Manuscript prepared for Biogeosciences
with version 2015/11/06 7.99 Copernicus papers of the LaTeX class copernicus.cls.
Date: 29 June 2016

Spring Blooms in the Baltic Sea have weakened but lengthened from 2000 to 2014

Philipp M. M. Groetsch[1, 2], Stefan G. H. Simis[3, 4], Marieke A. Eleveld[5, 1], and Steef W. M. Peters[2]

[1]Institute for Environmental Studies (IVM). De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands
[2]Water Insight. Marijkeweg 22, 6709 PG Wageningen, The Netherlands
[3]Plymouth Marine Laboratory. Prospect Place, The Hoe, PL1 3DH Plymouth, United Kingdom
[4]Finnish Environment Institute SYKE. Erik Palménin Aukio 1, 00560 Helsinki, Finland
[5]Deltares. P.O. Box 177, 2600 MH Delft, The Netherlands

Correspondence to: Philipp M. M. Groetsch (groetsch@waterinsight.nl)

Abstract. Phytoplankton spring bloom phenology was derived from a 15-year time-series (2000-2014) of ship-of-opportunity chlorophyll-*a* fluorescence observations collected in the Baltic Sea through the Alg@line network. Decadal trends were analysed against inter-annual variability in bloom timing and intensity, and environmental drivers (nutrient concentration, temperature, radiation level, wind speed).

Spring blooms developed from the south to the north with the first blooms peaking mid-March in the Bay of Mecklenburg and the latest bloom peaks occurring mid-April in the Gulf of Finland. Bloom duration was similar between sea areas (43 ± 2 d), except for shorter bloom duration in the Bay of Mecklenburg (36 ± 11 d). Variability in bloom timing increased towards the south. Bloom peak chlorophyll-*a* concentrations were highest (and most variable) in the Gulf of Finland (20.2 ± 5.7 mg m^{-3}) and the Bay of Mecklenburg (12.3 ± 5.2 mg m^{-3}).

Bloom peak chlorophyll-*a* concentration showed a negative trend of -0.31 ± 0.10 mg m^{-3} yr^{-1}. Trend-agnostic distribution-based (Weibull-type) bloom metrics showed a positive trend in bloom duration of 1.04 ± 0.20 d yr^{-1}, which was not found with any of the threshold-based metrics. The Weibull bloom metric results were considered representative in presence of bloom intensity trends.

Bloom intensity was mainly determined by winter nutrient concentration, while bloom timing and duration co-varied with meteorological conditions. Longer blooms corresponded to higher water temperature, more intense solar radiation, and lower wind speed. It is concluded that nutrient reduction efforts led to decreasing bloom intensity, while changes in Baltic Sea environmental conditions associated with global change correspond to a lengthening spring bloom period.

1 Introduction

Human influence and climate change transform terrestrial and marine ecosystems worldwide at unprecedented rates (Cleland et al., 2007; Cloern et al., 2015). Coastal marine systems experience anthropogenic pressure as well as indirect changes in climatic conditions, which affect the marine food-web (Heisler et al., 2008; Zhai et al., 2013; Paerl and Huisman, 2008). Ecosystem responses to these changes are difficult to relate to unique causes (HELCOM, 2007b; Winder and Cloern, 2010; Neumann et al., 2012). Experiments designed to support biogeochemical model scenarios (e.g. Neumann et al., 2002; Tamminen and Andersen, 2007; Seppälä and Olli, 2008) help to disentangle observed trends. However, the predictive capabilities of biogeochemical models (e.g Kuusisto et al., 1998; Meier et al., 2011; Gnanadesikan and Anderson, 2009) remain dependent on calibration against long and consistent multi-variable time series.

Phytoplankton bloom intensity and timing (bloom phenology) are indicators for ecosystem health at the base of the food web (e.g. Hays et al., 2005; Adrian et al., 2009; Vargas et al., 2009). Phenological studies are increasingly used to inspect regional ecosystem response to nutrient reduction efforts (HELCOM, 2007a; Voss et al., 2011; Fleming-Lehtinen et al., 2015) and changing climatic conditions (Sommer and Lengfellner, 2008; Paerl and Huisman, 2009). The Baltic Sea is a coastal ecosystem affected by eutrophication (Korpinen et al., 2012), which intensifies naturally occurring spring- and summer bloom (Bianchi and Engelhaupt, 2000; HELCOM, 2007a). The Helsinki Commission formulated a nutrient reduction scheme aimed at improving ecosystem health in 1992 (HELCOM, 2008), which entered into force in 2000. Monitoring of key ecosystem health indicators is implemented in the national monitoring programmes of HELCOM contracting parties. These programmes include traditional dedicated sampling campaigns at sea and increasingly the use of highly resolving observation platforms.

Ships-of-opportunity (typically cargo ships or passenger ferries) offer a largely weather-independent, reliable, and cost-effective platform for the collection of high frequency in situ observations (Leppänen et al., 1995; Ainsworth, 2008). Phytoplankton pigment fluorometers are included in most of these ferryboxes. In the Baltic sea, such systems have recorded phytoplankton blooms on the route from Helsinki to Travemünde (v.v.) since 1992 (Rantajärvi et al., 2003). On this route, ferryboxes have collected over 9.5 million chlorophyll-*a* pigment fluorescence observations from 1926 transects with a median revisit time of under two days in the last 15 years (2000-2014). Ship-based observations from merchant vessels provide continuity in monitoring, which is particularly important in seasons when other observation systems are less reliable. In spring, satellite observations are rare due to high average cloud cover, while high costs of dedicated research cruises and coastal laboratories limit their spatio temporal coverage. Ferrybox observations are therefore the primary source of observations to study spring bloom dynamics in this region.

Phytoplankton abundance and succession in the Baltic Sea is controlled by nutrient (Neumann et al., 2002; Tamminen and Andersen, 2007) and light availability (Sverdrup, 1953; Smetacek and

Passow, 1990; Nelson and Smith, 1991; Siegel et al., 2002), mixing-status (Ueyama and Monger, 2005; Sharples et al., 2006), temperature (Grayek and Staneva, 2011), ice cover (Kahru and Nommann, 1990; Omstedt et al., 2004; Sommer and Lengfellner, 2008), and salinity (Fennel, 1999; Tamminen and Andersen, 2007). In addition, the quantum yield of fluorescence is influenced by solar irradiance (Kiefer, 1973; Dandonneau and Neveux, 1997; Marra, 1997; Sackmann et al., 2008), species composition, and physiology (Kiefer et al., 1989). Hence, interpretation of unattended pigment fluorescence measurements in terms of phytoplankton biomass presents a number of challenges (Roesler and Barnard, 2013). Firstly, phytoplankton distribution exhibits high spatial and temporal variability, while ferryboxes measure pigment fluorescence at fixed depth (Ruokanen et al., 2003). Therefore, stratified conditions may not be well represented in the data (Groetsch et al., 2014). Secondly, in a typical ferrybox setup fluorescence yield is at best determined as a daily regional average, which disregards variability on smaller spatio-temporal scales. Despite these challenges, Fleming and Kaitala (2006) demonstrated that ferrybox observations in the Baltic Sea can be used to derive bloom timing and intensity for biomass-rich sea areas. They report a slightly negative trend in bloom initiation in the Northern Baltic Proper and the Gulf of Finland for the period 1992-2004. Recent studies also reported shifts in phytoplankton spring bloom biomass or species composition (e.g. Klais et al., 2011; Wasmund et al., 2011, 2013). Kahru and Elmgren (2014) reported that the timing of cyanobacterial surface accumulations has advanced approximately 20 days from 1979 to 2013. However, information about shifts in Baltic Sea spring bloom timing are still lacking.

Choosing an adequate bloom metric is not trivial as no clear guidelines exist that conclusively support one metric over others. Bloom metrics for both remotely sensed and in situ sampled time series are commonly divided into three groups: 1) fixed or variable concentration threshold metrics (Siegel et al., 2002; Fleming and Kaitala, 2006; Lips et al., 2014; Racault et al., 2015), 2) growth-rate-based metrics (Rolinski et al., 2007; Wiltshire et al., 2008), and 3) distribution-based metrics (Rolinski et al., 2007; Platt et al., 2009; Vargas et al., 2009; Zhai et al., 2011). Threshold- and growth-rate based metrics typically require data pre-processing (e.g. interpolation and smoothing), to mitigate the impact of gaps, noise, outliers, and multi-modal bloom distributions on the derived bloom phenology (Rolinski et al., 2007; Cole et al., 2012; Ferreira et al., 2014). Distribution-based metrics fit an analytical expression to observations using fitting routines designed to cope with imperfections in the input data while optimally preserving natural variability. Distribution-based bloom metrics are considered more robust than threshold- or growth-rate-based metrics, in the presence of complex, multi-modal bloom observations (Ji et al., 2010). Interpretation based on several, conceptually different bloom metrics can be used to obtain uncertainty estimates (Ho and Michalak, 2015). It also allows to screen for long-term trends in bloom phenology. The latter is because threshold-based metrics are biased by long-term bloom intensity trends, whereas growth-rate and distribution-based metrics are not. Figure 1 illustrates how a gradual decline (negative trend) in bloom peak concentration causes any metric based on fixed thresholds (e.g. derived from climatology or expert-judgement)

95 to introduce an artificial negative trend in bloom duration. In contrast, metrics based on growth rate, distribution, or annually derived thresholds yield a single bloom duration in this example because bloom intensity does not influence these metrics.

