Phytoplankton Blooms in Estuarine and Coastal Waters:

Seasonal Patterns and Key Species

Jacob Carstensen[1]*, Riina Klais[2], James E. Cloern[3]

[1]Dept. of Bioscience, Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde, Denmark

[2]Institute of Ecology and Earth Sciences, Tartu University, Lai 40, 51005 Tartu, Estonia

[3]United States Geological Survey, MS496, 345 Middlefield Road, Menlo Park, CA 94025, USA

*Corresponding author: jac@bios.au.dk

Abstract

Phytoplankton blooms are dynamic phenomena of great importance to the functioning of estuarine and coastal ecosystems. We analysed a unique (large) collection of phytoplankton monitoring data covering 86 coastal sites distributed over eight regions in North America and Europe, with the aim of investigating common patterns in the seasonal timing and species composition of the blooms. The spring bloom was the most common seasonal pattern across all regions, typically occurring early (February-March) at lower latitudes and later (April-May) at higher latitudes. Bloom frequency, defined as the probability of unusually high biomass, ranged from 5 to 35% between sites and followed no consistent patterns across gradients of latitude, temperature, salinity, water depth, stratification, tidal amplitude or nutrient concentrations. Blooms were mostly dominated by a single species, typically diatoms (58% of the blooms) and dinoflagellates (19%). Diatom-dominated spring blooms were a common feature in most systems, although dinoflagellate spring blooms were also observed in the Baltic Sea. Blooms dominated by chlorophytes and cyanobacteria were only common in low salinity waters and occurred mostly at higher temperatures. Key bloom species across the eight regions included the diatoms *Cerataulina pelagica* and *Dactyliosolen fragilissimus* and dinoflagellates *Heterocapsa triquetra* and *Prorocentrum cordatum*. Other frequent bloom-forming taxa were diatom genera *Chaetoceros*, *Coscinodiscus*, *Skeletonema*, and *Thalassiosira*. Our meta-analysis shows that these 86 estuarine-coastal sites function as diatom-producing systems, the timing of that production varies widely, and that bloom frequency is not associated with environmental factors measured in monitoring programs. We end with a perspective on the limitations of conclusions derived from meta-analyses of phytoplankton time series, and the grand challenges remaining to understand the wide range of bloom patterns and processes that select species as bloom dominants in coastal waters.

INTRODUCTION

Phytoplankton biomass, primary production and community composition are all highly dynamic at the land-sea interface where diverse human actions and climate variability intersect to drive complex patterns of change over time (Cloern and Jassby, 2008). An important pattern is the occurrence of seasonal or episodic bursts of biomass accumulation as blooms, and research in recent decades has identified processes that trigger blooms at the land-sea interface, including: pulsed inputs of nutrients from river inflow (Peierls et al., 2012; Hall et al., 2013), coastal upwelling (Brown and Ozretich, 2009), atmospheric deposition (Paerl, 1997), wind-induced entrainment of bottom water (Iverson et al., 1974; Carstensen et al., 2005), and neap-spring variability of tidal mixing and stratification (Cloern, 1996); seasonal winds that enhance water retention in bays (Yin, 2003); heat waves that set up thermal stratification (Cloern et al., 2005); increasing retention time in flushed systems (Odebrecht et al., this issue); release of benthic grazing pressure (Carstensen et al., 2007; Cloern et al., 2007; Petersen et al., 2008); and seasonal changes in temperature and solar radiation (Shikata et al., 2008). Phytoplankton blooms have ecological and biogeochemical significance because much of the annual primary production in estuarine-coastal ecosystems occurs during these events when photosynthesis exceeds system respiration (Caffrey et al., 1998). That production is the energy supply that fuels production in food webs supporting fisheries (Houde and Rutherford, 1993), aquaculture harvest (Bacher et al., 1998), system respiration (Hopkinson et al., 2005), and microbial processes that make estuaries biogeochemical hot spots (Cloern et al., 2014). Recent comparisons of chlorophyll *a* time series across a range of estuarine-coastal ecosystem types reveal a surprising diversity of seasonal biomass patterns, that these patterns differ from those in the open ocean (Cloern and Jassby, 2010), and they can change abruptly (Winder and Cloern, 2010).

Comparison of chlorophyll *a* time series across sites has been useful for discovering how the patterns of phytoplankton biomass variability are shaped by features that distinguish estuarine-coastal ecosystems from the open ocean – nutrient enrichment, tidal mixing, freshwater inflow, shallow depth and tight benthic-pelagic coupling, sharp vertical and horizontal gradients (Cloern 1996). Progress has been slower in solving the much more challenging problem of understanding how these and other processes select those phytoplankton species that grow fast enough to

develop blooms. Our general conceptual understanding recognizes one seasonal pattern that starts with a spring bloom dominated by large, fast-growing diatoms, followed by a number of summer blooms comprised of diatoms, flagellates, and dinoflagellates, and autumn blooms dominated by diatoms and dinoflagellates (Tett et al., 1986; Mallin et al., 1991). However, there are many deviations from this classical pattern. Blooms in San Francisco Bay are dominated by diatoms throughout the year (Cloern and Dufford, 2005), and blooms of nitrogen-fixing cyanobacteria typically develop in low salinity waters during summer (Jurgensone et al., 2011) when dissolved inorganic nitrogen is depleted from the surface layer and temperatures are high (Paerl and Huisman, 2009). Dinoflagellates dominate spring blooms in parts of the Baltic Sea (Klais et al., 2011). The highly variable physical environment and nutrient regime in estuaries and coastal waters promote different strategies at different times (Margalef, 1978), and bloom species are often selected among those present at suitable inoculum levels prior to the bloom (Smayda and Reynolds, 2001). Given the complexity of the problem, we have not yet identified consistent seasonal patterns of bloom occurrence by individual species or species groups in coastal waters.

Understanding bloom dynamics at the species level has been elusive partly because we are not making sufficient effort to study life-cycle processes such as sexual reproduction (Sarno et al., 2010), germination of resting stages (Shikata et al., 2008), allelopathy and mutualism between species (Smayda, 1997; Smayda and Reynolds, 2001). Second, the information contained in the many empirical records of phytoplankton community variability has not been synthesized to search for common patterns of bloom occurrence and composition. We extend the approach of comparing time series across sites to explore patterns of variability in phytoplankton communities and, particularly, species that develop blooms in estuaries, bays, and shallow coastal waters. To do this we assembled phytoplankton time series from 86 estuarine and coastal sites, and then probed this compilation to explore four basic ecological questions:

Q1. Which species and higher taxa dominate phytoplankton blooms in shallow, nutrient-enriched coastal waters?

Q2. Are there characteristic seasonal patterns of bloom occurrence?

Q3. Are blooms dominated by a common set of phytoplankton groups or species?

Q4: Does bloom frequency vary consistently along gradients of habitat attributes such as salinity, temperature, light availability, nutrients, or mixing?

Answers to these fundamental questions are essential for expanding our still-limited knowledge of the natural history of phytoplankton species succession and blooms.

METHODS

We used long-term monitoring data from a diverse set of marine ecosystems in North America and Northwestern Europe to identify blooms as observations of unusually high phytoplankton biomass. Differences in the frequencies and phytoplankton taxonomic composition of these blooms were examined across 86 coastal sites ranging from estuaries and lagoons typically affected by land runoff to embayments and nearshore coastal systems (Fig. 1). For simplicity we refer to these as estuarine-coastal sites, recognising their differences in landscape and hydroclimatic settings. These sites encompass a broad range of salinity, temperature, nutrient concentrations, tidal mixing, stratification patterns, water depth and transparency, providing a unique opportunity to explore phytoplankton bloom patterns across habitat gradients characteristic of the land-sea continuum. For some analyses we grouped phytoplankton data from the 86 sites into 8 geographic regions (Fig. 1), largely based on latitude, salinity, tidal amplitude and stratification patterns.

Data sources

Time series (minimum 5 years) of phytoplankton species counts and water quality data (all surface data) were collected from different national and regional monitoring programs (Table 1). In addition to salinity, temperature and Secchi depth, water samples were analysed for nutrient and chlorophyll *a* concentrations using standard measurements within the different monitoring programs. The taxonomical composition and biomass of the phytoplankton community was assumed to be analysed by standard techniques (inverted microscope; Utermöhl, 1958) in Lugol's-fixed samples. The taxonomic resolution varied among and even within monitoring programs due to differences in identification expertise of the microscopist and the level of taxonomic aggregation (e.g. specimens identified to genus level only). We assume that the most common bloom-forming species are well recognised throughout the diverse data sets and that differences

in taxonomic resolution are most problematic for the less common species that are not addressed in this study. The taxonomy used in all data sets was standardized according to the World Register of Marine Species (http://www.marinespecies.org/) to enable comparison of bloom species across sites (Olli et al., this issue). We recognise limitations to the taxonomy obtained by microscopy (Jakobsen et al., this issue) and that the taxonomical identification in some cases includes cryptic species that include multiple species, which we address below.

In all monitoring programs phytoplankton specimens were identified to a standard taxonomical level (mostly at the species or genus level) and size class. Results of microscopic analyses were reported as either biovolume (NRE and SFB; Table 1) or carbon biomass (all other data sets) of each species using different compendia for translating counts (details were not provided with the data). If carbon biomass was not reported we estimated it for each species using measured biovolumes and conversion factors for diatoms (0.11 pg C μm^{-3}; Strathmann, 1967) and non-diatoms (0.13 pg C μm^{-3}; Edler, 1979). More accurate scaling equations could not be employed because cell volumes were not reported. For each sample, we calculated the total phytoplankton carbon biomass by aggregating biomass of all autotrophic and mixotrophic species. We excluded the mixotrophic ciliate *Mesodinium rubrum* because it was not consistently identified in all monitoring programs, and we also excluded from analysis all samples having fewer than five taxonomical units reported.

Bloom definition

The time series of total phytoplankton carbon biomass from each coastal site was separated into bloom and non-bloom observations, using a modified algorithm of Carstensen et al. (2007). In the present study we employed a periodic spline function (gam(bs="cp") in R-package mgcv) as opposed to a harmonic function in Carstensen et al. (2007) to describe the seasonal variation in total carbon biomass. The periodic spline allowed for greater flexibility in describing the seasonal variation. The algorithm for identifying blooms is illustrated in supplemental information Fig. S1.

