth and mortality are balanced, thus loss terms are relevant before the maximum is reached, hence making the correlation between initial growth rate and the maximum densities a

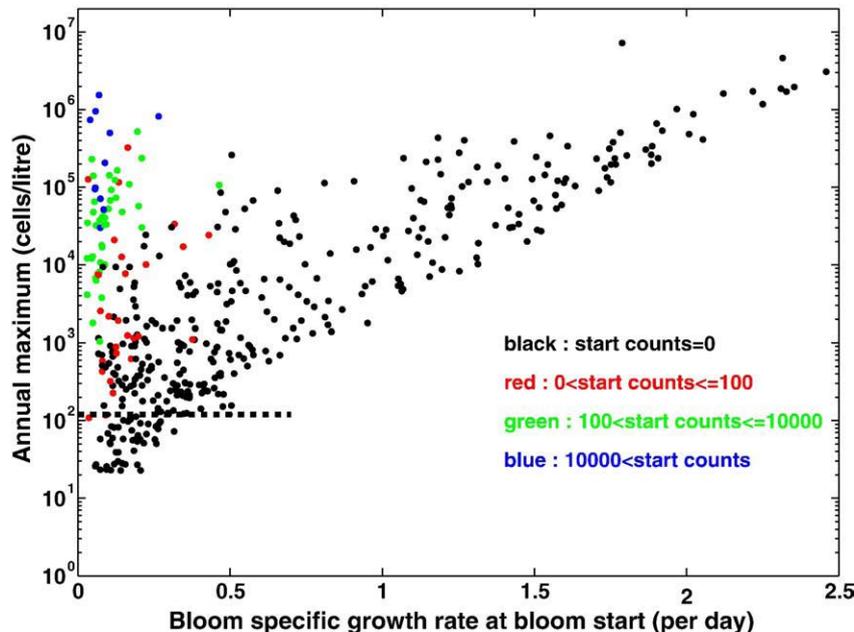


Fig. 3. Scatter plot of annual maximum counts vs. growth rate at bloom start with colours coding start counts. The values of the abscissa indicate the multiplication per day, i.e. one means 10^1 , two means 10^2 etc.. This plot clearly shows the event-based classification of Zero Starters (black) and Supra Zero Starters (red, green, blue). The black dashed line separates the statistical fluctuations from blooms.