The aims of this study are twofold: (1) to report long-term trends for Baltic Sea spring bloom intensity and timing, and (2) to attribute these trends to changes in environmental conditions. To meet

100 these objectives, we describe a methodology to derive quality controlled time-series of chlorophyll-a concentrations from observations collected under the Baltic Sea Alg@line program over a period of 15 years (2000-2014). Uncertainties arising from variability in the phytoplankton pigment fluorescence yield are estimated. Bloom phenology parameters, derived from threshold- and distribution-based bloom metrics, are explored for long-term trends. Inter-annual variability of bloom phenology

105 parameters are attributed to nutrient availability and meteorological conditions (temperature, radiation level, wind speed), which might help to relate long-term trends to unique causes. Finally, we summarize how these results contribute to the discussion on recent changes in the Baltic Sea, and the monitoring practices that need to be in place to detect such changes.

2 Materials and Methods

110 ### 2.1 Alg@line Data

In situ data in this study were collected until 2009 by the Finnish Institute of Marine Research, and by the Finnish Environment Institute (SYKE) from 2009 onwards, within the Alg@line network of Baltic Sea ferryboxes. Here we consider systems installed on two cargo vessels, M/S *Finnpartner* (2000-2006) and M/S *Finnmaid* (2007-2014), which served between Travemünde (Germany) and

115 Helsinki (Finland) as depicted in Fig. 2. Three routes were sailed during the study period. Depending on weather conditions the passage between Gotland and the mainland of Sweden (39 % of all transects) was favoured over the direct route east of Gotland (52 %), while the route with a lay-over in Gdansk (Poland) was only occasionally served during 2009 to 2012 (7 %). Several transects (2 %) were sailed for refuelling or maintenance in other ports and not used for this study.

120 Details on the instrumentation of the Alg@line ferrybox systems can be found in Leppänen et al. (1994); Rantajärvi et al. (2003); Ruokanen et al. (2003); Seppälä et al. (2007). In summary, the systems record in vivo fluorescence of chlorophyll-a (chla), salinity and temperature throughout the studied period (2000-2014). Turbidity and (in summer) phycocyanin pigment fluorescence were recorded from 2005 onwards and are not used here. At cruising speed (20 kn to 23 kn) the sampling

125 interval of 20 s resulted in a nominal spatial resolution of 200 m.

Quality control flags were derived from (1) sensor reading thresholds on speed, flow rate, hull and sampled water temperature, and (2) data variability, expressed as lower and upper bounds for standard deviation between neighbouring measurements, as described below. Measurements at low (< 5 kn) or zero ship speed are typically collected in harbour and were omitted. Erroneous records,

130   e.g. caused by instrument communication errors, were removed using a moving window mean filter. A window length of 25 observations (approximately $8.3\,\mathrm{min}$) was used for records of ship speed, and a window length of 100 observations ($33.3\,\mathrm{min}$) was used for flow rate and temperature records. Low flow rates can indicate blocked passages, pump failure, or leaks. Flow meter readings were available for approximately one-third of all records. A proxy for flow disruption is the difference in ship-hull

135   temperature and in-line temperature. Flow rates $< 0.3\,\mathrm{L\,min^{-1}}$ or a temperature difference $> 2\,°\mathrm{C}$ were used to flag records as suspect. Instrument failure, communication and digitizing errors may lead to 'stuck' values, which were detected by calculating standard deviation in a moving window of 100 samples. Observations corresponding to low standard deviation ($\sigma < 1e^{-4}$) of chla fluorescence measurements or GPS-derived latitude were omitted. GPS-derived latitude was additionally filtered

140   for exceptionally high short-term variability ($\sigma > 0.5$, window size 50 samples), caused by poor satellite reception or serial communication errors. Table 1 provides an overview of the applied quality control flags.

     Chla fluorescence data were corrected for sensor drift and discontinuities by transect-wise normalization (division by transect mean). This was necessary to account for changes in instrumentation,

145   signal contamination due to bio-fouling, trapped bubbles and particles, and changes in sensor sensitivity due to deterioration or manual adjustments. Laboratory analysis results of bottle samples are typically available from every 6th transect, with up to 24 samples collected by automated, refrigerated water samplers (Teledyne Isco). Laboratory analyses included inorganic nutrient concentrations (nitrate+nitrite, phosphate and silicate), chla concentration, and occasionally inverted light

150   microscopy counts of phytoplankton species. Laboratory chla concentration results were used to convert transect-normalized chla fluorescence to units of chla concentration (in $\mathrm{mg\,m^{-3}}$). First, a linear (generalized least squares) regression fit of normalized chla fluorescence against corresponding chla lab measurements was carried out for each sampled transect. If the regression failed ($R^2 < 0.3$ or $p > 1$) a moving window regression was carried out (window length 10 samples) and the subset with

155   the highest R^2 was used to determine the correction factor. The threshold for R^2 was determined manually based on the distribution of R^2, while $p > 1$ indicates numerical instabilities during the fitting procedure. Each transect without corresponding bottle samples was corrected by individually applying the regression parameters of the two neighbouring sampled transects. These two solutions were then interpolated linearly, weighted by their temporal distance to the respective transect. Nega-

160   tive concentration values occasionally occurred for weak fluorescence signals, and were set to zero.

     The diurnal variability of the fluorescence signal was estimated from quality-controlled observations in all seasons. First, these observations were divided by their respective transect mean to remove biomass-driven first-order variability in the fluorescence signal. Then, diurnal cycles were derived by dividing these observations into hourly bins and sun elevation angle ranges (0.1 rad bins).

2.2 Meteorological Data

Photosynthetically active radiation (PAR), sea surface temperature (SST) and wind speed (WIND) were derived from the ECMWF ERA-Interim reanalysis data set (Dee et al., 2011). The spatial resolution of the model is constrained by the underlying atmospheric model, which is stored on a spatial T255 grid corresponding to approximately 79 km cell size when projected to a reduced Gaussian grid. Four values per day were retrieved for each parameter and the entire Baltic Sea. Parameter values for each Alg@line observation were extracted using spatio-temporal spline interpolation of third order. The first order seasonal signal (e.g. rising PAR and SST in spring) was removed from the observations by subtracting multi-year (2000-2014) daily sea area averages, approximated by second order polynomials. The seasonally detrended parameters were then averaged over the bloom period and are further referred to as PAR, SST, and WIND.

2.3 Nutrient Concentration and Depletion Timing

A single term for nutrient availability was adopted from Fleming and Kaitala (2006), calculated as $\text{NUT} = \sqrt[3]{(NO_3 + NO_2) \times PO_4 \times SiO_4}$, where $NO_3 + NO_2$, PO_4 and SiO_4 are the concentrations of nitrite+nitrate, phosphate, and silicate, respectively. These concentrations were derived from laboratory analysis of bottle samples that were regularly collected along the transect (further detail in section 2.1). NUT was spatially binned for each investigated sea area and re-sampled to daily averages and consecutively smoothed with a 21-day centred-running-mean filter. This treatment resembles the processing of Alg@aline observations (see section 2.4) to enable consistent interpretation of the joint data set. Nutrient concentrations and depletion timing are described using the following metrics. The nutrient concentration prior to bloom start (NUT-PEAKVALUE) was defined as the yearly maximum nutrient concentration (day-of-year between 31 and 160). The day-of-year when the nutrient concentrations equalled 100 %, 50 %, and 25 % of their peak values are referred as NUT-PEAKDAY, NUT-DEPLDAY-50, and NUT-DEPLDAY-25. The day and value of the lowest nutrient concentration index are referred to as NUT-MINDAY and NUT-MINVALUE. The rate of nutrient depletion between 75 % and 25 % of the peak value (NUT-SLOPE) was determined through linear regression.

2.4 Extraction of Bloom Timing and Intensity

Extraction of bloom timing and intensity was carried out for five Baltic Sea areas, where each area follows definitions of the HELCOM Combine program (HELCOM, 2013). Figure 2 illustrates the location of the areas: the Western Gulf of Finland (GOF: >59.5 °N latitude, along-transect), the Northern Baltic Proper (NBP: 58.4-59.5 °N latitude, along-transect), the combined Western and Eastern Gotland basins (GOT: 56.2-58.4 °N latitude, along-transect), the Southern Baltic Proper (SBP: 54.5-56.2 °N latitude, along-transect), and the Bay of Mecklenburg (BOM: <54.5 °N latitude, along-transect). For the GOT and SBP areas only routes that passed by Gotland were selected whereas routes

via Gdansk were excluded. This is because the route through Gdansk was sailed only from 2009 to 2012. If not otherwise stated, all further steps are carried out individually for each of these areas and for day-of-year between 31 (31 January) and 160 (9 June). The ship-of-oppportunity (Alg@line) measurements typically commenced in the second half of January, which is why 31 January was chosen as the start of our analysis. The end date was chosen such that it covers all spring bloom events in all basins but excludes summer bloom.