The algorithm rests on the assumption that the total carbon biomass of non-bloom observations was normally distributed with a seasonal mean described by a periodic spline function. Blooms were identified as significant deviations above this pattern. The algorithm was initialised by setting all observations to the non-bloom population, and the periodic spline for the seasonal variation

was estimated. Occurrences of total phytoplankton carbon biomass exceeding the 99th percentile of the prediction interval of the periodic spline were defined as blooms. The periodic spline was estimated again on the remaining non-bloom observations and the 99th percentile of the prediction interval was used to single out additional bloom observations. This re-estimation of the periodic spline was continued until all non-bloom observations were below the 99th percentile of the prediction interval, and all bloom observations exceeded this upper limit.

For all phytoplankton samples, including those collected during bloom and non-bloom situations, the species (or taxonomical level reported) contributing the most to total carbon biomass was identified and its proportion of the total biomass was estimated. Even though the taxonomic units employed could be at the genus level or higher, we refer to the taxonomical unit with the highest carbon biomass as the "dominating species".

Statistical analyses to answer four fundamental questions

We examined seasonal patterns of bloom frequency across the eight regions by calculating the monthly probabilities of a phytoplankton sample being identified as a bloom. The monthly probabilities of bloom occurrence at each site were calculated using a generalised linear model (GLM) with bloom versus non-bloom as a binomial response variable and sampling month as an explanatory factor. Similarly, monthly means of the biomass proportion of the dominating species were calculated from both bloom and non-bloom samples to examine how much a dominant species contributed to the overall biomass in both cases. We described the region-specific seasonality in bloom frequency by averaging these monthly probabilities over all sites within regions to address Q2 (seasonal patterns).

To address Q3 (common bloom composition) we identified the taxonomical group of bloom-dominating species across the ranges of salinity, temperature and biomass proportion of the bloom-dominating species represented by the 86 sites. We also identified the most common bloom-dominating species (those found in >1 % of all bloom samples) for each region and across the entire data set, and we compared the frequencies that these species dominated biomass in bloom and non-bloom samples. We further analysed occurrences of dominant bloom species commonly found across all sites (only those identified to the species level) in relation to salinity and temperature.

We investigated differences in bloom frequency across sites by calculating site-specific mean bloom frequencies as the average of the site-specific monthly bloom frequencies. We identified differences in the site-specific bloom frequencies between regions using a one-way ANOVA. For comparison, we calculated site-specific annual means of salinity, temperature, total nitrogen (TN) and total phosphorus (TP) using a GLM approach accounting for differences in sampling frequencies between years and months (see Carstensen et al. 2006 for details). Time series of water quality data were generally longer than the phytoplankton time series, but we only analysed water quality data from years with phytoplankton samples. TN and TP were log-transformed prior to applying the GLM and the site means resulting from the GLM were subsequently back-transformed to their original scale. The frequency of stratification was calculated for each site as the relative number of samples having a density difference between bottom and surface samples larger than 0.5 g L^{-1}. Individual sites were categorized as: 'stratified' if the frequency of stratification by this definition was higher than 80 %; 'intermittent' if that frequency was between 20 and 80%; or 'mixed' if fewer than 20% of density profiles were stratified. Some data sets only had surface measurements and their stratification patterns were based on information from the literature and data providers. To address Q4 we looked for relationships between the 86 site-specific mean bloom frequencies and environmental factors (latitude, water depth, tidal amplitude, stratification pattern, and overall mean salinity, temperature, TN and TP) with a General Additive Model (GAM) using a spline function where the degree of smoothing was determined by generalised cross validation.

Results and Discussion

Phytoplankton blooms as essential features of estuarine-coastal ecosystems

While there is no universally accepted definition of what constitutes a bloom, the notion of a substantial deviation above background phytoplankton biomass is common to all definitions. Of the 29,462 carbon-biomass measurements compiled for this analysis, 7368 (25%) were identified as blooms defined as deviations from standard seasonal patterns at each site (Fig. 2). This result is robust evidence that blooms are common in estuarine-coastal waters, although their raw

frequency (number of bloom samples relative to total number of samples) varied across sites, ranging from 6% of observations in Laholmsbukten (transition area between the North Sea and Baltic Sea) to 43% of observations in South San Francisco Bay.

The composition of bloom communities is important because the phytoplankton include a diversity of taxonomical groups having a wide range of cell size, motility, nutritional requirements, life history, biochemical compositions, and food quality for consumers (Litchman and Klausmeier, 2008). Diatoms play central roles in silica cycling, some cyanobacteria species in nitrogen cycling, all species in the cycling of carbon, nitrogen, phosphorus and oxygen; large cells are directly accessible to metazoans and their energy content is transferred efficiently in food webs supporting fisheries; diatoms, dinoflagellates and cryptophytes are highly nutritious for consumers because they are enriched in essential fatty acids; blooms of some diatom, dinoflagellate, cyanobacteria and flagellate species are harmful and degrade habitat and water quality. Therefore, the biogeochemical and ecological responses to blooms vary depending on which phytoplankton species are selected by environmental conditions that promote biomass accumulation (Cloern and Dufford, 2005).

Bloom-dominant species

Our data compilation shows that blooms in estuarine-coastal ecosystems are events of biomass production by microphytoplankton (cell size > 20 μm) and, in particular, diatoms. Diatoms were the dominant biomass component in 58% of bloom samples, and dinoflagellates were the dominant biomass component in 19% of blooms (Table 2, Fig. 3). In this sense estuarine-coastal ecosystems function as analogs to diatom- and dinoflagellate-producing coastal upwelling systems (Lassiter et al., 2006; Silva et al., 2009), except the source of new nutrients is land runoff as well as nutrients below the photic zone that are mixed up by wind and tidal mixing rather than the deep ocean. Nanoplankton were minor components of phytoplankton blooms, including cryptophytes that were commonly present but rarely observed as bloom-dominating species (<2% of all blooms). For most samples cryptophytes contributed a modest proportion of the total carbon biomass (Fig. 3C). Cyanobacteria and chlorophytes dominated biomass in only 7 and 3% of bloom samples, respectively. In the remaining 12% of blooms, species from other algal groups

dominated. The most frequent genus in this group was *Phaeocystis* spp., found to dominate 2.5% of all blooms; however, it dominated only in the RhineMeuse-Scheldt (RMS) and Wadden Sea (WS) regions where it dominated 9-14% of blooms (Table S1).

Diatoms are the taxonomic group most often associated with phytoplankton blooms, seconded by dinoflagellates (Sarthou et al., 2005, Carstensen et al., 2007), and supported also by this analysis. This is a robust pattern, built from multi-year sampling records across a wide diversity of coastal habitats. The success of diatoms as a group derive from their capacity to grow rapidly in turbulent high-nutrient environments (Maranon et al. 2012). Diatom cells are capable of fast growth due to rapid nitrogen (especially NO_3-) uptake (Lomas and Glibert, 2000). However, in contrast to smaller nano- and picophytoplankton diatoms require high nutrient concentrations for optimal growth because of their small surface to volume ratios (Sarthou et al., 2005). Estuarine-coastal waters are generally nutrient rich: across the 86 study sites, median values of DIN (10.4 µM), DIP (0.38 µM), and silicate (14.7 µM) were high enough to support fast diatom growth for large parts of the year. Diatoms are also well adapted to varying light levels and physical stress characteristic of shallow coastal systems, especially during spring blooms (Lomas and Glibert, 2000). Diatoms, owing to their high species diversity (Armbrust, 2009), appear to be well adapted to the habitat gradients of estuaries because they were dominant bloom components across the entire temperature ranges and most of the salinity ranges represented by the eight regions (see below).

Dinoflagellates were the second most frequent taxonomic group dominating blooms. Dinoflagellate-dominated spring blooms were not common in all regions, but they regularly occurred in the Baltic Sea (BS), varying inter-annually in importance along with diatom-dominated blooms (Klais et al., 2011). The success of dinoflagellates in the Baltic Sea spring phytoplankton community remains poorly understood, as they are inferior competitors due to their low growth rates and nutrient uptake capacities under nutrient-replete spring conditions (Spilling and Markager, 2008). Chemical suppression of competitors is suggested as the possible mechanism promoting cold water bloom dinoflagellates, such as *Scrippsiella hangoei*, to outcompete the regular spring bloom diatoms, including *Skeletonema costatum* s.l. and *Thalassiosira baltica* (Suikkanen et al., 2011). Like the common bloom-forming diatoms the common bloom-forming

dinoflagellates *Peridiniella catenata* and *S. hangoei* also form benthic resting stages (cysts) that seed blooms in the Baltic Sea (Kremp, 2001).

Pico- and nanophytoplankton taxa generally did not develop blooms in the estuarine-coastal waters we considered. The exceptions were chlorophyte blooms that developed in low-salinity upper reaches of Scheldt Estuary of RMS region, a highly turbid estuary where light intensity is insufficient to support positive net production (Gazeau et al., 2005). This suggests that these chlorophyte blooms might develop upstream in freshwater rather than within the estuary. All other cases of small-celled species dominated blooms (e.g. cyanobacteria in the BS, *Phaeocystis* spp. in the RMS and WS) consisted of colonial forms. The generation time of grazers often determines the outcome of size competition (Kiørboe 1993). In the absence of grazers, especially the fast-reproducing microzooplankton including heterotrophic dinoflagellates, microphytoplankton would always be outcompeted by pico- and nanoplankton, regardless of nutrient concentrations.

Timing of blooms

Our analyses of bloom timing revealed that: (a) spring blooms were common in all regions, (b) timing of the spring blooms varied across regions, and (c) blooms developed any time of year, but the occurrence of other seasonal blooms varied between regions (Fig. 2). Spring blooms occurred with a probability of 30 to 60% across regions. They typically developed during April and May in the Gulf of Bothnia (GB) and Baltic Sea (BS) regions (Fig. 4A,B), but during March in the Danish Straits and Estuaries (DSE) region (in 33% of the samples, Fig. 4C). The spring bloom was even more pronounced in the WS, with 50% of the April observations categorised as blooms (Fig. 4D). Bloom patterns in the RMS region were similar to the WS, although with highest bloom frequency (49%) in May (Fig. 4E). Bloom frequency in Chesapeake Bay (CB) peaked (30-35%) in March-April (Fig. 4F). Blooms were common in San Francisco Bay (SFB) between February and April (Fig. 4G), and they developed in the Neuse River Estuary (NRE) during the first three months of the year, but at lower frequency (~30%) than spring-bloom frequency in other regions (Fig. 4H).

The timing of spring blooms followed latitudinal gradients in the annual solar radiation cycle. In the high-latitude GB and BS regions low incident irradiance from December to March suppresses phytoplankton growth and the bloom seasons were compressed (Fig. 2A,B). The bloom season

began earlier in the DSE region and even earlier in CB and NRE. However, the bloom season began later (April-May) in the RMS and WS regions despite their latitudes below DSE. Coastal sites in these two regions have stronger tidal mixing and are more turbid (see below), which probably delayed the onset of spring blooms. Thus, although the timing of spring blooms is generally governed by latitude, it can be modulated by local physical processes such as sediment suspension by tidal currents.