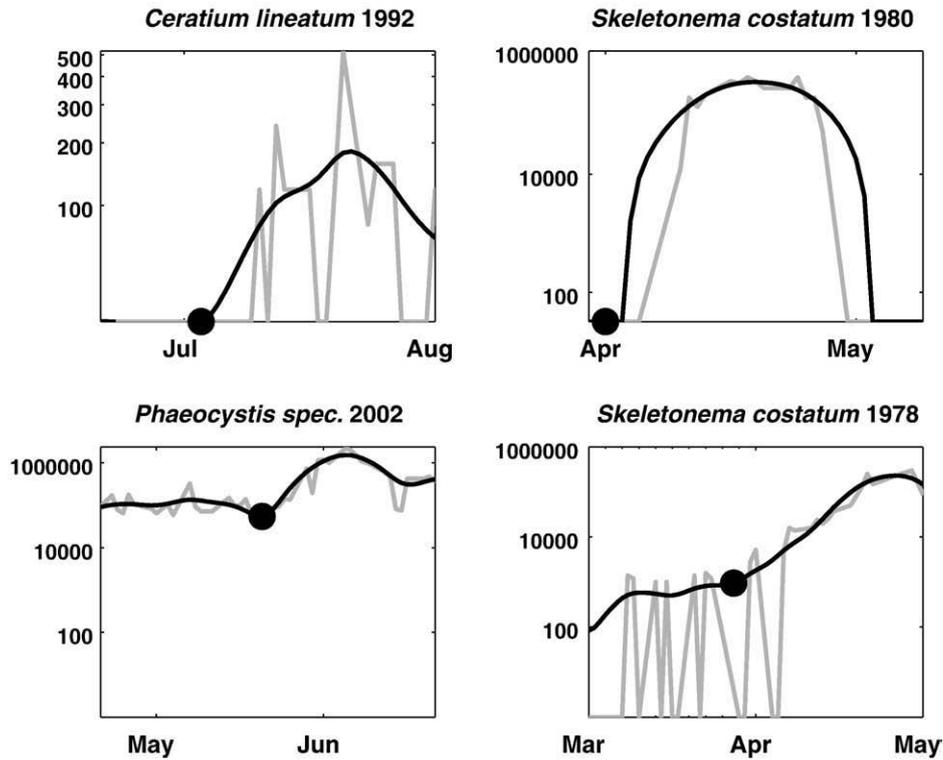


Fig. 4. Two examples of Zero Starters (top row), and two examples of Supra Zero Starters (bottom row). According to our classification scheme the upper left panel shows a type A bloomer (*Ceratium lineatum*), the lower left a type C bloomer (*Phaeocystis spec.*), and the two right panels relate to a type B bloomer (*Skeletonema costatum*). Cell concentrations are plotted on a logarithmic axis (note that Zero Starters are counted with value 20 (detection limit)). Black dots indicate the bloom start as identified by our algorithm applied to pre-processed data. Grey curves outline raw data, black curves represent the smoothing splines.

non-trivial one. The event-based classification (Fig. 3) of phytoplankton blooms results in finding two types of bloomers:

1. **Zero Starters** are marked as black dots (a highly significant correlation coefficient ($r = 0.86, p < 0.001$) is observed between the initial growth rate and the maximum counts), and these type of bloomers start growing from very low initial densities (examples: top panels of Fig. 4).

2. **Supra Zero Starters** (red, green, blue dots) start with varying start counts up to 100000 cells per litre and more (examples: bottom panels of Fig. 4).

The species-based classification is adapted from Fig. 5 where different species are indicated by different colours and symbols. Based on this structure of scattered blooms, we introduced a grouping of different types of bloomers indicated by ellipsoids shown schematically

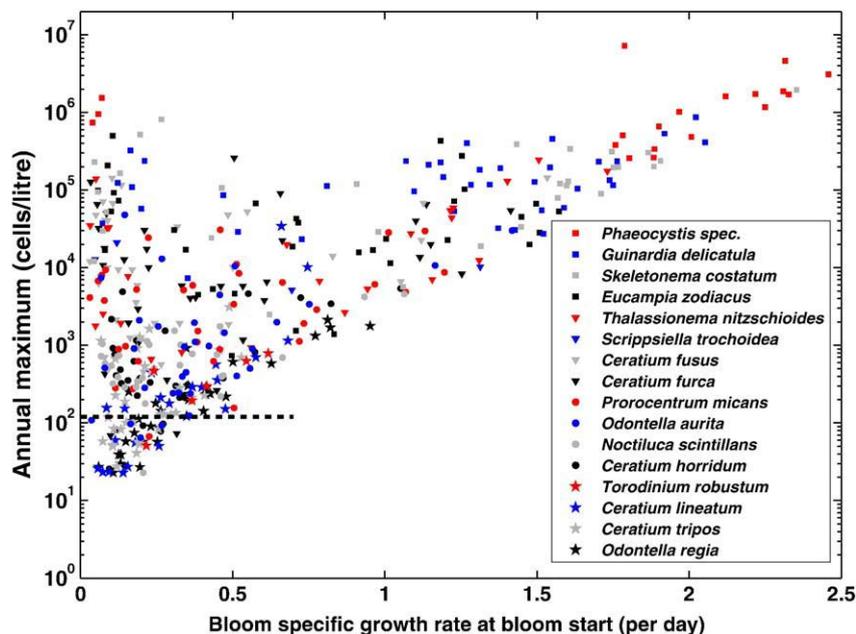


Fig. 5. Scatter plot of annual maximum counts vs. growth rate at bloom start with colours coding key species; note the decadal logarithmic scale of the ordinate. The basic features reflecting the statistical structure of this bloom ensemble are sketched in Fig. 6; further explanations see text. The black dashed line separates the statistical fluctuations from blooms.