Alg@line chla concentrations (see section 2.1) were resampled to daily sea area averages, using linear interpolation, and subsequently smoothed with a 21-day centred-running-mean filter (e.g. Ferreira et al., 2014; Racault et al., 2015) to fill in gaps and reduce short-term variability. We derive several metrics, all of which have in common that the bloom peak concentration (PEAKHEIGHT, see Table 2 for explanations of acronyms) and timing (PEAKDAY) are defined as the maximum chla value at the corresponding day-of-year, respectively. Two threshold-based metrics and one distribution-fit-based metric were calculated:

1) Chla concentration exceeding a fixed-threshold of $5\,\mathrm{mg\,m^{-3}}$ was defined as bloom by Fleming and Kaitala (2006), further referred to as const5. A 21-day centred-running-mean filter was used to keep results comparable to the other metrics considered, whereas Fleming and Kaitala (2006) used a 7-day centred-running-median filter.

2) Siegel et al. (2002) proposed a spatially variable-threshold metric based on the 5 %-above-median concentration, but reported small quantitative differences for thresholds between 1 and 30 %-above-median. Their threshold is based on the complete annual cycle, while here only the spring bloom period from day-of-year 31 to 160 is considered. We refer to this metric as median5.

3) Distributions proposed to describe bloom phenology include shifted-Gaussian (Platt et al., 2009), Gamma (Vargas et al., 2009), and Weibull distributions (Rolinski et al., 2007). The shifted Gaussian is symmetric in shape, whereas Gamma distributions allow for different slopes of bloom rise and decline. In addition, Weibull functions recognize non-zero offsets before and after the bloom phase. The latter has proven essential to obtain a good fit for the transition phase between spring and summer bloom with the here analysed data set. A modified Weibull-function, as proposed by Rolinski et al. (2007), was fitted non-linearly to the preprocessed and scaled (to a range of zero to one) chla concentrations. The bloom initiation and end are defined as the 10th and 90th percentiles before and after the bloom peak, respectively. This metric is further referred to as weibull.

For each metric, bloom initiation, peak, and end dates (STARTDAY, PEAKDAY, and ENDDAY) were extracted from the data set. Based on these dates, bloom duration (DURATION), concentration average (CONCAVG), and the sum of daily chla concentrations (BLOOMIDX) were calculated. The latter was proposed by Fleming and Kaitala (2006) to characterize bloom intensity. We assumed the bloom to have started prior to Alg@line service commence if the first data point already satisfied the bloom criterion for a given metric. Such cases were identified for 30 out of 225 combinations of sea region, year, and bloom metric (9 times for bloom metric const5, 16 times for median5, and

5 times for `weibull`). Corresponding bloom start days were replaced by the median value for the region over the 15 years studied in all subsequent calculations.

2.5 Principal Component Analysis

Principal component analysis (PCA) was carried out to attribute seasonally detrended meteorological conditions (SST, PAR, WIND) and nutrient concentrations (NUT-PEAKVALUE, NUT-MINVALUE) to the inter-annual variability in bloom intensity (BLOOMIDX, CONCAVG, PEAKHEIGHT) and timing (STARTDAY and PEAKDAY, DURATION). Outliers were defined for each parameter as departure by more than 3 standard deviations from the parameter mean, and replaced with the region-median. Z-score normalization (subtraction of mean, division by standard deviation) was carried out on a per-region basis. Region-equalized, zero-mean and unit-variance data were then subjected to the PCA function in the python framework scikit-learn (Pedregosa and Varoquaux, 2011).

3 Results

3.1 Quality-controlled Chlorophyll-*a* Concentration Time Series

The Alg@line ferrybox systems collected over 9.5×10^6 observations between 2000 and 2014, of which 3.8×10^6 observations were sampled during spring (day-of-year 31 to 160). Availability and rejection rates for each quality control parameter are listed in Table 1. In total, quality control procedures removed 4.55 % of all observations.

Determination of the fluorescence yield was supported by an 'adaptive regression' method. Where necessary ($R^2 < 0.3$ or $p > 1$), it selected the subset of bottle-sampled and laboratory-analysed chla concentrations that yielded the best linear fit to chla fluorescence observations for a given transect. This procedure allowed to successfully fit 318 (98 %) out of 324 transects for which bottle samples were collected. Only 266 (82 %) transects could have been used ($R^2 >= 0.3$ and $p \ll 1$) without applying this technique.

Figure 3A shows normalized fluorescence observations as a function of sampling time-of-day. Results are presented separately for summer (May to August), winter (November to February) and the transition periods (autumn, spring). Diurnal variability was most pronounced in summer, when the fluorescence signal varied on average 50 % over the course of a day. In winter and during the transition periods (spring, autumn) a diurnal variability of 35 and 38 %, respectively, was contained in the fluorescence signals. This seasonal effect is likely caused by variations in average irradiance intensity, which are modulated primarily by sun elevation, but also by atmospheric conditions (e.g. cloud cover, aerosol optical thickness) and optical properties of the water body (e.g. ice cover, attenuation). Figure 3B depicts normalized fluorescence as a function of solar elevation. In this representation seasonal differences in diurnal variability are essentially absent and the correspondence between solar elevation and average fluorescence response was approximately linear for daytime observations.

3.2 Bloom Intensity and Timing

Blooms generally developed first in the south and progressed towards the north (see Fig. 4 and Table 3). Bloom peak timing (not influenced by choice of metric) followed this pattern, as did metric-dependent bloom start and end dates. The fixed-threshold bloom metric `const5` suggested longer blooms in high-biomass sea areas like the GOF, compared to low-biomass areas such as the SBS. The spatially variable-threshold metric `median5` applies area-specific bloom thresholds (NBP: $3.52 \, \mathrm{mg \, m^{-3}}$, GOF: $4.95 \, \mathrm{mg \, m^{-3}}$, GOT: $2.51 \, \mathrm{mg \, m^{-3}}$, SBS: $2.62 \, \mathrm{mg \, m^{-3}}$, BOM: $4.02 \, \mathrm{mg \, m^{-3}}$) and resulted in approximately stable bloom duration in all sea areas. The `weibull` metric, which is not sensitive to absolute bloom intensity, also resulted in comparable bloom durations for all sea areas. The year-to-year variability of start, peak, and end days generally increased towards the south for all metrics.

Spring bloom intensity was described by three parameters: the metric-independent bloom peak concentration (PEAKHEIGHT), the chla concentration average during bloom conditions (CONCAVG), and the sum of daily chla concentrations over the bloom period (BLOOMIDX). Similar patterns were observed for all these parameters and bloom metrics, as illustrated in Fig. 5. The highest bloom intensity was found in the GOF and NBP, followed by the BOM. Low-intensity blooms were observed in the SBP and the GOT. Variability was generally proportional to bloom intensity, highest in the high-biomass and coastal GOF and BOM. Variability in BLOOMIDX was comparable to that in PEAKHEIGHT, while CONCAVG was considerably more stable. All calculated bloom phenology parameters can be found in the supplementary material.

3.3 Trends

Figure 6 shows normalized (subtraction of area-average concentration) CONCAVG and PEAKHEIGHT for all sea areas combined, as a function of bloom year. PEAKHEIGHT is independent of bloom metric and shows a highly significant ($R^2 = 0.12, p \ll 0.01$) negative trend of $-0.30 \pm 0.10 \, \mathrm{mg \, m^{-3} \, yr^{-1}}$. CONCAVG is dependent on bloom start and end days and was therefore calculated for all applied metrics. Statistically significant, negative trends resulted from all metrics: $-0.12 \pm 0.04 \, \mathrm{mg \, m^{-3} \, yr^{-1}}$ for `const5` ($R^2 = 0.11, p \ll 0.01$), $-0.11 \pm 0.05 \, \mathrm{mg \, m^{-3} \, yr^{-1}}$ for `median5` ($R^2 = 0.12, p < 0.05$), and $-0.22 \pm 0.07 \, \mathrm{mg \, m^{-3} \, yr^{-1}}$ for `weibull` ($R^2 = 0.11, p \ll 0.01$).

No significant trends were found for BLOOMIDX, STARTDAY, and PEAKDAY with any of the applied metrics, while ENDDAY showed weakly correlated but statistically significant ($R^2 = 0.06, 0.08, p < 0.05$) positive trends for `const5` and `weibull` with slopes $0.6, 0.7 \pm 0.3 \, \mathrm{d \, yr^{-1}}$, respectively.

Bloom duration resulting from the `weibull` metric stands out in the result set with a positive trend of $1.04 \pm 0.20 \, \mathrm{d \, yr^{-1}}$ ($R^2 = 0.28, p \ll 0.01$, Fig. 7). No significant trend in bloom duration was found for any fixed- or variable-threshold metric.

305   Peak nutrient concentrations showed no significant trend, in contrast to post-bloom nutrient concentrations with a highly significant, negative trend $-0.020 \pm 0.004 \, \mu\text{mol L}^{-1} \, \text{yr}^{-1}$ ($R^2 = 0.23, p \ll 0.01$). Peak nutrient concentration timing shifted to earlier dates ($-0.7 \pm 0.3 \, \text{d yr}^{-1}$ ($R^2 = 0.06, p < 0.05$)), while the 25 %-of-peak-value was reached progressively later ($0.67 \pm 0.31 \, \text{d yr}^{-1}$, ($R^2 = 0.06, p < 0.05$)). No significant trends were found for nutrient depletion slope, 50 %-of-peak-value-
310   timing, or day of minimal nutrient concentrations.