Beyond the common occurrences of blooms during spring, there was a wide range across regions in the occurrences of blooms during other seasons. Summer and autumn blooms developed in DSE, the WS and the RMS regions. Autumn blooms developed in SFB, winter blooms were common in CB, and summer blooms developed in the NRE. The scatterplots of bloom occurrences in regions DSE (Fig. 2C) and WS (Fig. 2D) suggest that blooms in these regions might be characterized more accurately as episodic rather than seasonal components of phytoplankton variability. Thus, our analyses reveal that there is no one canonical pattern of bloom occurrence. This result is not surprising, given the many processes that regulate algal growth, mortality and transport and their distinct seasonal patterns and frequencies of variability in estuarine-coastal waters (Cloern and Jassby, 2008). For example, river inflow is both a source of nutrients to promote blooms and a flushing process to remove phytoplankton biomass, so blooms develop at intermediate flows that optimize the balance between these processes (Peierls et al., 2012). Pulses of riverine inflow provide both nutrients and fresh water as a source of buoyancy to stratify estuaries, so pulse events promote large (and harmful) dinoflagellate blooms in the NRE (Hall et al., 2013). Seasonal changes in wind direction alter the retention time of water in coastal bays and promote development of red-tide dinoflagellates in Hong Kong coastal waters during the winter monsoon season (Yin, 2003). Seasonal thermal stratification of the BS leads to nutrient depletion in surface waters and low bloom frequency during summer (Fig. 4B), but short-term events of thermal stratification during heat waves can promote summer or autumn blooms in nutrient-rich estuaries such as SFB (Cloern et al., 2005). Some blooms are initiated in shallow coastal waters when light penetrates deep enough to germinate diatom resting stages in bottom sediments (Shikata et al., 2008) or when storms transport dinoflagellate cysts from sediments into the water column (Kremp, 2001). In tidal systems the neap-spring cycle induces biweekly patterns of mixing and stratification (an increase in nutrients by tidal mixing during spring tides and stratification

during neap tides), with blooms developing on neap tides and dissipating on spring tides (Cloern, 1996). Tidal currents in SFB have a semi-annual component with largest spring tides around the solstices, so intense mixing suppresses stratification and blooms during summer and winter (Fig. 4G). The characteristic winter-spring diatom bloom disappeared from Narragansett Bay after decades of warming (Nixon et al., 2009), and the autumn blooms in SFB (Fig. 4G) appeared for the first time in 1999 after a shift in climate forcing of the North Pacific Ocean (Cloern et al., 2007). Our synthesis of many data sets confirms that beyond the common occurrences of spring blooms there is no consistent seasonal pattern of bloom timing in these estuarine-coastal ecosystems.

Key species

A small number of common species developed blooms in both North American and European waters: diatoms *Skeletonema costatum* s.l. (representing several species, see below), *Cerataulina pelagica*, *Dactyliosolen fragilissimus*, and dinoflagellates *Heterocapsa triquetra* and *Prorocentrum cordatum*. Each of these species developed blooms within different ranges of temperature and salinity. Blooms dominated by *S. costatum* s.l. were most common at salinities between 14 and 28 and at low temperatures (Fig. 5A,B), and this diatom did not dominate blooms in the low salinity GB or in the warm NRE regions. *C. pelagica* blooms were most common at salinities between 8 and 20 and temperatures between 5 and 12 ˚C (Fig. 5A,B) and rarely occurred in waters with salinity less than 8, e.g. the coastal sites in the GB and BS. It did not develop blooms in the high-salinity SFB or in the warm waters of NRE. Blooms dominated by *D. fragilissimus* occurred in the CB, DSE, RMS and WS regions (Table 2) and most commonly in the salinity range 10-27 and when temperature exceeded 12 ˚C. Blooms of *H. triquetra* and *P. cordatum* were most common in the salinity range 6-15 (Fig. 5C), and *P. cordatum* blooms generally occurred at higher temperatures than *H. triquetra* blooms (Fig. 5D). Both dinoflagellate species formed blooms in all regions except the low-salinity regions of GB and BS (Table 2).

The most frequent and widespread bloom forming "species", *Skeletonema costatum* s.l., is perceived as a fast-growing and highly adaptable diatom that thrives in coastal waters throughout the world. In many areas, species identified as *S. costatum* s.l. are among the most important contributors to phytoplankton blooms. The apparent cosmopolitan distribution of *S. costatum* s.l. in coastal areas worldwide can now be attributed partly to its inclusion of cryptic species. Recent

molecular and morphological analyses revealed that this "species" is a complex comprising at least eight morphologically similar species in addition to *S. costatum* (Sarno et al., 2005). Therefore, the ubiquity of *S. costatum* s.l. blooms might be explained by the genetic and physiological variability within this morphospecies, reflecting a limitation in our capacity to understand population dynamics and biogeography of phytoplankton species from microscopy (Jakobsen et al., this issue).

The other two diatom species with intercontinental distribution and high frequency of bloom formation – *Cerataulina pelagica* and *Dactyliosolen fragilissimus* - are also chain forming centric diatoms having a form similar to *S. costatum* s.l., although these species are larger (Harrison et al., this issue). Together, *S. costatum* s.l., *C. pelagica* and *D. fragilissimus* dominated 17% of all blooms identified in our data compilation, i.e. almost every third diatom bloom. Other frequent and globally distributed bloom-forming diatoms were chain-forming taxa (*Thalassiosira* spp., *Chaetoceros* spp.) or, if solitary, very large cells (*Coscinodiscus* spp.). Thus, while there is no canonical seasonal pattern of bloom formation in estuarine-coastal waters, there are life forms -- chains of smaller celled or solitary large celled diatoms -- that are highly adapted to the opportunities for fast population growth in shallow, turbulent, turbid and nutrient-enriched coastal ecosystems.

Dinoflagellates *H. triquetra* and *P. cordatum* dominated blooms in all regions except the low-salinity GB and the mesotidal WS. In the BS the non-toxic *H. triquetra* forms extensive blooms during summer, when thermal stratification is strong, and the blooms disappear in late August or early September (Olli, 2004). This pattern was also observed in the DSE, whereas in the CB and NRE, *H. triquetra* typically dominated blooms in winter-spring when freshwater discharge establishes strong salinity stratification. *P. cordatum* blooms occur in systems affected by large freshwater discharges, which are rich in nutrients and dissolved organic matter (Grzebyk and Berland, 1996). However, *P. cordatum* blooms were most common in late summer and autumn, when freshwater discharges are normally at minimum, suggesting that *P. cordatum* blooms could be based on organic nutrients (Heil et al., 2005). Preferred water temperatures (20-24 °C) and salinities (10-16) reported by Velikova and Larsen (1999) are consistent with our results (Fig. 5).

Blooms as responses to environmental variability

The 86 sites included in our data compilation represent a wide range of physical and chemical conditions (Table S2), providing a robust empirical record to search for general patterns of association between environmental factors and bloom dynamics. Among these factors we considered: nutrient concentrations that set the potential magnitude for bloom development; latitude that represents gradients of climate (solar radiation, temperature) cycles; water depth that influences light availability and strength of benthic-pelagic coupling; salinity as a habitat attribute and indicator of river influence; tidal amplitude as an indicator of mixing energy; and density stratification that isolates phytoplankton from benthic consumers and damps mixing to establish vertical gradients of light and nutrients. First we looked for associations between these factors and frequency of bloom occurrence. Although ANOVA revealed no significant differences in bloom frequency between the eight regions ($F_{7,78}$=1.13; p=0.3553), differences among sites were large ranging from 5% bloom frequency at Laholmsbukten (DSE region) to 35% bloom frequency at Veerse Meer (RMS region). Our GAM analysis showed that no significant component of this variability could be explained by any of the environmental factors we considered or their interactions, despite the large ranges of latitude, depth, tidal amplitude, stratification, and annual mean temperature, salinity, TN and TP (Fig. 6).

This result suggests that there may be no general relationships between physical-chemical factors and variability of bloom occurrence across the wide diversity of coastal ecosystems. These relationships do exist within individual ecosystems, but they appear to be site specific and they change over time. For example, nutrient supply sets the potential for phytoplankton production, but the realisation of that potential varies greatly across estuarine-coastal ecosystems. Some, such as Chesapeake Bay, appear to respond strongly to changes in nutrient supply whereas others, such as San Francisco Bay, have shown resistance to nutrient enrichment (Cloern 2001). This implies that the efficiency with which nitrogen and phosphorus are converted into phytoplankton biomass, and the potential for bloom development, varies across sites. Moreover, grand experiments of nutrient reduction have shown that this efficiency changes over time (Carstensen et al., 2004, 2007). Perhaps then we should not be surprised by the absence of general relationships between nutrient loading, or any other environmental factor, and bloom dynamics indexed as frequency of occurrence. Instead, phytoplankton production dynamics appear to be regulated by site-specific, idiosyncratic and time-varying combinations of all the factors that

regulate the balance between production, consumption and transport (Cloern et al., 2014) instead of global relationships that operate uniformly across all ecosystems.

On the other hand, we did find general patterns in the taxonomic composition of blooms and, in particular, variability of bloom communities along the gradients of salinity and temperature. Diatoms developed blooms across the full ranges of temperature and salinity sampled (Fig. 3A,B), although some niche partitioning was evident among the most frequently blooming diatoms – *Skeletonema costatum* s.l. prevailing at lowest, *Cerataulina pelagica* at intermediate and *Dactyliosolen fragilissimus* at higher temperatures (Fig. 5B). At the group level, chlorophytes and cyanobacteria developed blooms only at low salinities in systems with strong freshwater input, and dinoflagellate blooms occurred at low salinities where they take advantage of their vertical migration ability to swim down and take up nutrients (Fig. 3A). Dinoflagellate blooms were more frequent at both high and low temperatures, but less so in intermediate temperatures (Fig. 3B). The dominance of dinoflagellates at low temperatures is due to cold water dinoflagellates *Peridiniella catenata* and *Scrippsiella hangoei* that frequently dominated the BS spring blooms, where they not only outcompeted, but even exceeded the typical diatom spring bloom biomass (Klais et al., 2011). These two species showed highest frequency of bloom dominance among dinoflagellates, but they only occurred in the Baltic Sea (Table 2). Lists of the top bloom dominants (Table 3) separate the regions with classical diatom and dinoflagellate dominated communities from coastal areas that frequently experienced blooms formed by other algal groups – e.g. the chlorophytes in NRE and RMS, and haptophyte *Phaeocystis* spp. in WS. Chlorophytes tend to respond positively to elevated freshwater input events (Paerl et al., 2014), which is typical of both regions where chlorophytes dominated blooms. Thus, although there is great variability in environmental conditions, phytoplankton biomass and bloom frequency across the diversity of ecosystem types, one general rule does emerge from our synthesis: estuarine-coastal waters support blooms that are primarily composed of high diatom and dinoflagellate biomass, and blooms of chlorophytes and cyanobacteria are restricted to low-salinity habitats.