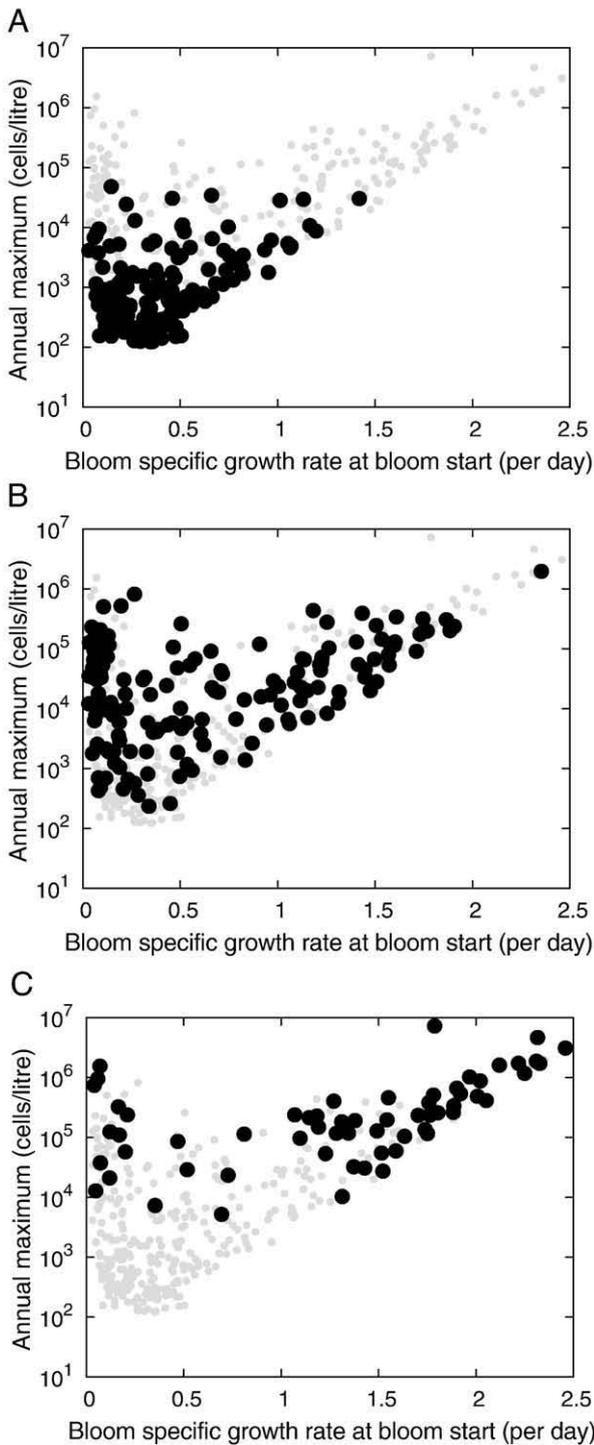


Fig. 6. The three bloom types A: Pure Zero Starters, B: Mixed Zero and Supra Zero Starters and C: Constant Peak Bloomers, obtained by grouping the computed bloom descriptors are plotted as black dots.

in Fig. 6. So from two bloom types (**Zero Starters**, **Supra Zero Starters**), ignoring the species we go to three:

- **Pure Zero Starters:** The faster they start the higher they get. Even though the correlation between the maximum count and start-to-peak duration was significant, the coefficient of determination was not very strong and the slope of the relation not very steep. Hence, it seems to be the case that the growth rate at the start of a bloom is a good indication of the growth during the whole bloom in these type blooms. For instance *Odontella regia* is a type A bloomer.

- **Mixed Zero and Supra Zero Starters:** These are species that have a mixed strategy: They can start growing with relatively small growth rates at bloom start and reach a broad range of different maximum densities. This dispersion is strongly correlated with substantially varying start counts and not with variation in start-to-peak durations. Furthermore type B can grow like type A. As an example we mention the species *Ceratium fusus* in Fig. 5.
- **Constant Peak Bloomers:** Irrespective of substantially varying growth rates at bloom start they reach narrow-distributed (on a logarithmic scale) bloom peak counts, a fact which is understandable when noting that both start counts and start-to-peak duration can vary substantially. A typical representative is *Guinardia delicatula* (blue squares in Fig. 5). The crucial difference between B and C is the stable character of C mostly reaching constant maximum densities.