3.4 Inter-annual Variability

Pre-bloom nutrient concentrations were positively correlated to bloom peak height (no normalization, $R^2 = 0.39, p \ll 0.01$) and concentration average (no normalization, $R^2 = 0.37 - 0.57, p \ll 0.01$, depending on metric). After applying area-wise mean and variance (z-score) normalization,
315   however, a negative correlation was found for PEAKHEIGHT ($R^2 = 0.11, p \ll 0.01$, metric independent) and CONCAVG ($R^2 = 0.12, 0.11, p \ll 0.01$ for `const5` and `weibull`, respectively).

The timing of nutrient depletion, specifically NUT-DEPLDAY-50, was positively correlated to the bloom peak day ($R^2 = 0.47, p \ll 0.01$), and to bloom-averaged, detrended PAR-levels ($R^2 = 0.14 - 0.29, p \ll 0.01$). Average wind speed and PAR were negatively correlated during bloom conditions
320   ($R^2 = 0.10 - 0.23, p \ll 0.01$). The bloom timing parameters (STARTDAY, PEAKDAY, ENDDAY) were weakly but statistically significantly inter-correlated (results not shown).

PCA scores and loadings of the first three principal components (PC) are shown as biplots in Fig. 8. The first PC is dominated by negative correlations to bloom intensity parameters (PEAKHEIGHT, CONCAVG, BLOOMIDX). This component is positively correlated to pre-bloom nutrient concentra-
325   tion (NUT-PEAKVALUE) and bloom duration, illustrating that bloom intensity is driven by pre-bloom nutrient availability. The second PC is linked to bloom timing, with strong positive correlations to STARTDAY and PEAKDAY. Correlations to PAR (positive), SST (positive), and WIND (negative) suggest that weather conditions affect bloom timing. Bloom duration is positively correlated to the third PC, as well as to BLOOMIDX. Additional negative correlations to NUT-MINVALUE and WIND, as
330   well as a positive correlation to PAR, suggest a link between favourable meteorological conditions (low wind-mixing, high light level) and efficient nutrient depletion.

4 Discussion

Trends in spring bloom phenology can be interpreted as responses to nutrient reduction as well as to slowly acting environmental processes, such as climate change. To disentangle or even quantify
335   these trends, suitable observation platforms and subsequent analytical approaches must be chosen. We present evidence that fundamental challenges of ferrybox observations can be overcome to yield an internally consistent data source. Subsequently, the behaviour of commonly used bloom metrics in presence of decadal trends can be scrutinized in the context of previously reported system

knowledge. Finally, we attempt to disentangle the effects of nutrient availability and meteorological

340 conditions on inter-annual variability in bloom phenology.

4.1 Automated Processing of Ferrybox Observations

Thresholds for speed, flow rate, and data variability were iteratively adjusted to the data set and may not be applicable to other ferrybox implementations. Particularly flow rate, derived from differences in line and hull temperature will likely require tuning to each ferrybox installation. However,

345 here we analysed data from two ferrybox installations, which could be treated with the same set of thresholds. Transect-wise normalization of the quality controlled fluorescence data was adequate to consistently interpret observations collected by different generations of instrumentation. However, this approach crucially depends on continuous temporal coverage of reference measurements for calibration to chla concentrations. Adaptive regression analysis improved the handling of sta-

350 tistical outliers which would otherwise hamper determination of fluorescence yield, while transects for which no bottle samples are available were corrected with an interpolated fluorescence yield derived from the closest bottle-sampled transects. The present procedure allows for automated and reproducible processing which is an improvement over manual quality control. Applying the proposed interpolated fluorescence yield helps in reprocessing and long-term data analysis of ferrybox

355 fluorescence observations to better represent natural variability.

4.2 Variability in Fluorescence Yield

Diurnal fluorescence patterns showed low seasonal dependence after accounting for solar elevation. Unsurprisingly, light intensity is the predominant factor in Baltic Sea phytoplankton fluorescence yield variability. Other seasonal differences in fluorescence response can be attributed to typically

360 higher cloud cover in winter compared to summer and spring/autumn, which was not accounted for in our analysis. The seasonal cycle of species composition, from dinoflagelate and diatom dominated spring communities (Klais et al., 2011) to cyanobacterial summer bloom (Kahru and Elmgren, 2014), influenced fluorescence yield considerably less than diel cycles.

The diurnal variability in fluorescence response of 50 % during an average summer day is within

365 the range of earlier findings, e.g. 66 % ($\pm 33\%$) for near surface observations in upwelled waters of the equatorial Pacific reported by Dandonneau and Neveux (1997) or 30 % for near-surface seaglider observations in Northeast Pacific waters off the Washington coast, USA (Sackmann et al., 2008), although differences in normalization impede direct comparison. The sampling depth of 5 m for Alg@line systems and the high attenuation of the Baltic Sea in comparison to clear Pacific Ocean

370 waters are likely to dampen the observed diurnal variability.

In this study, fluorescence observations during spring, when diurnal variability reached on average 38 %, were binned for five large Baltic Sea areas. At a typical cruising speed of approximately 23 kn each sea area is sampled for at least several hours. This limits the influence of diurnal variability

in fluorescence yield along a transect on derived chla concentration, which is therefore of lesser relevance for the present study. However, if fluorescence measurements were to be quantitatively evaluated at a higher spatial resolution, locally varying fluorescence yield should be accounted for. Analysis of signal-coherence (Groetsch et al., 2014) offers an alternative to quantitative interpretation of fluorescence observations and can be used to qualitatively detect cyanobacterial surface bloom. If light history is known, e.g. from a dedicated irradiance sensor, a correction of diurnal fluorescence yield variability might be possible and further research in this direction is recommended.

4.3 Spring Bloom Timing and Intensity

The presented bloom phenology expands the time series presented by Fleming and Kaitala (2006) and is in good agreement for the overlapping period (2000 - 2004) when comparing the `const5` metric results. Remaining differences are likely due to quality-control and pre-processing procedures on the fluorescence records. The authors reported for GOF, NBP, and the Arkona Sea that bloom typically started in the south and ended in the north, while bloom intensity increased towards the north. These observations are confirmed here. Sea areas not covered in Fleming and Kaitala (2006), e.g the high-biomass BOM and low-biomass SBP and GOT, followed the reported south-north trend in bloom development. Present results also support and expand the findings of Fennel (1999), who showed with simulations and monitoring data from 1994-1996 for the Western Baltic Sea that surface heating in early spring needs to overcome the temperature of maximum density to repress convective mixing and allow spring bloom to emerge. The temperature of maximum density increases with decreasing salinity, so that convective mixing is sustained longer in less saline northern Baltic Sea waters when spring temperature is on the rise. At the same time, incident solar radiation increases slower in the north due to lower solar elevation.

4.4 Trends

Interannual variability in coastal systems exceeds long-term trends by orders of magnitude (Cole et al., 2012). Consequently, trends were observed at relatively low coefficients of correlation. The importance of appropriate data preprocessing and gap handling (e.g Cole et al., 2012; Racault et al., 2014) and choice of metric (Ferreira et al., 2014) has been demonstrated in literature and is further emphasized by the present analysis. Robustness of the reported decadal trends is documented by high statistical significance levels ($p \ll 0.01$, Figs. 7 and 6), which were supported by spatially binning phenology parameters from all examined Baltic Sea areas. Similar trends were observed earlier for individual Baltic Sea areas, however, usually outside 95 % confidence intervals (e.g Wasmund and Uhlig, 2003).

Łysiak Pastuszak et al. (2014) reported stable or increasing chla concentrations for the period 2007-2011 in several Baltic Sea areas despite signs of declining nutrient concentrations. More recently, eutrophication trend reversal and oligotrophication processes were reported by Andersen et al.

(2015), based on analysis of 112 years of consolidated Baltic Sea observations. Both reports considered surface-layer chla concentration in summer as one of the direct indicators for eutrophication, but did not include spring bloom in their assessment. The time series for 2000-2014 that we present here fills this gap: a negative trend in bloom intensity was also found for spring bloom, providing further evidence for their hypothesis of gradual nutrient load reduction.

Thresholds of `const5` and `median5` are fixed for the whole time series. The observed negative trend in peak concentration was expected to introduce an artificial negative trend in bloom duration because an increasingly higher percentile of the distribution is seen below the bloom threshold (Fig. 1). Contrary to this expected behaviour, however, `const5` and `median5` revealed no significant trends in bloom duration. This indicates that the anticipated negative trend in bloom duration was countered by a positive trend, e.g. in bloom intensity. The Weibull-metric is based on concentration distribution-ratios that are calculated individually for each bloom. Therefore, Weibull-metric results for bloom duration are not sensitive to long-term trends in peak concentration. Weibull-distribution metrics confirmed a highly significant, positive trend in bloom duration. These two sets of results corroborate the conclusion that spring blooms in the Baltic Sea have become longer, while chla peak and average concentration levels declined.