PERSPECTIVES

16

Meta-analyses can be powerful approaches for identifying general patterns of community dynamics through synthesis of many site-based studies. For example, a meta-analysis spanning 203 terrestrial species identified trends of advanced timing of animal migrations and plant flowering associated with global warming (Parmesan, 2007). This discovery was grounded in direct observations of plants and animals identified to the species level at many sites. We cannot observe phytoplankton directly in their environment, many specimens can not be identified to species level by microscopy, and species identified by their morphology can be complexes of an unknown number of cryptic species (Amato et al., 2007). Moreover, there is a mismatch in the time scales of phytoplankton biomass variability (days-weeks) and sampling (weeks-months), so bloom events are not well documented and resolved. Finally, there is significant variability among laboratories in the methods used to sample, preserve and count, identify and measure phytoplankton cell volume, the taxonomic nomenclature used and resolution reported. These all place substantial constraints on the reliability of conclusions that can be drawn from meta-analyses of phytoplankton time series. However, these constraints notwithstanding, several key results emerged from our meta-analysis of phytoplankton community variability in estuarine-coastal sites.

First, phytoplankton blooms across these sites did not follow a common seasonal pattern; events of high biomass were most common in spring, but they occurred at any time during the year. This result appears robust because it corroborates another meta-analysis showing that the timing of annual chlorophyll *a* peaks varies across and within estuarine-coastal ecosystems (Cloern and Jassby, 2008). The within-ecosystem variability includes shifting seasonal patterns over time, including loss of the winter-spring diatom from Narragansett Bay (Borkman and Smayda, 2009), earlier onset of cyanobacteria blooms in the Baltic Sea (Kahru and Elmgren, 2014), and new occurrences of autumn blooms in San Francisco Bay (Cloern et al., 2007). One characteristic of estuarine-coastal ecosystems appears to be their diverse and variaable seasonal patterns of bloom occurrence.

Second, blooms at these sites were characteristically dominated by microplankton (cell size >20 µm) – a distinct contrast from biomass dominance by picoplankton in oligotrophic regions of the world ocean. This predominance of larger algal forms is consistent with other meta-analyses

showing that the microplankton component of phytoplankton communities increases in proportion to resources and biomass (Marañón et al., 2012). Microplankton are consumed directly by metazoans so their predominance implies efficient energy transfer from producers to consumers in these high-biomass, high-nutrient systems.

Third, phytoplankton blooms were most commonly dominated by diatoms, and this was true across the full ranges of salinity and temperature. Therefore, diatoms as a group appear to be highly adapted to exploit opportunities for biomass growth along the land-ocean continuum. Our meta-analysis suggests that estuarine-coastal waters function as diatom-producing systems. This has important ecological and biogeochemical implications because of the large amounts of high food value of diatoms to consumers and the unique role played by diatoms in the cycling of silica as it is carried from rivers to oceans.

Finally, our meta-analysis illustrates the elusiveness of global rules defining how phytoplankton respond to environmental variability in coastal areas. This implies that bloom dynamics are regulated by site-specific variability of those factors and/or processes not captured in monitoring programs such as species interactions and life cycle events. A grand challenge remains to synthesize observations from place-based studies into a general model that explains the processes underlying phytoplankton patterns in estuarine-coastal waters. Keys to success will be enhanced efforts to: measure processes of species interaction; apply emerging technologies such as imaging flow cytometry and molecular tools to better resolve biomass variability and taxonomic composition of phytoplankton communities at higher temporal resolution; and compare time series from tropical and subtropical ecosystems that are under-sampled and may not follow the patterns of temperate estuaries, such as diatom-dominance (Cotovicz et al., 2015).

Agency), the DEVOTES project (funded by the EU 7th framework program), and the COCOA project (funded by the BONUS program for Baltic Sea research).

REFERENCES

Amato, A., Kooistra, W.H.C.F., Ghiron, J.H.L., Mann, D.G., Proschold, T. and Montresor, M., 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. Protist 158, 193–207.

Armbrust, E.V., 2009. The life of diatoms in the world's oceans. Nature 459, 185-192.

Bacher, C., Duarte, P., Ferreira, J.G., Heral, M., Raillard, O., 1998. Assessment and comparison of the Marennes-Oléron Bay (France) and Carlingford Lough (Ireland) carrying capacity with ecosystem models. Aquatic Ecology 31, 379–394.

Borkman, D. G., Smayda, T., 2009. Multidecadal (1959–1997) changes in Skeletonema abundance and seasonal bloom patterns in Narragansett Bay, Rhode Island, USA. Journal of Sea Research 61,84–94.

Brown, C.A., Ozretich, R.J., 2009. Coupling between the coastal ocean and Yaquina Bay, Oregon: Importance of oceanic inputs relative to other nitrogen sources. Estuaries and Coasts 32, 219–237.

Caffrey, J., Cloern, J.E., Grenz, C., 1998. Changes in production and respiration during a spring phytoplankton bloom in San Francisco Bay, California, USA: implications for net ecosystem metabolism. Marine Ecology Progress Series 172, 1–12.

Carstensen, J., Conley, D.J., Henriksen, P., 2004. Frequency, composition, and causes of summer phytoplankton blooms in a shallow coastal ecosystem, the Kattegat. Limnology and Oceanography 49, 190–201.

Carstensen, J., Frohn, L.M., Hasager, C.B., Gustafsson, B.G., 2005. Summer algal blooms in a coastal ecosystem: the role of atmospheric deposition versus entrainment fluxes. Estuarine Coastal Shelf Science 62, 595–608.

Carstensen, J., Conley, D.J., Andersen, J.H., Ærtebjerg, G., 2006. Coastal eutrophication and trend reversal: a Danish case study. Limnology and Oceanography 51, 398–408.

Carstensen, J., Henriksen, P., Heiskanen, A.-S., 2007. Summer algal blooms in shallow estuaries: Definition, mechanisms, and link to eutrophication. Limnology and Oceanography 52, 370–384.

Cloern, J.E., 1996. Phytoplankton bloom dynamics in coastal ecosystems: A review with some general lessons from sustained investigation of San Francisco Bay, California. Reviews of Geophysics 34, 127–168.

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

Cloern, J.E., Dufford, R., 2005. Phytoplankton community ecology: principles applied in San Francisco Bay. Marine Ecology Progress Series 285, 1–28.

Cloern, J.E., Schraga, T., Lopez, C., Knowles, N., Labiosa, R., Dugdale, R., 2005. Climate anomalies generate an exceptional dinoflagellate bloom in San Francisco Bay. Geophysical Research Letters 32(14), 1-5.

Cloern, J.E., Jassby, A.D., Thompson, J.K., Hieb, K.A., 2007. A cold phase of the East Pacific triggers new phytoplankton blooms in San Francisco Bay. Proceedings of the National Academy of Sciences 104, 18561–18565.

Cloern, J.E., Jassby, A., 2008. Complex seasonal patterns of primary producers at the land-sea interface. Ecology Letters 11, 1294–1303.

Cloern, J.E., Jassby, A., 2010. Patterns and scales of phytoplankton variability in estuarine-coastal ecosystems. Estuaries and Coasts 33, 230–241.

Cloern, J.E., Foster, S.Q., Kleckner, A.E., 2014. Phytoplankton primary production in the world's estuarine-coastal ecosystems. Biogeosciences 11, 2477–2501.

Cotovicz, L. C. J., Knoppers, B. A., Brandini, N., Costa Santos, S. J., Abril, G., 2015. A large CO_2 sink enhanced by eutrophication in a tropical coastal embayment (Guanabara Bay, Rio de Janeiro, Brazil). Biogeosciences Discussions 12, 4671–4720.

Edler, L., 1979. Recommendations on methods to marine biological studies in the Baltic Sea. Phytoplankton and chlorophyll. The Baltic Marine Biologists Publications, 5, 38 pp.

Gazeau, F., Gattuso, J.-P., Middelburg, J., Schiettecatte, L.-S., Frankignoulle, M., Borges, A., 2005. Planktonic and whole system metabolism in a nutrient-rich estuary (the Scheldt estuary). Estuaries and Coasts 28, 868–883.

Grzebyk, D., Berland, B., 1996. Influnce of temperature, salinity and irradiance on growth of *Prorocentrum minimum* (*Dinophyceae*) from the Mediterranean Sea. Journal of Plankton Research 18, 1837–1849.

Hall, N.S., Paerl, H.W., Peierls, B.L., Whipple, A.C., Rossignol, K.C., 2013. Effects of climatic variability on phytoplankton community structure and bloom development in the eutrophic, microtidal, New River Estuary, North Carolina, USA. Estuarine Coastal Shelf Science 117, 70–82.

Heil, C.A., Glibert, P.M., Fan, C., 2005. *Prorocentrum minimum* (Pavillard) Schiller: A review of a harmful algal bloom species of growing worldwide importance. Harmful Algae 4, 449–470.

Hopkinson, C.J., Smith, E.M., del Giorgio, P.A., Williams, P.J.l.B., 2005. Estuarine respiration: an overview of benthic, pelagic and whole system respiration, in P.A. del Giorgio and P.J.l.B. Williams, editors. Respiration in Aquatic Ecosystems. Oxford University Press.

Houde, E.D., Rutherford, E.S. 1993. Recent trends in estuarine fisheries - predictions of fish production and yield. Estuaries 16, 161–176.

Iverson, R.L., Curl, H.C.Jr., O'Connors, H.B.J., Kirk, D., Zakar, K. 1974. Summer phytoplankton blooms in Auke Bay, Alaska, driven by wind mixing of the water column. Limnology and Oceanography 19, 271–278.

Jakobsen, H.H., Carstensen, J., Harrison, P.J., Zingone, A., In press. Estimating time series phytoplankton biomass: Species identifications and comparing volume to carbon scaling ratios. Estuarine and Coastal and Shelf Science (this issue).

Jurgensone, I., Carstensen, J., Ikauniece, A., Kalveka, B., 2011. Long-term changes and controlling factors of phytoplankton community in the Gulf of Riga (Baltic Sea). Estuaries and Coasts 34, 1205–1219.

Kahru, M., Elmgren, R., 2014. Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. Biogeosciences 11, 3619–3633.