3.2. Annual maximum counts vs. start-to-peak duration

When constructing scatter plots (not shown here), we only found a weak correlation ($r=0.10, p=0.035$) between the start-to-peak duration and the annual maximum cell counts, thus emphasising the weak influence of the start-to-peak duration on maximum cell count.

3.3. Start-to-peak duration vs. growth rate at bloom start

Plotting the start-to-peak duration vs. the growth rate at bloom start (not shown here) reveals that there is only a weak but statistically significant ($p<0.001$) anti-correlation (correlation coefficient $r=-0.17$) which reflects the fact that due to nutrient depletion algal species cannot grow too long with a large growth rate. Similar effects can be expected from temperature dependence of growth rates.

3.4. Statistical support for our classification

Based on our classification above we assigned each of the 16 species to one of these types (Table 1). Additionally we calculated species specific mean (μ_{cc} and μ_{gr}) and standard deviations (σ_{cc} and σ_{gr}) of logarithmic maximum cell counts and growth rates, respectively. We tested this classification, using all of the available parameters described above in a discriminant analysis with the

Table 1

Assignment of bloom types A, B and C to each of the 16 key species from Helgoland Roads data.

Species	Bloom types	$\mu_{cc} \pm \sigma_{cc}$	$\mu_{gr} \pm \sigma_{gr}$	# blooms
<i>Phaeocystis spec</i>	C	6.0 ± 0.4	1.7 ± 0.8	19
<i>Guinardia delicatula</i>	C	5.1 ± 0.4	1.2 ± 0.6	38
<i>Skeletonema costatum</i>	B	5.0 ± 0.6	1.1 ± 0.7	30
<i>Eucampia zodiacus</i>	B	4.5 ± 0.7	0.8 ± 0.5	33
<i>Thalassionema nitzschioides</i>	B	4.1 ± 0.8	0.7 ± 0.6	24
<i>Scrippsiella trochoidea</i>	(C)	4.0 ± 0.3	0.5 ± 0.6	4
<i>Ceratium fusus</i>	B	4.0 ± 0.8	0.3 ± 0.4	34
<i>Ceratium furca</i>	B	3.9 ± 0.8	0.5 ± 0.4	34
<i>Prorocentrum micans</i>	A	3.5 ± 0.6	0.5 ± 0.3	32
<i>Odontella aurita</i>	A	3.2 ± 0.7	0.4 ± 0.3	28
<i>Noctiluca scintillans</i>	A	2.8 ± 0.3	0.3 ± 0.2	34
<i>Ceratium horridum</i>	A	2.9 ± 0.5	0.3 ± 0.2	26
<i>Torodinium robustum</i>	(A)	2.6 ± 0.2	0.4 ± 0.1	5
<i>Ceratium lineatum</i>	A	2.7 ± 0.7	0.4 ± 0.2	14
<i>Ceratium tripos</i>	A	2.6 ± 0.4	0.3 ± 0.1	18
<i>Odontella regia</i>	A	2.6 ± 0.4	0.5 ± 0.2	19

This assignment (cf. Fig. 5) is based on our semi-quantitative classification and supported by the combination of standard deviations of logarithmic maximum cell counts σ_{cc} and growth rate at start σ_{gr} . Related mean values are denoted μ_{cc} and μ_{gr} . Bracketed bloom types indicate an assignment that we consider to be statistically non significant as the number of observed blooms (last column) was too small (<10).