This 'flattening' of the concentration distribution is supported by the absence of a trend in time-integrated biomass BLOOMIDX and by shifts in nutrient concentration timing (earlier nutrient peak concentration, later 25 %-of-peak-value day). These results indicate that annually generated spring bloom biomass has not changed significantly over the study period, in contrast to bloom timing. Kahru and Elmgren (2014) found a similar development for cyanobacterial summer surface bloom, and reported decadal oscillations, yet no long-term trend, of surface area covered by cyanobacteria in the period 1979-2013. In the same period, summer bloom initiation moved to earlier dates by $-0.6\,\mathrm{d\,yr^{-1}}$. These results suggest that the gap has decreased between dinoflagelate- and diatom-dominated spring bloom and cyanobacterial summer bloom. Due to the shorter period covered here as compared to the time series presented by Kahru and Elmgren (2014), it cannot be ruled out that the spring bloom trends are caused by decadal oscillation. Moreover, Alg@line nutrient records often did not commence sufficiently early in the season to record bloom onset. Trends in bloom start and nutrient peak timing can therefore not be derived at the same accuracy and precision as the other phenological parameters. In future, additional data and longer time series may revise this analysis. To this end, nutrient metrics derived in this work are provided in the appendix.

Our findings emphasize that bloom timing is an essential indicator to monitor marine ecosystem dynamics, and thus eutrophication status. Observations at high temporal resolution and choice of bloom metrics are crucial to derive bloom timing trends. Eutrophication status assessment frameworks such as HEAT3.0 (Andersen et al., 2015) may be adapted to embrace available high-frequency data sources to include bloom timing in their analysis. The present results may also prove useful in the calibration and validation of ecosystem models of the Baltic Sea.

4.5 Environmental Forcing

Gradually decreasing nutrient concentrations (Łysiak Pastuszak et al., 2014; Andersen et al., 2015), as well as rising average air- and sea-surface temperatures (Omstedt et al., 2004; Borsenkova et al., 2013) have been reported for recent years, corresponding to a combination of nutrient reduction efforts and global climate change. Several scenarios for future change are plausible (Duarte et al., 2009) but extrapolation of the present results to climate scenarios is beyond the scope of this study. We nevertheless make an attempt to attribute the observed bloom phenology shifts to reported changes in environmental drivers.

Winter-time nutrient concentration and bloom intensity were positively correlated if no spatial normalization was applied. This supports the paradigm that the first-order driver of bloom intensity is nutrient availability. Lacking alternative explanations, we attribute the reported negative trend in bloom peak concentration to declining nutrient concentrations. First-order spatial trends in bloom intensity and timing can be removed by an area-wise z-score normalization, which effectively constrains the analysis to inter-annual variability. After this normalization both regression and PCA resulted in negative correlation between winter-time nutrient concentration and bloom intensity. This negative feedback can be understood as a subtle interaction between meteorological forcing and nutrient supply: strong wind-forced mixing can cause upwelling of deep, nutrient rich waters to surface layers. Wind speed, however, was found to be negatively correlated to the prevalent light level, as well as to bloom duration and bloom index. Therefore, in years when additional nutrients are available due to strong wind forced mixing, low-light regimes that can slow down bloom development are also likely to prevail.

Bloom duration co-varied primarily with weather conditions, e.g. high irradiance levels and low wind speeds were frequently observed for long-lasting blooms (and vice versa). Although the same pattern was observed for bloom timing, no trend was found for bloom start- and peak-day. Increasingly favourable meteorological conditions in late bloom phases are thus a likely driver for the observed increase in bloom duration. Similar weather-driven modulations of bloom timing were reported earlier (Fleming and Kaitala, 2006; Meier et al., 2011; Neumann et al., 2012) for spring, and especially cyanobacterial summer bloom (Wasmund, 1997; Kanoshina et al., 2003; Wynne et al., 2010, 2011).

5 Conclusions

A Baltic Sea spring bloom phenology was derived from 15 years of automated ferrybox chla fluorescence observations. Procedures for automated quality control and processing were introduced and uncertainty due to diurnal variability in phytoplankton fluorescence response was resolved. Both innovations promote increased use of ferrybox observations for scientific research and monitoring purposes, such as the periodic HELCOM eutrophication status assessments. Negative trends in spring

bloom peak- and average-concentration were found and an increase in bloom duration was derived from conceptually differing bloom metrics. Inter annual variability in bloom intensity was primarily linked to nutrient availability, while bloom timing and duration was found to be related to meteorological conditions. In the future, these findings might help to better disentangle ecosystem response
485 to changing nutrient availability and climatic conditions.

Acknowledgements. The authors thank the Alg@line consortium, specifically scientists and technical personnel at SYKE and FIMR, for the ferrybox in situ data set. Acknowledgement is made to ECMWF for the use of their ERA-Interim data set in this research. MAE, SWMP and PMMG were co-funded by the European Community Seventh Framework Programme under grant agreement 607325 AQUA-USERS, and grant agreement 313256
490 GLaSS. PMMG also received support from EC/IAPP project WaterS (Grant 251527). We sincerely thank the anonymous reviewers for their detailed comments and constructive criticism on the manuscript.

References

Adrian, R., Reilly, C. M. O., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingstone, D. M., Sommaruga, R., Straile, D., and Van Donk, E.: Lakes as sentinels of climate change, Limnology and Oceanography, 54, 2283–2297, 2009.

Ainsworth, C.: FerryBoxes begin to make waves, Science, 322, 1627–1629, 2008.

Andersen, J. H., Carstensen, J., Conley, D. J., Dromph, K., Fleming-Lehtinen, V., Gustafsson, B. G., Josefson, A. B., Norkko, A., Villnäs, A., and Murray, C.: Long-term temporal and spatial trends in eutrophication status of the Baltic Sea, Biological Reviews, pp. n/a–n/a, 2015.

Bianchi, T. and Engelhaupt, E.: Cyanobacterial blooms in the Baltic Sea: natural or human-induced?, Limnology and Oceanography, 45, 716–726, 2000.

Borsenkova, I., Christensen, O. B., Elken, J., Haapala, J., Hünicke, B., Käyhkö, J., Niedźwiedź, T., Niemelä, P., Rasmus, S., Rutgersson, A., Schneider, B., Viitasalo, M., Wibig, J., and Zorita-Calvo, E.: Climate Change in the Baltic Sea Are, HELCOM thematic assessment in 2013, Tech. rep., Helcom, 2013.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. a., and Schwartz, M. D.: Shifting plant phenology in response to global change, Trends in Ecology and Evolution, 22, 357–365, 2007.

Cloern, J. E., Abreu, P. C., Carstensen, J., Chauvaud, L., Elmgren, R., Grall, J., Greening, H., Johansson, J. R., Kahru, M., Sherwood, E. T., Xu, J., and Yin, K.: Human Activities and Climate Variability Drive Fast-Paced Change across the World's Estuarine-Coastal Ecosystems, Global Change Biology, pp. n/a–n/a, 2015.

Cole, H., Henson, S., Martin, A., and Yool, A.: Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics, Journal of Geophysical Research, 117, 1–8, 2012.

Dandonneau, Y. and Neveux, J.: Diel variations of in vivo fluorescence in the eastern equatorial Pacific: an unvarying pattern, Deep Sea Research Part II, 44, 1869–1880, 1997.

Dee, D. P., Uppala, S. M., Simmons, a. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. a., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, a. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, a. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, a. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, 2011.

Duarte, C. M., Conley, D. J., Carstensen, J., and Sánchez-Camacho, M.: Return to Neverland: Shifting Baselines Affect Eutrophication Restoration Targets, Estuaries and Coasts, 32, 29–36, 2009.

Fennel, K.: Convection and the timing of phytoplankton spring blooms in the western Baltic Sea, Estuarine, Coastal and Shelf Science, 49, 113–128, 1999.

Ferreira, A. S., Visser, A. W., MacKenzie, B. R., and Payne, M. R.: Accuracy and precision in the calculation of phenologymetrics, Journal of Geophysical Research: Oceans RESEARCH, pp. 2121–2128, 2014.

Fleming, V. and Kaitala, S.: Phytoplankton spring bloom intensity index for the Baltic Sea estimated for the years 1992 to 2004, Hydrobiologia, 554, 57–65, 2006.

Fleming-Lehtinen, V., Andersen, J. H., Carstensen, J., Łysiak Pastuszak, E., Murray, C., Pyhälä, M., and Laamanen, M.: Recent developments in assessment methodology reveal that the Baltic Sea eutrophication problem is expanding, Ecological Indicators, 48, 380–388, 2015.

Gnanadesikan, A. and Anderson, W. G.: Ocean Water Clarity and the Ocean General Circulation in a Coupled Climate Model, Journal of Physical Oceanography, 39, 314–332, 2009.

Grayek, S. and Staneva, J.: Use of FerryBox surface temperature and salinity measurements to improve model
535   based state estimates for the German Bight, Journal of Marine Systems, 2011.

Groetsch, P. M., Simis, S. G., Eleveld, M. A., and Peters, S. W.: Cyanobacterial bloom detection based on coherence between ferrybox observations, Journal of Marine Systems, 140, 50–58, 2014.

Hays, G. C., Richardson, A. J., and Robinson, C.: Climate change and marine plankton, Trends in Ecology and Evolution, 20, 337–344, 2005.