Kiørboe, T., 1993. Turbulence, phytoplankton cell size and the structure of pelagic food webs. Advances in Marine Biology 29, 1–72.

Klais, R., Tamminen, T., Kremp, A., Spilling, K., Olli, K., 2011. Decadal-scale changes of dinoflagellates and diatoms in the anomalous Baltic Sea spring bloom. PLoS ONE 6, e21567.

Kremp, A., 2001. Effects of cyst resuspension on germination and seeding of two bloom forming dinoflagellates in the Baltic Sea. Marine Ecology Progress Series 216, 57–66.

Lassiter. A.M., Wilkerson, F.P., Dugdale, R.C., Hogue, V.E., 2006. Phytoplankton assemblages in the CoOP-WEST coastal upwelling area. Deep-Sea Research Part II 53, 3063–3077.

Litchman, E., Klausmeier, C.A., 2008. Trait-based community ecology of phytoplankton. Annual Review of Ecology, Evolution, and Systematics 39, 615–639.

Lomas, M.W., Glibert, P.M., 2000. Comparison of nitrate uptake, storage, and reduction in marine diatoms and flagellates. Journal of Phycology 36, 903–913.

Mallin, M.A., Paerl, H.W., Rudek, J., 1991. Seasonal phytoplankton composition, productivity and biomass in the Neuse River estuary, North Carolina. Estuarine Coastal Shelf Science 32, 609–623.

Marañón, E., Cermeño, P., Latasa, M., Tadonléké, R.D., 2012. Temperature, resources, and phytoplankton size structure in the ocean. Limnology and Oceanography 57, 1266–1278.

Margalef, R., 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. Oceanologica Acta 1, 493–509.

Nixon, S.W., Fulweiler, R.W., Buckley, B.A., Granger, S.L., Nowicki, B.L., Henry, K.M., 2009. The impact of changing climate on phenology, productivity, and benthic-pelagic coupling in Narragansett Bay. Estuarine Coastal and Shelf Science 82, 1–18.

Odebrecht, C., Abreu, P.C., Carstensen, J., In press. Retention time as controlling factor of short-term chlorophyll a variability in a shallow microtidal subtropical estuary. Estuarine Coastal Shelf Science (this issue).

Olli, K., 2004. Temporary cyst formation of *Heterocapsa triquetra* (*Dinophyceae*) in natural populations. Marine Biology 145, 1–8.

Olli, K., Paerl, H.W., Klais, R., In press. Diversity of coastal phytoplankton assemblages – cross ecosystem comparison. Estuarine Coastal Shelf Science (this issue).

Paerl, H.W., 1997. Coastal eutrophication and harmful algal blooms: Importance of atmospheric deposition and groundwater as "new" nitrogen and other nutrient sources. Limnology and Oceanography 42, 1154–1165.

Paerl, H.W., Huisman, J., 2009. Climate Change: A Catalyst for Global Expansion of Harmful Cyanobacterial Blooms. Environmental Microbiology Reports 1, 27–37.

Paerl, H.W., Hall, N.S., Peierls, B.J., Rossignol, K.L., Joyner, A.R., 2014. Hydrologic variability and its control of phytoplankton community structure and function in two shallow, coastal, lagoonal ecosystems: the Neuse and New River Estuaries, North Carolina, USA. Estuaries and Coasts 37 (Suppl. 1), S31-S45.

Parmesan, C., 2007. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. Global Change Biology 13,1860–1872.

Peierls, B.L., Hall, N.S., Paerl, H.W., 2012. Non-monotonic responses of phytoplankton biomass accumulation to hydrologic variability: A comparison of two coastal plain North Carolina estuaries. Estuaries and Coasts, 1–17.

Petersen, J.K., Hansen, J.W., Laursen, M.B., Clausen, P., Carstensen, J., Conley, D.J., 2008 Regime shift in a coastal marine ecosystem. Ecological Applications 18, 497–510.

Sarno, D., Zingone, A., Montresor, M., 2010. A massive and simultaneous sex event of two *Pseudo-nitzschia* species. Deep Sea Research Part II: Topical Studies in Oceanography 57, 248–255.

Sarno, D., Kooistra, W.H.C.F., Medlin, L.K., Percopo, I., Zingone, A., 2005. Diversity in the genus *Skeletonema* (Bacillariophyceae). II. An assessment of the taxonomy of *S. costatum*-like species with the description of four new species. Journal of Phycology 41, 151–176.

Sarthou, G., Timmermanns, K.R., Blain, S., Tréguer, P., 2005. Growth physiology and fate of diatoms in the ocean: a review. Journal of Sea Research 53, 25–42.

Shikata, T., Nagasoe, S., Matsubara, T., Yoshikawa, S., Yamasaki, Y., Shimasaki, Y., Oshima, Y., Jenkinson, I. R., Honjo, T., 2008. Factors influencing the initiation of blooms of the raphidophyte *Heterosigma akashiwo* and the diatom *Skeletonema costatum* in a port in Japan. Limnology and Oceanography 53, 2503–2518.

Silva, A., Palma, S., Oliveira, P.B., Moita, M.T., 2009. Composition and interannual variability of phytoplankton in a coastal upwelling region (Lisbon Bay, Portugal). Journal of Sea Research 62, 238–249.

Smayda, T.J., 1997. Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. Limnology and Oceanography 42, 1137–1153.

Smayda, T.J., Reynolds, C.S., 2001. Community assembly in marine phytoplankton: application of recent models to harmful dinoflagellate blooms. Journal of Plankton Research 23, 447–461.

Spilling, K., Markager, S., 2008. Ecophysiological growth characteristics and modelling of the onset of the spring bloom in the Baltic Sea. Journal of Marine Systems 73, 323–337.

Strathmann, R.R., 1967. Estimating the organic carbon content of phytoplankton from cell volume or plasma volume. Limnology and Oceanography 12, 411–418.

Suikkanen, S., Hakanen, P., Spilling, K., Kremp, A., 2011. Allelopathic effects of Baltic Sea spring bloom dinoflagellates on co-occurring phytoplankton. Marine Ecology Progress Series 439, 45–55.

Tett, P., Gowen, R., Grantham, B., Jones, K., Miller, B.S., 1986. The phytoplankton ecology of the Firth of Clyde sealochs Striven and Fyne. Proceedings of the Royal Society Edinburgh, Section B 90, 223–238.

Utermöhl, H., 1958. Zur Vervollkommung der quantitativen phytoplankton-methodik. Internationale Vereinigung fuer Theoretische and Angewandte Limnologie and Verhandlungen 17, 47–49.

Velikova, V., Larsen, J., 1999. The *Prorocentrum cordatum/Prorocentrum minimum* taxonomic problem. GRANA 8(3-4), 108–112.

Winder, M., Cloern, J.E., 2010. The annual cycles of phytoplankton biomass. Philosophical Transactions of the Royal Society B-Biological Sciences 365, 3215–3226.

Yin, K., 2003. Influence of monsoons and oceanographic processes on red tides in Hong Kong waters. Marine Ecology Progress Series 262, 27–41.

FIGURE LEGENDS

Fig. 1: Location of the 86 coastal sites distributed among regions (specified by different colours): A) San Francisco Bay (SFB), B) Neuse River Estuary (NRE), C) Chesapeake Bay (CB), D) Rhine-Meuse-Scheldt delta (RMS), E) Wadden Sea (WS), F) Danish Strait and Estuaries (DSE), G) Baltic Sea (BS), and H) Gulf of Bothnia (GB). Note that one Wadden Sea station is plotted in F). Scaling varies between charts. Sites are located in North America (A-C) and in Northwestern Europe (D-H).

Fig. 2: Partitioning samples of total phytoplankton carbon biomass into bloom and non-bloom observations for eight different coastal sites; one example from each region. Dark solid line shows the estimated mean seasonal carbon biomass for non-bloom observations and the grey solid line is the 99th percentile of the non-bloom prediction interval, separating non-bloom and bloom observations. The eight coastal sites are displayed in a north-south latitudinal sequence. A few bloom observations were outside of the plotting range and not shown. Note the difference in biomass scale between sites.

Fig. 3: Functional groups represented in observed blooms versus salinity, temperature and the carbon biomass proportion of the dominating species. Bloom observations were divided into groups based on ambient salinity and temperature rounded to the nearest integer and groups of carbon biomass proportion rounded to the nearest 5 %.

Fig. 4: Seasonal distribution of bloom frequency (blue line) and the proportion of the total carbon biomass for the dominant species (right axis), separated into blooms (red line) and non-blooms (grey line), averaged over stations within each of the eight regions (n_{stat}=number of stations). The eight regions are ordered according to decreasing latitude.

Fig. 5: Proportion of blooms dominated by selected diatom (top) and dinoflagellates (bottom) species versus salinity (A, C) and temperature (B, D). Bloom observations were divided into groups for salinity and temperature rounded to the nearest integer. The selected dominant bloom species were chosen to have a taxonomical resolution at the species level and have a broad regional presence among the most common dominating bloom species in Table 2.

Fig. 6: Bloom frequency versus environmental conditions (Table S2): A) Latitude, B) temperature, C) water depth, D) stratification pattern, E) salinity, F) tidal amplitude, G) total nitrogen (TN) and H) total phosphorus (TP). The eight regions are shown in different colours using the abbreviations from Table 1. The estimated generalized additive models (GAM) are shown as solid lines with statistics inserted (n=86, except for TN and TP where n=81).

Table 1: Data sets used in the present study listing time span, number of stations and phytoplankton surface samples included, as well as the source of the data sets. Note that all the stations were not monitored throughout the entire period. The data were grouped into eight regions: BS=Baltic Sea, CB=Chesapeake Bay, DSE=Danish Straits and Estuaries, GB=Gulf of Bothnia, NRE=Neuse River Estuary, RMS=Rhine-Meuse-Scheldt, SFB=San Francisco Bay, WS=Wadden Sea.

Monitoring program	Period	# of stations	# of samples	Provider	Region
Chesapeake Bay	1984-2009	21	7197	Chesapeake Bay Monitoring Program	CB
Denmark	1979-2013	18	7249	Aarhus University, DCE	DSE, WS
Finland	1966-2010	11	4542	Finnish Environment Institute	BS, GB
Germany, regional	1999-2012	1	1002	Niedersachsen, NLPV	WS
Latvia	1976-2012	4	905	Inst. of Aquatic Ecology, Univ. of Latvia	BS
Netherlands	1990-2011	14	5645	Rijkswaterstaat	RMS, WS
North Carolina	2000-2013	4	624	NC Dept. of Environ. & Natural Res.	NRE
San Francisco Bay	1992-2013	4	727	US Geological Survey, Menlo Park, CA	SFB
Sweden	1983-2012	9	1571	Swedish Meteorological Hydrological Inst.	BS, DSE, GB

Table 2: Listing of the most common dominant bloom species (>1 % of the bloom observations). For each dominant bloom species the number of bloom observations and the number of phytoplankton observations (as totals and as percentages) are listed, where it contributed most to the carbon biomass in the sample. Regions, where the species dominated at least one sample, are also listed (abbreviations, see Table 1).