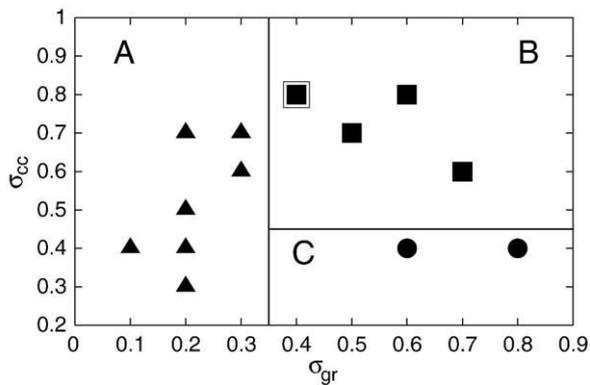


Fig. 7. Our assignment of bloom types listed in Table 1 is reflected by clusters in the plane spanned by standard deviations σ_{gr} and σ_{cc} of growth rates and logarithmic maximum cell counts, respectively. Statistically non significant classifications of algae are neglected (cf. Table 1). *C. furca* and *C. fusus* have identical σ_{gr} and σ_{cc} values and are indicated by the open square overlaid on the filled square.

three type bloomers (A, B, C) as the grouping variables, and the five bloom descriptors as the measured variables. Based on this, we observed a highly significant difference between groups (Wilks Lambda = 0.035; $F_{8,872} = 78$; $p < 0.001$), with maximum cell counts, initial growth rate and start date of the bloom as those variables that significantly differentiate the groups, resulting in a correct classification of the blooms in 73% of all cases (392). Most mistakes were made between C and B, which seems understandable given the overlap as shown in Fig. 6. Plotting the standard deviations of the logarithmic maximum cell counts (σ_{cc}) and growth rates (σ_{gr}), yields the same picture (Fig. 7). Bloom type A is typically connected with $\sigma_{gr} < 0.35$, while larger σ_{gr} comprise types B and C. The latter-mentioned types can be separated by the threshold of $\sigma_{cc} = 0.45$; while type C typically falls below this value type B is found at larger values.

So, we have shown that it is possible to differentiate between different bloomers based on the characteristics of the bloom, and we can do so with about 73% accuracy.

4. Discussion

In this paper we applied a simple method to differentiate algal growth types in blooming algal species. To date, such an examination of algal growth is missing in the literature and most studies deal with blooms in a fairly descriptive way. The only rate evaluations are based on temperature, light and nutrient related differences (Keller, 1989; Sommer, 1989; Wright, 1964). Particularly, precise and generally accepted criteria defining algal blooms in situ are missing. This is due to the fact that most of our current knowledge on microalgal growth has been obtained from experiments under ideal conditions (Fogg and Thake, 1987; Spencer, 1954). At the other end of the spectrum are models describing unicellular growth (Albinet & Pelce, 1996; Hentschel and Fine, 1994) and process-oriented models (Edwards and Brindley, 1996; Huppert et al., 2002; Truscott and Brindley, 1994). However, algal growth under natural conditions never occurs under ideal conditions, because fluctuations e.g. in temperature, light, nutrients etc. take effect.

Using our algorithmic methods we were successfully able to extract the following descriptors in over 80% of the 392 major blooms that occurred in the 16 key species in the period 1962–2002: maximum abundance, bloom start, start abundance, growth rate at start and start-to-peak duration. The algorithmic criterion presented in this paper is motivated by experimental observations and defines the bloom start as the point of maximal acceleration of cell proliferation. To avoid being trapped in fluctuations it must be applied to pre-processed data and we found that (adapted) smoothing splines

are appropriate to render the bloom start of realistic data in agreement with biological expertise.

Using these descriptors we differentiated between different types of blooms. Bloomers of the type C which reach a prototypical abundance in each year seem to be less sensitive to changes in their environment and to shifts of the bloom timing. Their peak abundance is often close to the saturation limit (potential carrying capacity) and not so much affected by competing species. In contrast, bloomers of the types A or B can show a large variation in their peak abundance that is either rooted in largely varying growth rates at the start (type A) or in their abundance right at the start (type B). Both factors indicate that they are much more susceptible to a changing environment, competitors or zooplankton grazing pressure. The distinction between Zero Starters and Supra Zero Starters (event-based) links the single bloom to the history preceding the bloom (cf. Fig. 4). Those species which are identified as type A always start from scratch and reach a peak abundance determined by the growth rate at the bloom start. This is remarkable when seen together with the fact that the start-to-peak duration is weakly correlated with the maximum cell counts. These empirical findings can be explained by the observations of a tight coupling between growth rate and initial nutrient concentration as reported by (Kudela and Dugdale, 2000).