540   Heisler, J., Glibert, P., Burkholder, J., Anderson, D., Cochlan, W., Dennison, W., Dortch, Q., Gobler, C., Heil, C., Humphries, E., Lewitus, a., Magnien, R., Marshall, H., Sellner, K., Stockwell, D., Stoecker, D., and Suddleson, M.: Eutrophication and harmful algal blooms: A scientific consensus, Harmful Algae, 8, 3–13, 2008.

HELCOM: Baltic Sea Action Plan, Tech. Rep. November, Helcom, 2007a.

545   HELCOM: Climate Change in the Baltic Sea area, Baltic Sea Environment Proceedings, 111, 2007b.

HELCOM: Convention on the Protection of the Marine Environment of the Baltic Sea Area , 1992 (Helsinki Convention), 2008.

HELCOM: Manual for Marine Monitoring in the COMBINE Program of HELCOM, Tech. Rep. September, Helcom, 2013.

550   Ho, J. C. and Michalak, A. M.: Challenges in tracking harmful algal blooms : A synthesis of evidence from Lake Erie, Journal of Great Lakes Research, 41, 317–325, 2015.

Ji, R., Edwards, M., Mackas, D. L., Runge, J. a., and Thomas, A. C.: Marine plankton phenology and life history in a changing climate: current research and future directions., Journal of plankton research, 32, 1355–1368, 2010.

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

Kahru, M. and Nommann, S.: The phytoplankton spring bloom in the Baltic Sea in 1985, 1986: multitude of spatio-temporal scales, Continental Shelf Research, 10, 329–354, 1990.

Kanoshina, I., Lips, U., and Leppänen, J.-M.: The influence of weather conditions (temperature and wind) on
560   cyanobacterial bloom development in the Gulf of Finland (Baltic Sea), Harmful Algae, 2, 29–41, 2003.

Kiefer, D. A.: Fluorescence properties of natural phytoplankton populations, Marine Biology, 22, 263–269, 1973.

Kiefer, D. a., Chamberlin, W. S., and Booth, C. R.: Natural fluorescence of chlorophyll a: Relationship to photosynthesis and chlorophyll concentration in the western South Pacific gyre, Limnology and Oceanography,
565   34, 868–881, 1989.

Klais, R., Tamminen, T., Kremp, A., Spilling, K., and Olli, K.: Decadal-scale changes of dinoflagellates and diatoms in the anomalous baltic sea spring bloom., PloS one, 6, e21 567, 2011.

Korpinen, S., Meski, L., Andersen, J. H., and Laamanen, M.: Human pressures and their potential impact on the Baltic Sea ecosystem, Ecological Indicators, 15, 105–114, 2012.

570   Kuusisto, M., Koponen, J., and Sarkkula, J.: Modelled phytoplankton dynamics in the Gulf of Finland, Environmental Modelling & Software, 13, 461–470, 1998.

Leppänen, J. M., Rantajärvi, E., Maunumaa, M., Larinmaa, M., and Pajala, J.: Unattended algal monitoring system-a high resolution method for detection of phytoplankton blooms in the Baltic Sea, Proceedings of OCEANS94, 1, 461–463, 1994.

575 Leppänen, J.-M., Rantajärvi, E., Hällfors, S., Kruskopf, M., and Laine, V.: Unattended monitoring of potentially toxic phytoplankton species in the Baltic Sea in 1993, Journal of Plankton Research, 17, 891–902, 1995.

Lips, I., Rünk, N., Kikas, V., Meerits, A., and Lips, U.: High-resolution dynamics of the spring bloom in the Gulf of Finland of the Baltic Sea, Journal of Marine Systems, 129, 135–149, 2014.

Marra, J.: Analysis of diel variability in chlorophyll fluorescence, Journal of marine research, pp. 767–784, 580 1997.

Meier, H. E. M., Eilola, K., and Almroth, E.: Climate-related changes in marine ecosystems simulated with a 3-dimensional coupled physical-biogeochemical model of the Baltic sea, Climate Research, 48, 31–55, 2011.

Nelson, D. and Smith, W.: Sverdrup revisited: Critical depths, maximum chlorophyll and the control of Southern Ocean productivity by the irradiance-mixing regime, Limnol. Oceanog, 36, 1650–1661, 1991.

585 Neumann, T., Fennel, W., and Kremp, C.: Experimental simulations with an ecosystem model of the Baltic Sea: a nutrient load reduction experiment, Global Biogeochemical Cycles, 16, 1–12, 2002.

Neumann, T., Eilola, K., Gustafsson, B., Müller-Karulis, B., Kuznetsov, I., Meier, H. E. M., and Savchuk, O. P.: Extremes of temperature, oxygen and blooms in the baltic sea in a changing climate, Ambio, 41, 574–585, 2012.

590 Omstedt, A., Pettersen, C., Rodhe, J., and Winsor, P.: Baltic Sea climate: 200 yr of data on air temperature, sea level variation, ice cover, and atmospheric circulation, Climate Research, 25, 205–216, 2004.

Paerl, H. and Huisman, J.: Blooms like it hot, Science, 320, 57–58, 2008.

Paerl, H. W. and Huisman, J.: Climate change: a catalyst for global expansion of harmful cyanobacterial blooms, Environmental Microbiology Reports, 1, 27–37, 2009.

595 Pedregosa, F. and Varoquaux, G.: Scikit-learn: Machine learning in Python, Journal of Machine Learning, 12, 2825–2830, 2011.

Platt, T., White, G. N., Zhai, L., Sathyendranath, S., and Roy, S.: The phenology of phytoplankton blooms: Ecosystem indicators from remote sensing, Ecological Modelling, 220, 3057–3069, 2009.

Racault, M. F., Platt, T., Sathyendranath, S., Agirbas, E., Martinez Vicente, V., and Brewin, R.: Plankton indi-600 cators and ocean observing systems: Support to the marine ecosystem state assessment, Journal of Plankton Research, 36, 621–629, 2014.

Racault, M.-F., Raitsos, D. E., Berumen, M. L., Brewin, R. J., Platt, T., Sathyendranath, S., and Hoteit, I.: Phytoplankton phenology indices in coral reef ecosystems: Application to ocean-color observations in the Red Sea, Remote Sensing of Environment, 160, 222–234, 2015.

605 Rantajärvi, E., Flinkman, J., Ruokanen, L., Hällfors, S., Stipa, T., Suominen, T., Kaitala, S., Maunula, P., Fleming, V., Lips, U., London, L., Vepsäläinen, J., Nyman, E., Neuvonen, S., Kankaanpää, H., Seppälä, J., Perttilä, M., Raateoja, M., and Haahti, H.: Alg@line in 2003: 10 Years of Innovative Plankton Monitoring and Research an Operational Information Service in the Baltic Sea, Report Series of the Finnish Institute of Marine Research, 48, 55, 2003.

610 Roesler, C. S. and Barnard, A. H.: Optical proxy for phytoplankton biomass in the absence of photophysiology: Rethinking the absorption line height, Methods in Oceanography, 7, 79–94, 2013.

Rolinski, S., Horn, H., Petzoldt, T., and Paul, L.: Identifying cardinal dates in phytoplankton time series to enable the analysis of long-term trends., Oecologia, 153, 997–1008, 2007.

Ruokanen, L., Kaitala, S., Flemming, V., and Maunula, P.: Alg@line: joint operational unattended phytoplankton monitoring in the Baltic Sea, Elsevier Oceanography, 69, 519–522, 2003.

Sackmann, B. S., Perry, M. J., and Eriksen, C. C.: Seaglider observations of variability in daytime fluorescence quenching of chlorophyll-*a* in Northeastern Pacific coastal waters, Biogeosciences Discussions, 5, 2839–2865, 2008.

Seppälä, J. and Olli, K.: Multivariate analysis of phytoplankton spectral in vivo fluorescence: estimation of phytoplankton biomass during a mesocosm study in the Baltic Sea, Marine Ecology Progress Series, 370, 69–85, 2008.

Seppälä, J., Ylöstalo, P., Kaitala, S., Hällfors, S., Raateoja, M., and Maunula, P.: Ship-of-opportunity based phycocyanin fluorescence monitoring of the filamentous cyanobacteria bloom dynamics in the Baltic Sea, Estuarine, Coastal and Shelf Science, 73, 489–500, 2007.

Sharples, J., Ross, O. N., Scott, B. E., Greenstreet, S. P., and Fraser, H.: Inter-annual variability in the timing of stratification and the spring bloom in the North-western North Sea, Continental Shelf Research, 26, 733–751, 2006.

Siegel, D. a., Doney, S. C., and Yoder, J. a.: The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis., Science (New York, N.Y.), 296, 730–733, 2002.

Smetacek, V. and Passow, U.: Spring bloom initiation and Sverdrup's critical-depth model, Limnology and Oceanography, 35, 228–234, 1990.

Sommer, U. and Lengfellner, K.: Climate change and the timing, magnitude, and composition of the phytoplankton spring bloom, Global Change Biology, 14, 1199–1208, 2008.

Sverdrup, H.: On conditions for the vernal blooming of phytoplankton, Journal du Conseil, 18, 287–295, 1953.

Tamminen, T. and Andersen, T.: Seasonal phytoplankton nutrient limitation patterns as revealed by bioassays over Baltic Sea gradients of salinity and eutrophication, Marine Ecology Progress Series, 340, 121–138, 2007.