Taxonomical group and dominating bloom genus or species	# of blooms dominated		# of samples dominated		Regions with observed dominant bloom species
Chlorophytes					
- *Chlorophyceae*	127	1.72%	201	0.68%	CB, DSE, RMS, WS
Cryptophytes					
- *Cryptophyceae*	94	1.28%	1827	6.20%	CB, DSE, RMS, WS
Cyanobacteria					
- *Planktothrix agardhii*	147	1.99%	318	1.08%	BS, GB, RMS, WS
Diatoms					
- *Skeletonema costatum* s.l.	673	9.13%	2065	7.01%	BS, CB, DSE, RMS, SFB, WS
- *Cerataulina pelagica*	406	5.51%	798	2.71%	CB, DSE, RMS, WS
- *Coscinodiscus* spp.	274	3.72%	1379	4.68%	BS, CB, DSE, NRE, RMS, SFB, WS
- *Odontella sinensis*	208	2.82%	787	2.67%	DSE, RMS, WS
- *Dactyliosolen fragilissimus*	207	2.81%	430	1.46%	CB, DSE, RMS, WS
- *Achnanthes taeniata*	144	1.95%	219	0.74%	BS, DSE,GB
- *Thalassiosira* spp.	144	1.95%	565	1.92%	BS, CB, DSE, RMS, WS
- *Chaetoceros* spp.	90	1.22%	378	1.28%	BS, CB, DSE, GB, RMS, WS
- *Pseudo-nitzschia delicatissima*	84	1.14%	156	0.53%	DSE, WS
- *Thalassiosira baltica*	84	1.14%	192	0.65%	BS
- *Cyclotella* spp.	83	1.13%	309	1.05%	CB, DSE, RMS
- *Rhizosolenia setigera*	77	1.05%	271	0.92%	CB, DSE, RMS, SFB, WS
- *Odontella regia*	76	1.03%	297	1.01%	RMS. WS
- *Proboscia alata*	74	1.00%	180	0.61%	CB, DSE
Dinoflagellates					
- *Peridiniella catenata*	361	4.90%	699	2.37%	BS, DSE, GB
- *Scrippsiella hangoei*	160	2.17%	218	0.74%	BS
- *Heterocapsa triquetra*	109	1.48%	374	1.27%	BS, CB, DSE, NRE, RMS, SFB
- *Glenodinium* spp.	87	1.18%	174	0.59%	BS, CB
- *Neoceratium tripos*	80	1.08%	549	1.86%	DSE
- *Gymnodinium* spp.	76	1.03%	474	1.61%	BS, CB, DSE, WS
- *Prorocentrum cordatum*	75	1.02%	262	0.89%	BS, CB, DSE, NRE, RMS, SFB
Other groups					
- *Phaeocystis* spp.	184	2.50%	298	1.01%	RMS, WS
- *Unidentified*	127	1.72%	902	3.06%	DSE, GB, RMS, WS

<Fig. 1>
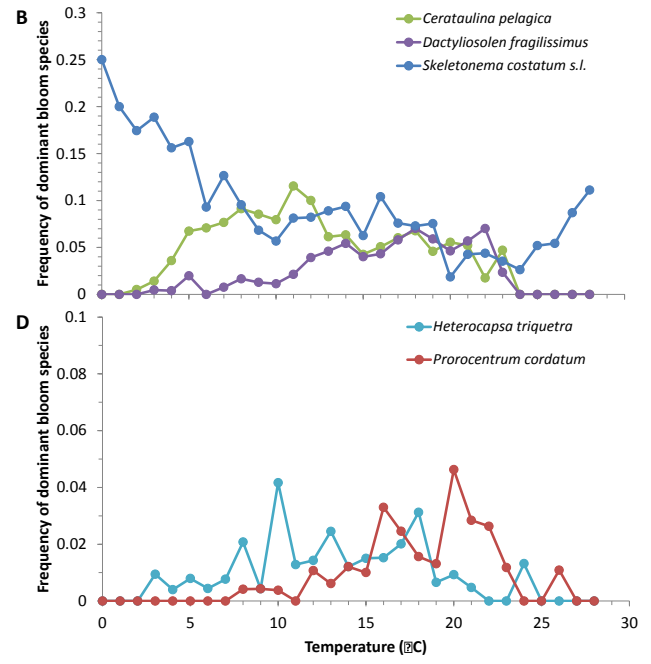
<Fig. 2>

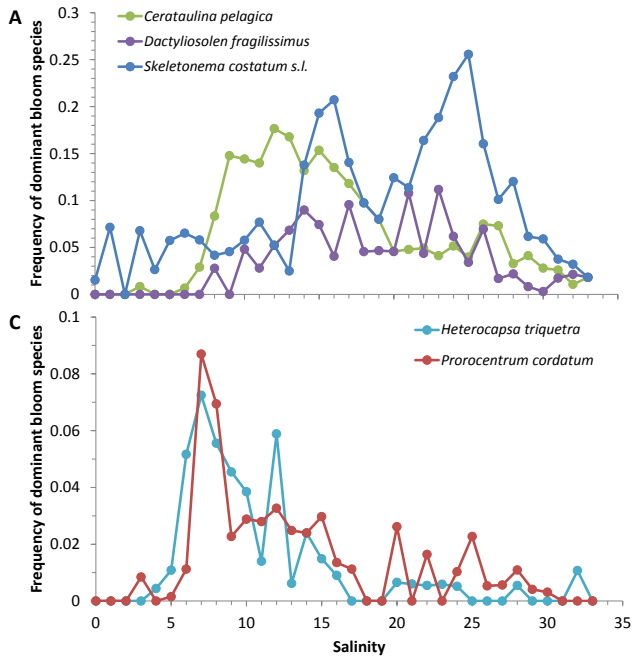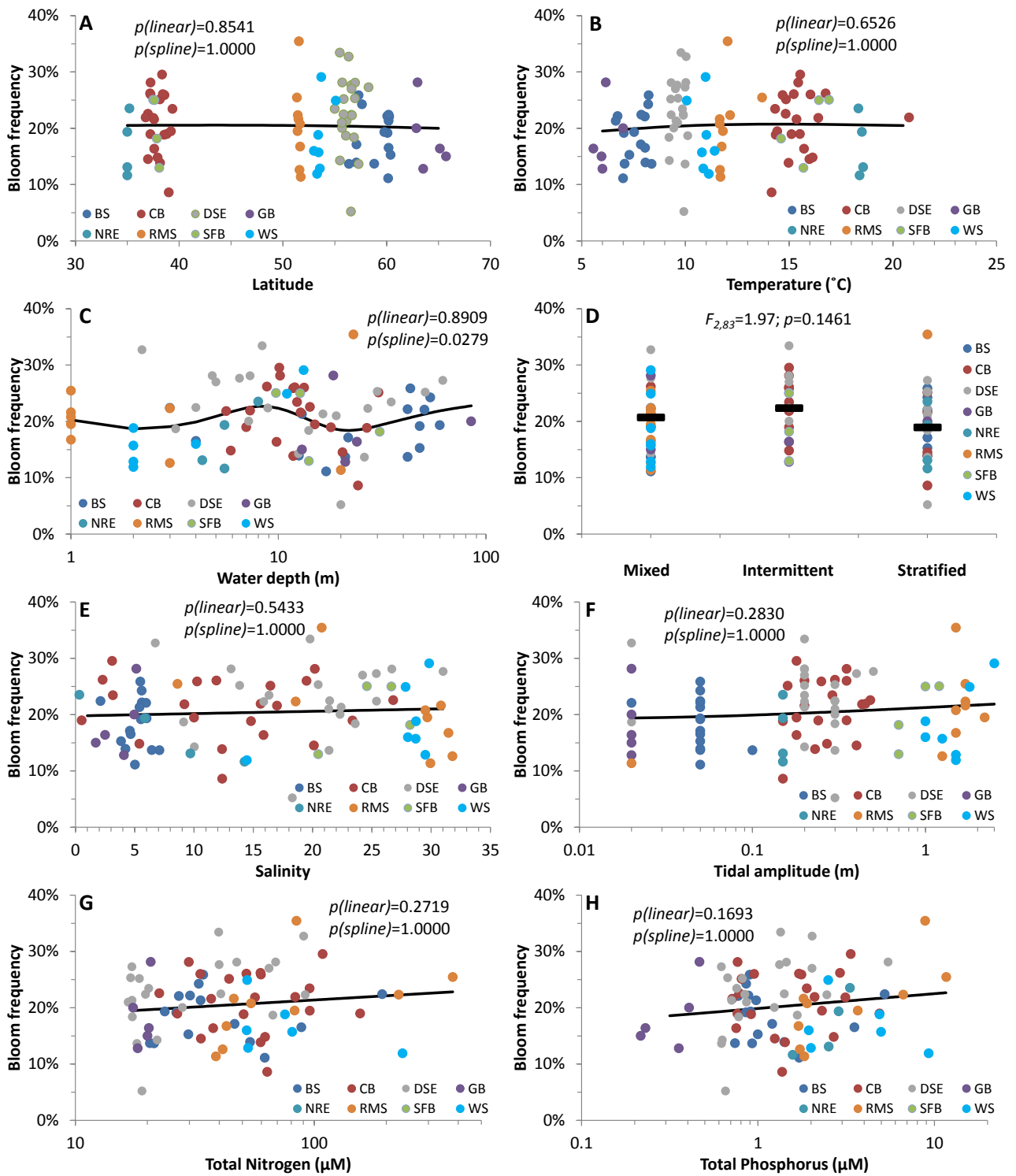<Fig. 3>