The significance of this work is, of course, based upon the ecological relevance of these observations. For example we can evaluate two microalgae which overlap in their timing: *Guinardia delicatula* and *Odontella aurita*. These are characterised as C and A type bloomers respectively. From these descriptions (see also Table 1) one would assume that *Guinardia delicatula* is the more robust species of the two, not susceptible to negative changes in its environment. In our evaluations of the co-occurrence of species it would seem that *Guinardia* is becoming the more abundant algae, with its period of occurrence becoming seasonally wider relative to the other algae (Freund et al., 2006b). This could be due to its robustness as such and, in addition, as *Guinardia* is a warm water loving species this is to its advantage with regard to the evinced warming at Helgoland (Wiltshire and Manly, 2004). When one examines algae of the same genus e.g., *Ceratium*, it is clear that *C. fusus* and *furca* are similarly classified, i.e., B types. This seems rather logical as they also often overlap in their occurrence and even seem to alternate. *C. lineatum*, *C. horridum* and *tripos* are classified as A bloomers, they occur earlier in the year and also together. *Odontella regia* and *Odontella aurita* also co-occur, they are both A types. If one classifies the μ_{cc} values and the μ_{gr} as some sort of success index then the order is C > B > A. Using the new criteria it should be checked whether the behaviour of the investigated algae has changed over the years related to temperature shifts and regime shifts of the North Sea. This would include timing of the blooms, changes in succession and growth rates.

We propose that the above described results should be compared with existing in situ data for isolated blooms of the same and other oceanic algae as well as data from laboratory or mesocosm experiments. Possible discrepancies should be investigated as they could be attributed to the strongly fluctuating environment and the effects it has on the competition and succession of species.

Our algorithms and analysis methods are implemented in Matlab scripts that can be downloaded (Mieruch and Freund, 2005) and readily applied to other (densely recorded) phytoplankton data or short term data recorded in mesocosm experiments. In the case of less dense sampled data the description of blooms using parametric methods, as applied in Rolinski et al. (2007) could enable the transfer of our classification scheme. We hope that our proposed algorithmic determination of the bloom start and concurrent examples will open a discussion among aquatic scientists about the most appropriate definition of the bloom start and, in addition, stimulate further research devoted to bloom trigger mechanisms.

The achieved classification may be characteristic for successional blooms or a specific habitat. In order to investigate the latter point

comparisons with other marine habitats are desirable. The developed classification scheme may thus be suitable for a monitoring programme in the context of climate-induced changes in a marine ecosystem. In addition, it allows to consider the sensitivity of single algal species to their natural environment or to the spectrum of their competitors.

Acknowledgements

S. Mieruch acknowledges support by the Wolfgang-Schulenberg Programme. J.A.F. and K.H.W. acknowledge financial support by the DFG in the framework of a DFG priority programme (AQUASHIFT 1162). We acknowledge fruitful discussions with C.D. Dürselen, M. Dowd, K. Wirtz and N. Grüner. We also acknowledge the Crew of the RV AADE of the BAH-AWI as well as the past counters of the Helgoland Roads Data series, particularly H. Treutner.