Ueyama, R. and Monger, B. C.: Wind-induced modulation of seasonal phytoplankton blooms in the North Atlantic derived from satellite observations, Limnology and Oceanography, 50, 1820–1829, 2005.

Vargas, M., Brown, C. W., and Sapiano, M. R. P.: Phenology of marine phytoplankton from satellite ocean color measurements, Geophysical Research Letters, 36, 2–6, 2009.

Voss, M., Dippner, J. W., Humborg, C., Hürdler, J., Korth, F., Neumann, T., Schernewski, G., and Venohr, M.: History and scenarios of future development of Baltic Sea eutrophication, Estuarine, Coastal and Shelf Science, 92, 307–322, 2011.

Wasmund, N.: Occurrence of Cyanobacterial Blooms in the Baltic Sea in relation to environmental conditions, Internationale Revue der gesamten Hydrobiologie und Hydrographie, 1997.

Wasmund, N. and Uhlig, S.: Phytoplankton trends in the Baltic Sea, ICES Journal of Marine Science: . . . , 3139, 177–186, 2003.

Wasmund, N., Tuimala, J., Suikkanen, S., Vandepitte, L., and Kraberg, A.: Long-term trends in phytoplankton composition in the western and central Baltic Sea, Journal of Marine Systems, 87, 145–159, 2011.

Wasmund, N., Nausch, G., and Feistel, R.: Silicate consumption: An indicator for long-term trends in spring diatom development in the Baltic Sea, Journal of Plankton Research, 35, 393–406, 2013.

Wiltshire, K. H., Malzahn, A. M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., Manly, B. F. J., and Boersma, M.: Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads, Limnology and Oceanography, 53, 1294–1302, 2008.

Winder, M. and Cloern, J. E.: The annual cycles of phytoplankton biomass., Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 365, 3215–3226, 2010.

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., and Dyble, J.: Characterizing a cyanobacterial bloom in Western Lake Erie using satellite imagery and meteorological data, Limnology and Oceanography, 55, 2025–2036, 2010.

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Schwab, D. J., Watabayashi, G. Y., and Christensen, J. D.: Estimating cyanobacterial bloom transport by coupling remotely sensed imagery and a hydrodynamic model., Ecological Applications, 21, 2709–21, 2011.

Zhai, L., Platt, T., Tang, C., Sathyendranath, S., and Hernández Walls, R.: Phytoplankton phenology on the Scotian Shelf, ICES Journal of Marine Science, 68, 781–791, 2011.

Zhai, L., Platt, T., Tang, C., Sathyendranath, S., and Walne, A.: The response of phytoplankton to climate variability associated with the North Atlantic Oscillation, Deep Sea Research Part II: Topical Studies in Oceanography, 93, 159–168, 2013.

Łysiak Pastuszak, E., Carstens, M., Leppänen, J.-M., Leujak, W., Nausch, G., Murray, C., and Jesper H., A.: Eutrophication status of the Baltic Sea 2007-2011, Tech. Rep. 143, Helcom, 2014.

Table 1. Quality control flag definitions and statistics. Observations were omitted if any of the flags exceeded the respective threshold. Absolute temperature difference is measured between the water intake and the flow-through sensors. Availability and rejection rates were calculated relative to the total number of observations.

	Sign	Threshold	Availability [%]	Rejection Rate [%]
Speed, [kn]	$<$	5	100	1.33
Flow, [L min^{-1}]	$<$	0.3	35.95	1.38
Abs. Temp. Diff. [°C]	$>$	2	67.17	2.12
STD Latitude, [°]	$<, >$	$1e^{-4}, 0.5$	100	0.96
STD CHLa Fl., [mg m^{-3}]	$<$	$1e^{-4}$	87.65	0.75
All				4.55

Table 2. Description and acronyms of bloom phenology, nutrient, and meteorological parameters that were used in the trend and multi-variate analysis.

Parameter	Unit	Description
BLOOMIDX	mg d m^{-3}	Integrated chlorophyll-*a* concentration during bloom
CONCAVG	mg m^{-3}	Average (mean) chlorophyll-*a* concentration during bloom
PEAKHEIGHT	mg m^{-3}	Highest chlorophyll-*a* concentration during bloom
STARTDAY	Julian Day	Bloom start day
PEAKDAY	Julian Day	Bloom peak day
ENDDAY	Julian Day	Bloom end day
NUT-MINVALUE	μmol L^{-1}	Nutrient concentration at end of bloom
NUT-MINDAY-50	Julian Day	Day when nutrients equalled 50 % of NUT-PEAKVALUE
NUT-PEAKVALUE	μmol L^{-1}	Pre-bloom (winter-time) nutrient concentration
NUT-PEAKDAY	Julian Day	Day of NUT-PEAKVALUE
NUT-DEPLAY-25	Julian Day	Day when nutrient concentration equalled 25 % of NUT-PEAKVALUE
NUT-DEPLAY-50	Julian Day	Day when nutrients concentration equalled 50 % of NUT-PEAKVALUE
NUT-SLOPE	μmol L^{-1} d^{-1}	Rate of nutrient depletion between 75 and 25 % of NUT-PEAKVALUE
PAR	W m^{-2} d^{-1}	Average (seasonally detrended) photosynthetically active radiation level
SST	°C	Average (seasonally detrended) sea surface temperature
WIND	m s^{-1}	Average (seasonally detrended) wind speed

Table 3. Bloom timing and intensity for each investigated sea area (Figure 2) and for all applied bloom metrics.