<Fig. 4>

4

<Fig. 5>
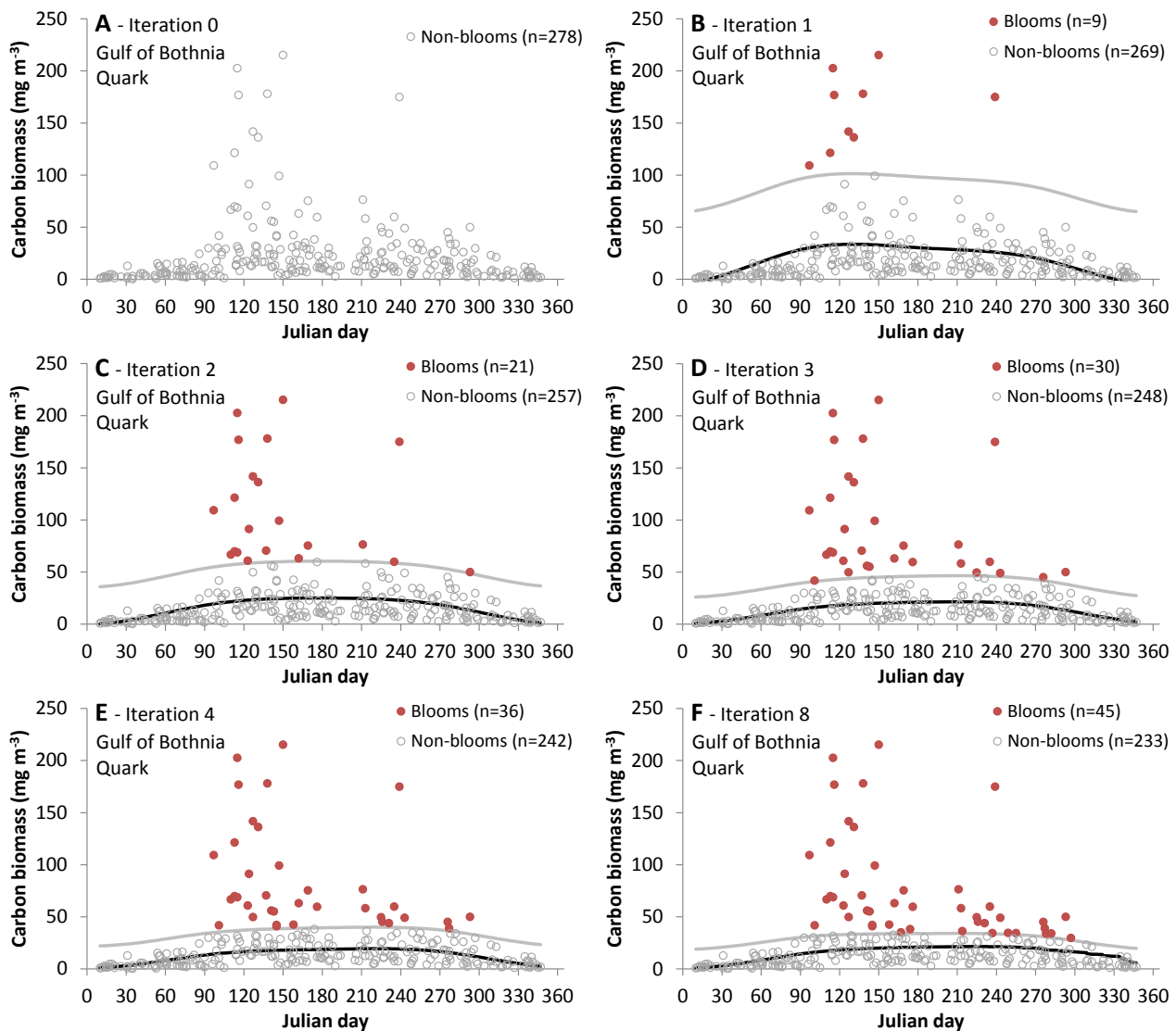
<Fig. 6>
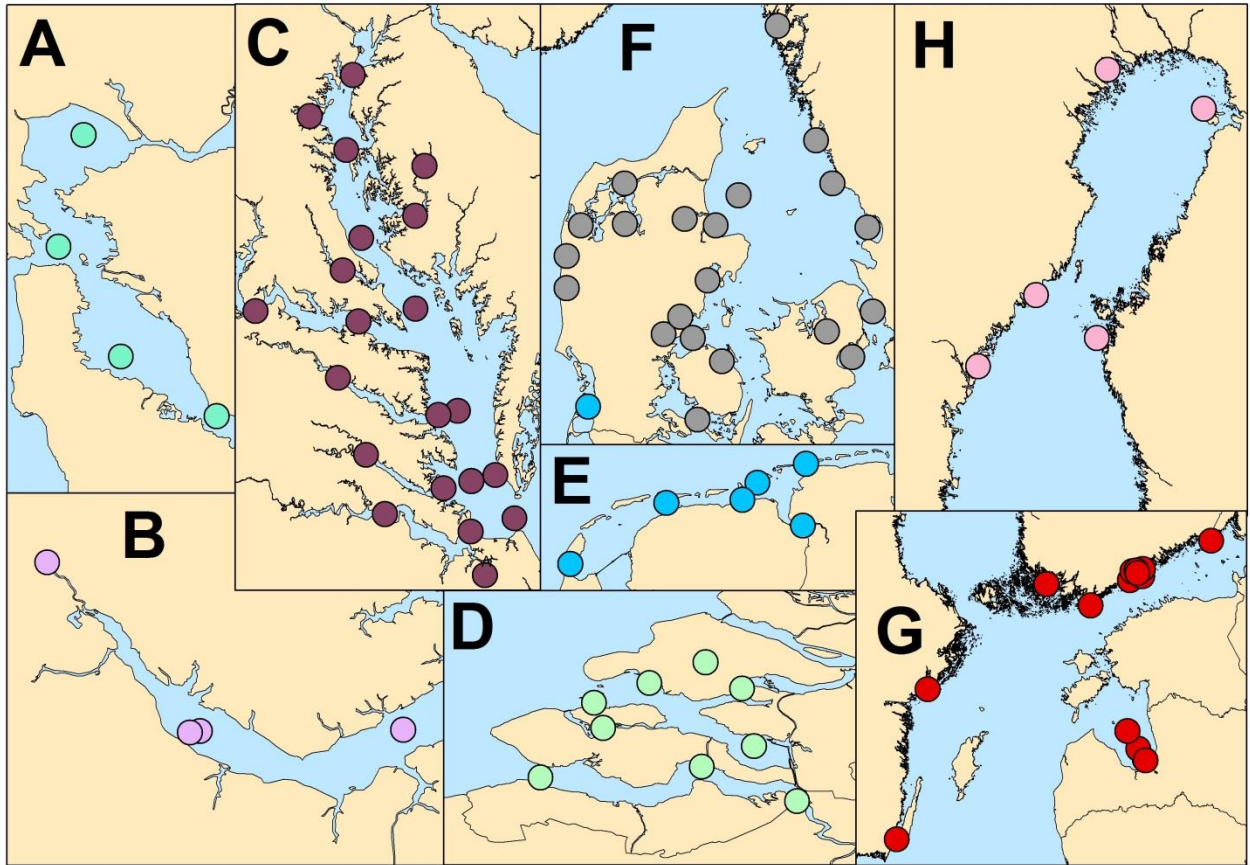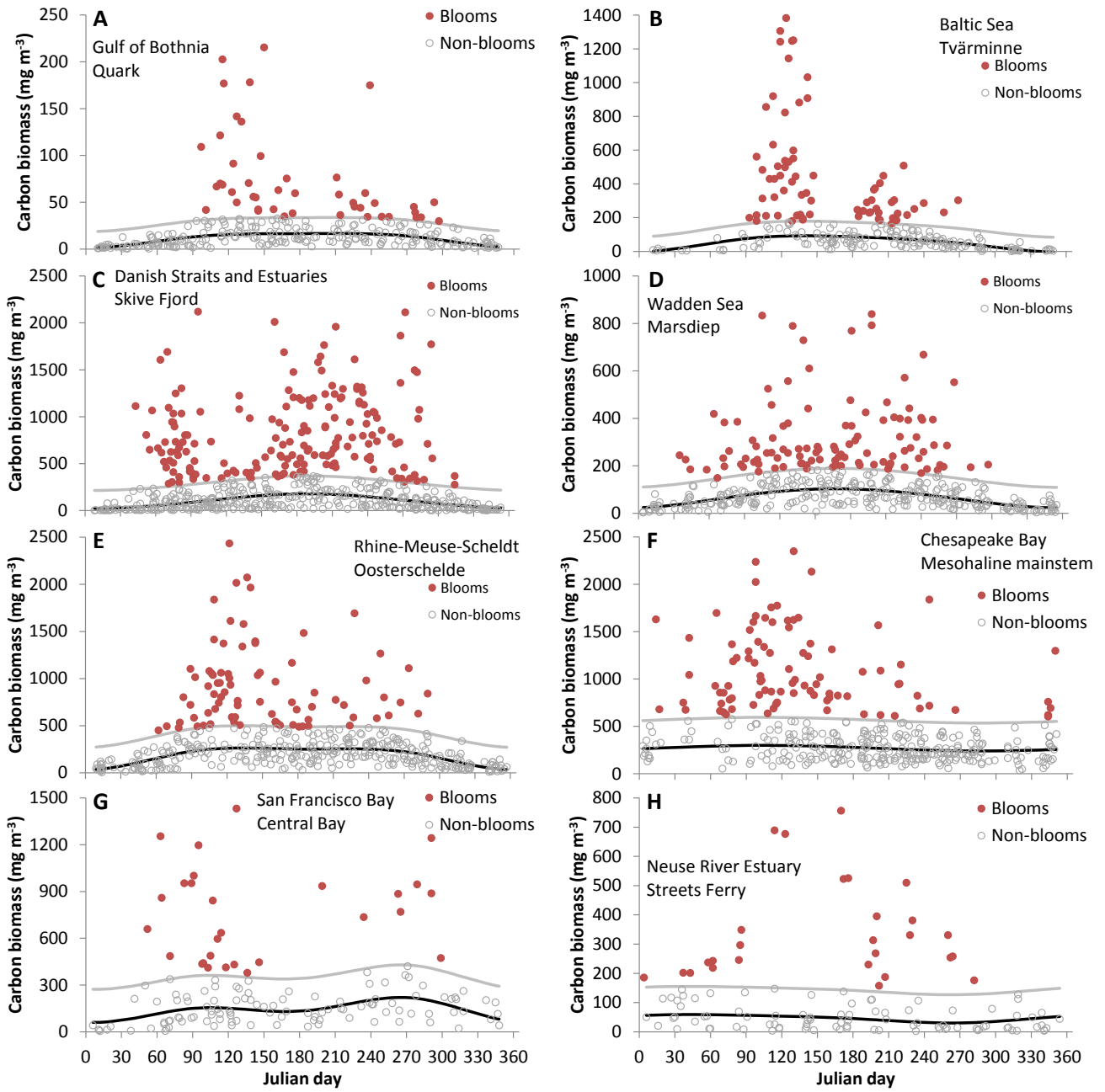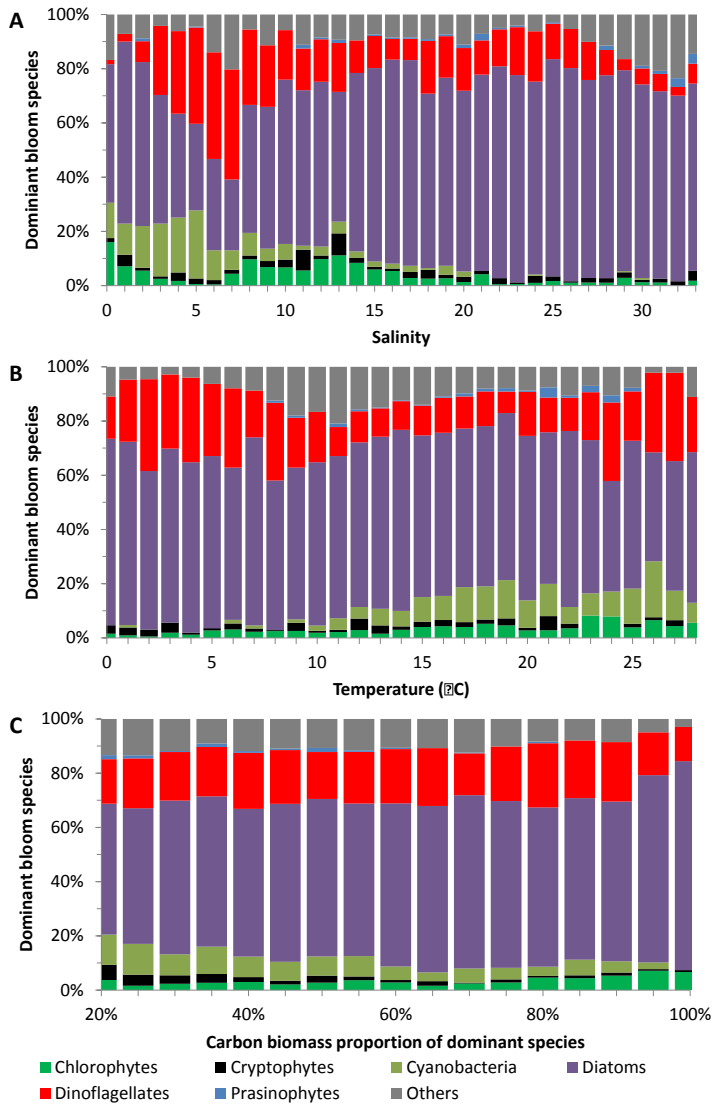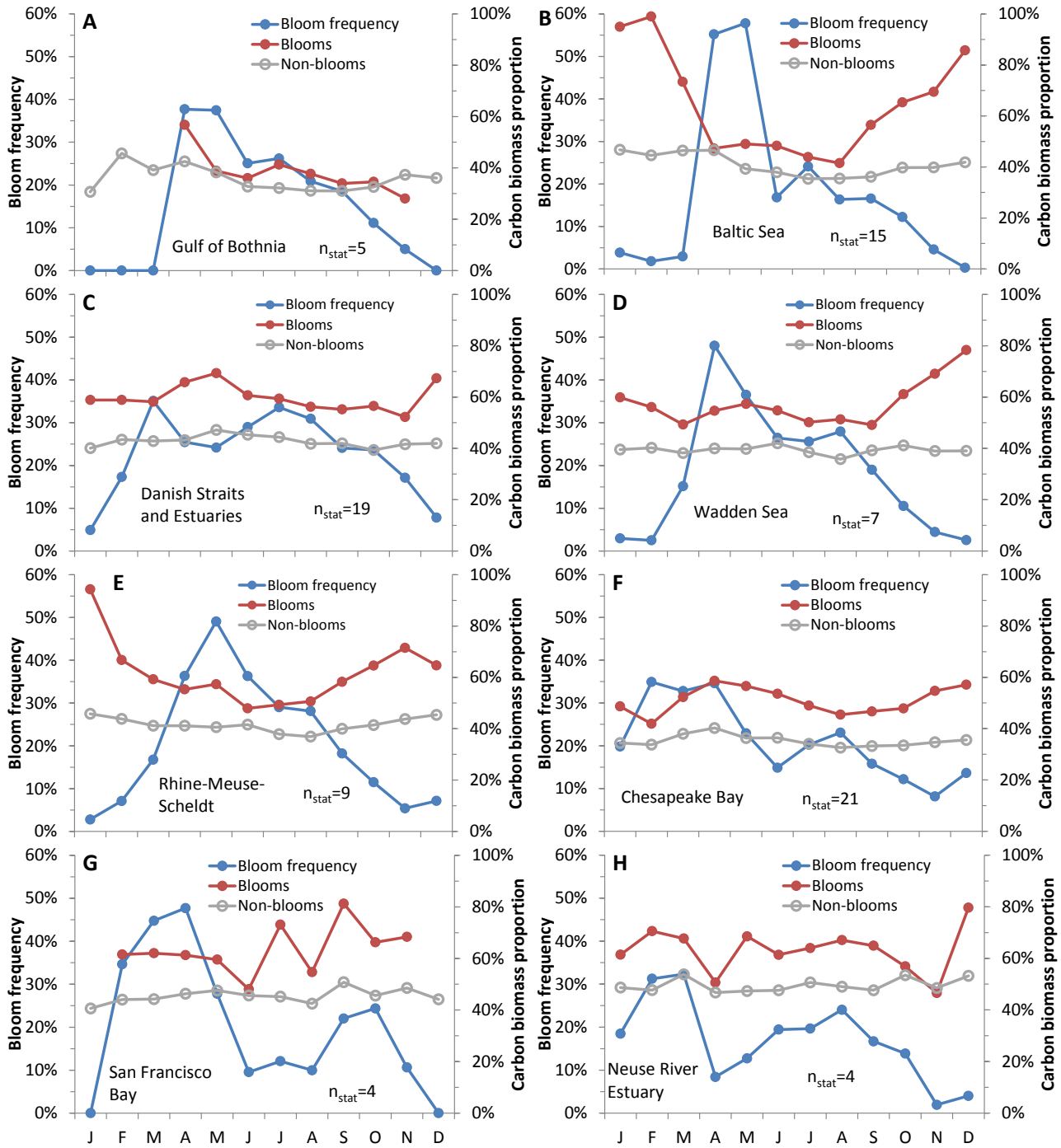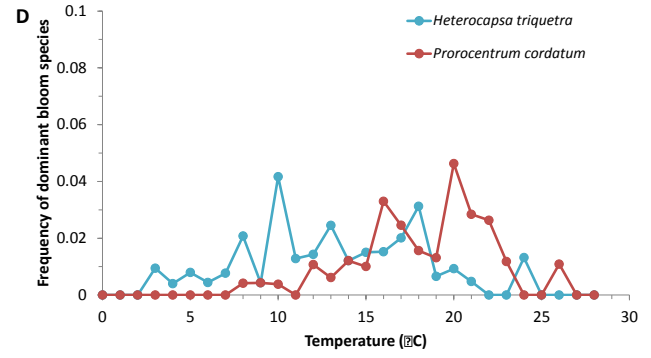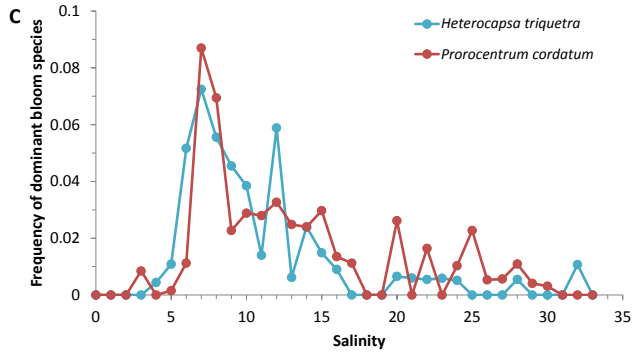
Supplementary Information