References

- Albinet, G., Pelce, P., 1996. Computer simulation of neurite outgrowth. *Europhysics Letters* 33, 569–574.
- Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S., Reid, P.C., 2003. Plankton effect on cod recruitment in the North Sea. *Nature* 426, 661–664.
- Beliaeff, B., Gros, P., Belin, C., Raffin, B., Gailhard, I., Durbec, J.P., 2001. 'Phytoplankton events' in French coastal waters during 1987–1997. *Oceanologica Acta* 24, 425–433.
- Cushing, D.H., 1990. Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in Marine Biology* 26, 249–292.
- Edwards, A.M., Brindley, J., 1996. Oscillatory behaviour in a three-component plankton population model. *Dynamics and Stability of Systems* 11, 347–370.
- Fogg, G.E., Thake, B., 1987. *Algal cultures and phytoplankton ecology*. University of Wisconsin Press, Wisconsin.
- Freund, J.A., Mieruch, S., Scholze, B., Wiltshire, K., Feudel, U., 2006a. Bloom dynamics in a seasonally forced phytoplankton-zooplankton model: Trigger mechanisms and timing effects. *Ecological Complexity* 3, 129–139.
- Freund, J.A., Pöschel, T., Wiltshire, K.H., 2006b. Markovsche Analyse nasser Gemeinschaften, pp. 99–110. Logos, Berlin.
- Greve, W., Reimers, F., Nast, J., Hoffmann, S., 2004. Helgoland Roads meso- and macrozooplankton time-series 1974 to 2004: lessons from 30 years of single spot, high frequency sampling at the only off-shore island of the North Sea. In: Wiltshire, K.H. (Ed.), *Helgoland Marine Research*, Vol. 58. Springer, Berlin, pp. 274–288.
- Hentschel, H.G.E., Fine, A., 1994. Instabilities in cellular dendritic morphogenesis. *Physical Review Letters* 73, 3592–3595.
- Huppert, A., Blasius, B., Stone, L., 2002. A model of phytoplankton blooms. *The American Naturalist* 159, 156–171.
- Keller, A.A., 1989. Modeling the effects of temperature, light, and nutrients on primary productivity: An empirical and a mechanistic approach compared. *Limnology and Oceanography* 34, 82–95.
- Kudela, R.M., Dugdale, R.C., 2000. Nutrient regulation of phytoplankton productivity in Monterey Bay, California. *Deep-Sea Research II* 47, 1023–1053.
- Lund, J.W.G., 1950. Studies on *Asterionella formosa* Hass. II. Nutrient depletion and spring maximum. *Journal of Ecology* 38, 1–35.
- Mieruch, S., Freund, J. A., 2005. Quick Start Guide. <http://www.icbm.de/~material>.
- Platt, T., Fuentes-Yaco, C., Frank, K.T., 2003. Spring algal bloom and larval fish survival. *Nature* 423, 398–399.
- Platt, T., Subba Rao, D.V., 1970. Energy flow and species diversity in a marine phytoplankton bloom. *Nature* 227, 1059–1060.
- Reinsch, C.H., 1967. Smoothing by spline functions. *Numerical Mathematics* 10, 177–183.
- Rolinski, S., Horn, H., Petzoldt, T., Paul, L., 2007. Identifying cardinal dates in phytoplankton time series to enable the analysis of long-term trends. *Oecologia* 153, 997–1008.
- Sommer, U., 1989. Maximal growth rates of antarctic phytoplankton: Only weak dependence on cell size. *Limnology and Oceanography* 34, 1109–1112.
- Spencer, C.P., 1954. Studies on the culture of a marine diatom. *Journal of the Marine Biological Association of the United Kingdom* 33, 265–290.
- Truscott, J.E., Brindley, J., 1994. Ocean plankton populations as excitable media. *Bulletin of Mathematical Biology* 56, 981–998.
- Wiltshire, K.H., 2004. Editorial. In: Wiltshire, K.H. (Ed.), *Helgoland Marine Research*, Vol. 58. Springer, Berlin, pp. 221–222.
- Wiltshire, K.H., Dürselen, C.-D., 2004. Revision and quality analyses of the Helgoland Reede long-term phytoplankton data archive. In: Wiltshire, K.H. (Ed.), *Helgoland Marine Research*, Vol. 58. Springer, Berlin, pp. 252–268.
- Wiltshire, K.H., Manly, B.F.J., 2004. The warming trend at Helgoland Roads, North Sea: phytoplankton response. In: Wiltshire, K.H. (Ed.), *Helgoland Marine Research*, Vol. 58. Springer, Berlin, pp. 269–273.
- Wiltshire, K.H., Malzahn, A.M., Greve, W., Wirtz, K., Janisch, S., Mangelsdorf, P., et al., 2008. Resilience of North Sea phytoplankton spring blooms dynamics: an analysis of long term data at Helgoland Roads. *Limnology and Oceanography* 53, 1294–1302.
- Wright, R.T., 1964. Dynamics of a phytoplankton community in an ice-covered lake. *Limnology and Oceanography* 9, 163–178.