Parameter	Metric	BOM Mean	BOM Std	SBP Mean	SBP Std	GOT Mean	GOT Std	NBP Mean	NBP Std	GOF Mean	GOF Std
STARTDAY, [Julian Day]	const5	68	9.3	87	13.6	95	8	89	5.6	81	8.4
	median5	65	9.4	73	7.3	83	9.4	86	4.4	81	8.5
	weibull	64	12.7	73	7.2	84	6	87	4.1	89	4.7
PEAKDAY, [Julian Day]	all metrics	75	14.7	92	14.9	106	7.4	108	4.4	112	4.7
ENDDAY, [Julian Day]	const5	95	15.2	102	12.5	118	10.7	130	5.2	143	5.6
	median5	107	20.3	115	13.1	133	7.1	142	3.3	143	5.2
	weibull	94	18.4	116	15.4	128	9	126	5.6	132	5.7
DURATION, [d]	const5	35	12.5	16	12.4	23	13.6	41	6.4	62	11.7
	median5	54	18.9	46	15.4	47	8.8	54	4.4	62	12.3
	weibull	36	10.8	43	16.4	44	10.9	40	6.1	43	6.7
BLOOMIDX, [mg d m^{-3}]	const5	283	167.6	98.4	77.1	162.4	114.2	352.4	84.9	691.6	157.6
	median5	334.3	135.1	197	92.6	224.5	77.2	386.9	72.9	694	165.9
	weibull	356	178.5	196.7	74.1	232.9	64.1	340.1	62.1	673.7	175.9
CONCAVG, [mg m^{-3}]	const5	7.3	2.1	5.3	0.9	6.2	1.3	8.4	1.6	11.7	2.2
	median5	6	1.2	4.1	0.7	4.6	1.1	7	1.1	11.7	2.3
	weibull	9.9	4.3	4.6	1.1	5.5	1.7	8.5	2.2	13.6	3.3
PEAKHEIGHT, [mg m^{-3}]	all metrics	12.3	5.2	6.1	1.7	7.2	2.3	11.3	2.9	20.2	5.7
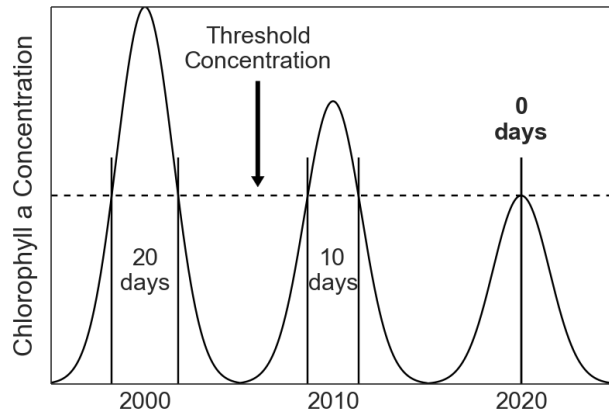
Figure 1. Illustration of threshold-based bloom metric behaviour when applied to a dataset with a negative peak concentration trend.
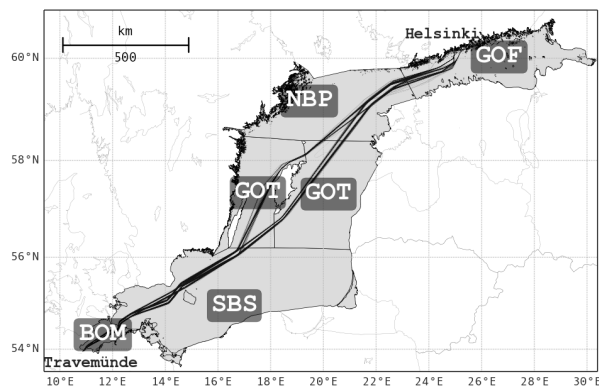
Figure 2. Transect of M/S *Finnmaid* and M/S *Finnpartner* through the Baltic Sea from Helsinki (Finland) to Travemünde (Germany) (v.v.). The following sea areas are considered in this study: the Western Gulf of Finland (GOF: $>59.5\,^\circ N$ latitude, along transect), the Northern Baltic Proper (NBP: 58.4-59.5 $^\circ N$ latitude, along transect), the Western and Eastern Gotland basins (GOT: 56.2-58.4 $^\circ N$ latitude, along transect), the Southern Baltic Proper (SBP: 54.5-56.2 $^\circ N$ latitude, along transect) and the Bay of Mecklenburg (BOM: $<54.5\,^\circ N$ latitude, along transect). Depending on weather conditions the North- or South of Gotland routes were sailed.
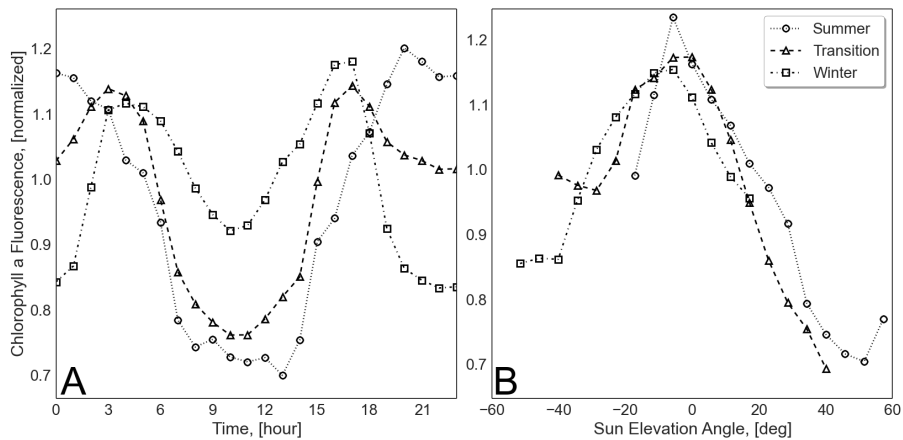
Figure 3. Diurnal variability in the chlorophyll-*a* fluorescence yield: (A) normalized (division by transect-mean) chlorophyll-*a* fluorescence observations plotted against time-of-day, and (B) sun elevation angle. The analysis was carried out on four subsets: winter (November - February), summer (May - August), and transition periods (March, April, September, October) using all ferrybox observations along the routes shown in Figure 2.
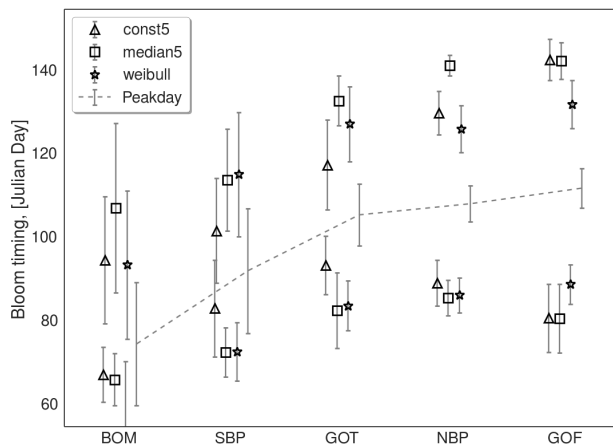
Figure 4. Bloom timing (bloom start, peak, and end day) for each sea area along the routes in Figure 2, averaged over the period 2000 to 2014, and for all applied bloom metrics. Whiskers indicate standard deviations over the 15-year study period. The bloom peak-day is independent of the chosen metric and plotted separately. The sea areas are ordered by latitude, from south to north.
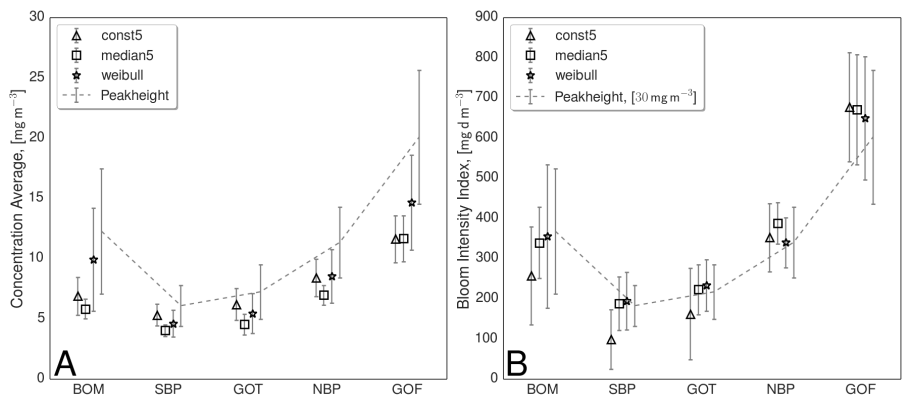
Figure 5. (A) Concentration average and (B) bloom intensity index for each sea area along the routes in Figure 2, averaged over the years 2000 to 2014, and for all applied bloom metrics. Whiskers indicate standard deviations over the 15-year study period. The sea areas are ordered by latitude. The metric-independent bloom peak concentration is added in both plots for visual comparison.
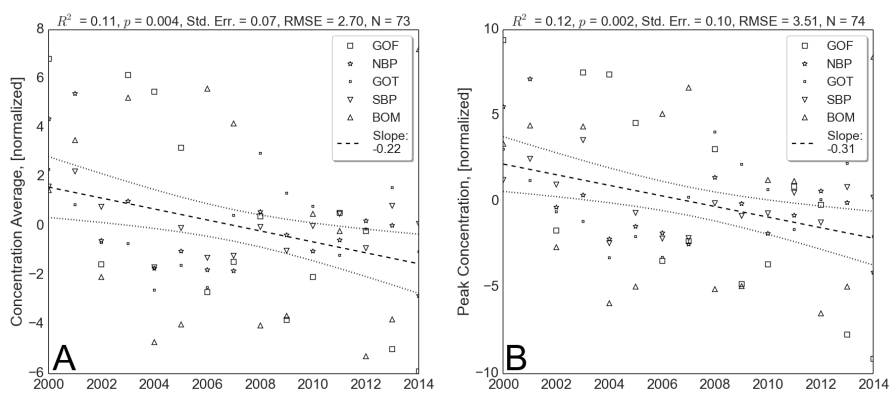
Figure 6. (A) Decadal trend of average (CONCAVG) and (B) peak (PEAKHEIGHT) chlorophyll-*a* concentration during bloom conditions, derived from the Weibull-distribtion metric. Concentrations were normalized prior to regression (subtraction of area-average concentration). Dashed lines indicate the trend line (bold) and its confidence intervals (5 %, small dashes).
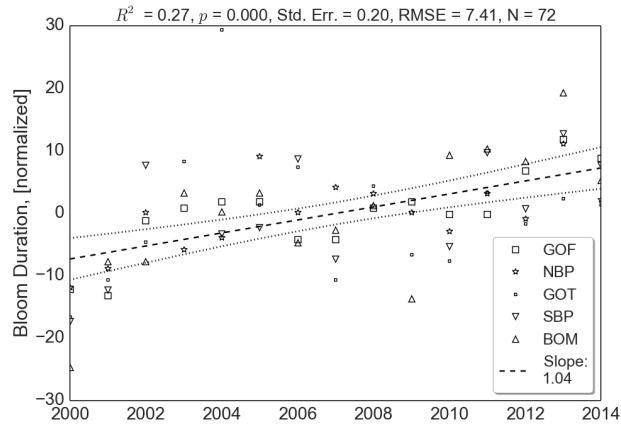
Figure 7. Decadal trend of bloom duration (DURATION), calculated with the Weibull-distribution metric. Durations were normalized prior to regression (subtraction of area-average duration). Dashed lines indicate the trend line (bold) and its confidence intervals (5 %, small dashes).
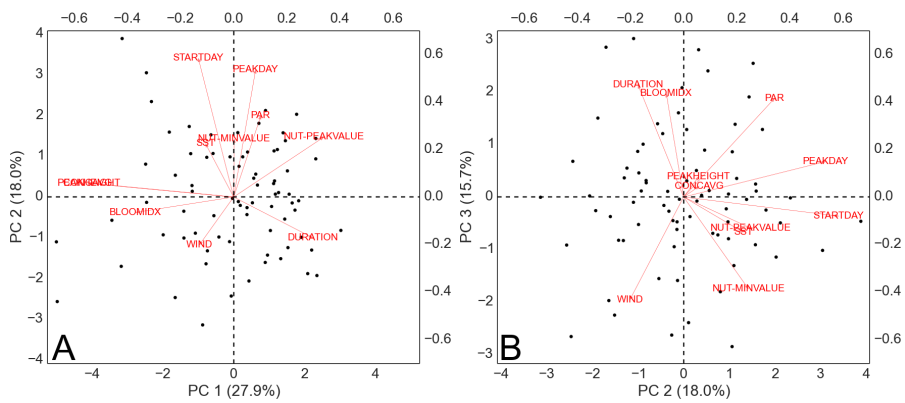
Figure 8. Principal component analysis bi-plots: arrows indicate correlation of a parameter with the principal components (bottom- and left-axes, percentages refer to the variability explained by the principal component), and black dots indicate scores of individual observations (top- and right-axes) on the principal components (A: component 1 and 2, B: component 2 and 3).