Fig S1: Illustration of the bloom identification algorithm. A) Initialisation of the algorithm with all observations categorised as blooms. B) First iteration where the periodic spline has been fit to non-bloom observations from the previous iteration and observations exceeding the 99th percentile of the prediction interval are categorised as blooms. C), D) and E) show second, third and fourth iteration. F) Final iteration, where the algorithm converges and no more non-bloom observations are categorised as blooms. The numbers of bloom and non-bloom observations in each iteration are shown in the symbol legend.

Table 3: The five most frequent dominant bloom species (taxonomical unit) in each region with the proportion of bloom observations dominated. The full list of bloom dominant species for all regions is given in Table S1.

Baltic Sea	Chesapeake Bay	Danish Straits and Estuaries	Gulf of Bothnia
Peridiniella catenata (19.8%)	*Cerataulina pelagica* (18.8%)	*Skeletonema costatum* s.l. (21.8%)	*Diatoma tenuis* (23.2%)
Scrippsiella hangoei (9.34 %)	*Coscinodiscus spp.* (11.4%)	*Dactyliosolen fragilissimus* (6.60%)	*Peridiniella catenata* (14.5%)
Planktothrix agardhii (8.11%)	*Skeletonema costatum* s.l. (5.64%)	*Cerataulina pelagica* (4.57%)	*Aphanizomenon spp.* (7.97%)
Achnanthes taeniata (8.11%)	*Cyclotella spp.* (5.24%)	*Neoceratium tripos* (3.97%)	*Synechococcus spp.* (7.97%)
Skeletonema costatum s.l. (5.78%)	*Dactyliosolen fragilissimus* (4.31%)	*Pseudo-nitzschia delicatissima* (3.92%)	*Chaetoceros spp.* (5.80%)

Neuse River Estuary	Rhine-Meuse-Scheldt delta	San Francisco Bay	Wadden Sea
Coscinodiscus spp. (33.0%)	*Odontella sinensis* (13.7%)	*Thalassiosira punctigera* (16.3%)	*Phaeocystis spp.* (13.8%)
Heterocapsa triquetra (15.1%)	*Chlorophyceae* (11.7%)	*Ditylum brightwellii* (7.76%)	*Odontella sinensis* (10.7%)
Trachelomonas spp. (14.2%)	*Unidentified* (10.85%)	*Coscinodiscus oculus-iridis* (7.35%)	*Odontella regia* (10.0%)
Gymnodinium instriatum (8.49%)	*Phaeocystis spp.* (8.97%)	*Thalassiosira rotula* (7.35%)	*Mediopyxis helysia* (8.03%)
Thalassiosira nordenskioeldii (5.66%)	*Thalassiosira spp.* (5.43%)	*Skeletonema costatum* s.l. (6.53%)	*Rhizosolenia imbricata* (5.35%)