Response to reviewers on 'Spring Blooms in the Baltic Sea have weakened but lengthened from 2000 to 2014' by P. M. M. Groetsch et al.

We would like to thank the two reviewers for the comments provided. After implementing revision changes the manuscript was proofread leading to a number of minor corrections (punctuation, grammar, clarity). An error in section 2.1 was corrected were two numbers were switched around (percentages of routes sailed east and west of Gotland). We also introduced some minor simplifications to the figures (no grid lines, no legend shading, open symbols).

Anonymous Referee #1
Received and published: 5 February 2016

1. Could you please provide some information on the source(s) of nutrient data in chap. 2.3.

Authors’ response: These concentrations were derived from laboratory analysis of bottle samples that were regularly collected along the transect (for further detail see section 2.1. in the manuscript). We added this information to section 2.3 to clarify this point.

2. The title is not much appealing to me. However, I don’t have any good suggestion. Probably you may include alg@line.

Authors’ response: The title was chosen to highlight the scope of the paper (phenological study of Baltic Sea spring bloom) rather than the methods used to obtain this value-adding dataset. We hope that this will draw readers to the paper who would not normally consider ship-of-opportunity pigment fluorescence data for this type of analysis. We included an early reference to the 'Alg@line' network in the 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'

Anonymous Referee #2
Received and published: 15 February 2016

1. Lines 90-95 / Figure 1. Text mentions that “any threshold-based metric” would introduce artificial trends in bloom duration. This is a clear problem for “fixed threshold” metrics, but not for “variable thresholds” as Siegel et al. (2002), which is later introduced.

Authors’ response: Please see our detailed response with question 2).

Furthermore, results using the fixed threshold const5 show a negative trend in peak concentrations, but no significant trend in bloom duration. This seems a somewhat inconsistent. Further discussion would
help clarify why the expected artificial trends do not occur.

Authors’ response: Peak concentrations are derived the same for all metrics, thus the negative trend is not metric dependent. Indeed, no artificial negative trend in bloom duration was observed for any threshold-based (fixed, or derived from climatology; see question 2)) metrics. We argue that this observation can be expected if blooms also became longer over the study period. An independent confirmation of this hypothesis is the strong positive trend in bloom duration in the Weibull-metric results. This is now stated more explicitly in section 4.4:

‘Thresholds of const5 and median5 are fixed for the whole time-series. The observed negative trend in peak concentration introduces an artificial negative trend in bloom duration due to a shortening of the part of the curve seen above the bloom threshold, but this is solely by decreased amplitude of the curve (see Fig. 1). Contrary to this expected behavior, however, const5 and median5 revealed no significant trends in bloom duration. This indicates that the anticipated negative trend was countered by a positive trend. The Weibull-metric is based on concentration distribution-ratios that are calculated for each bloom individually. 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 mutually support the conclusion that spring blooms in the Baltic Sea have become longer, while chla peak and average concentration levels declined.’

2. It is not clear to me whether median5 (Siegel et al. 2002) is calculated for each individual annual median or for all years together. The latter would indeed produce a fixed threshold for each region (see previous comment). That detail is unclear in Siegel et al. 2002 as well, but see Henson et al. 2009 (Decadal variability in North Atlantic phytoplankton blooms – J. Geophys. Res.) and Brody et al. 2013 (A comparison of methods to determine phytoplankton bloom initiation – J. Geophys. Res.).

Authors’ response: Indeed we assumed that Siegel et al. 2002 referred to the climatological median, rather than the annual median. Brody et al. 2013, however, state that both thresholds can be applied:

‘The threshold bloom initiation method was introduced for marine phenology studies in Siegel et al. [2002]. This method finds the yearly or climatological median of a chlorophyll time series, then identifies the bloom start date as the first point at which chlorophyll levels rise a certain percentage above the median.’

To make the distinction clear we changed the respective paragraph in the introduction:

‘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) 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 for the given example, because bloom intensity does not influence these metrics.’

3. Lines 195-200: Day-of-year 31 is January 1?

Authors’ response: This is in error and should read 'day-of-year between 31 (31 January) and 160 (9 June).
4. Why was the time frame between day 31 -160 selected? Is it possible that nutrient peak concentration occur prior to the minimum date considered? A shift to earlier peak nutrient concentrations is mentioned, but results of the nutrient metrics are not presented. I suggest extending Table 3 and/or including plots to support this.

Authors’ response:
The ship-of-opportunity (Alg@line) measurements typically commenced in late January, which is why we chose 31 January as the start of our analysis. The end date was chosen such that it covers all spring bloom events in all basins but not summer bloom. We added this information after the first paragraph of section 2.4.
The nutrient peak concentration is closely related to the day of bloom initiation, which is typically at least a month later (see table 3). Of more concern is that for several years, Alg@line data collection commenced after bloom initiation. These data were consequently omitted from statistical analysis (replaced with multi-year median), including the nutrient statistics. Further detail on this issue is also given in response to question 7. Table 3 shows multi-year averages of calculated parameters, so we can not expand to trends in nutrient timing or intensity from these. Since these results are nevertheless available, we added all nutrient metric results to the appendix to aid future research.

5. Lines 230-235: In 30 out 225 data combinations there were no ferrybox observations to properly identify bloom initiation. In these cases, bloom initiation date was replaced by the median value. It is not clear if this treatment was used only for the principal component analysis or the regressions as well. Cases identified by each timing method only account for 29 (const5:9, median5: 15, weibull: 5). I find it also unclear how these methods identified that the bloom had already started. A few words to clarify would be useful.

Authors’ response: Median-filling of missing dates was applied prior to both PCA and regression analysis. We assumed bloom initiated prior to Alg@line data collection if the first data point already satisfied the bloom criterion for a given metric. This is now stated more clearly already in section 2.4 instead of in section 2.5. The number of missed bloom initiation events is incorrectly stated as 15 for the median5 metric and should be changed to 16.

6. The time series analyzed is relatively short to claim long-term trends, especially when considering the large interannual variability observed in all of the metrics. A study between 1979-2013 where decadal-oscillations were found is mentioned in the text. I would recommend extending the discussion a bit to include how that analysis compares with this one during the same time frame.

Authors’ response: The authors of the mentioned study (Kahru2014) describe surface accumulations of cyanobacterial summer bloom. Links between spring bloom and cyanobacterial summer bloom are certainly worthwhile exploring. However, the complex interactions between light- and nutrient-limited spring bloom, and largely wind-modulated cyanobacterial surface bloom accumulations seem out of scope for the present paper (and may quite possibly be too complex). In section 4.4. we acknowledge the finding of Kahru2014 that summer bloom initiation moved to earlier dates, and thus that the period between dinoflagelate- and diatom-dominated spring bloom and cyanobacterial summer bloom decreased. The following sentence was added to section 4.4 to clarify that we can neither prove nor disprove a decadal oscillation signal based on our time series:
‘However, due to the shorter period covered here as compared to the time series presented by Kahru2014, it cannot be ruled out that the derived trends are caused by decadal oscillation.’
7. The final discussion and conclusions attribute the declining trend in bloom peak concentration to declining nutrient concentrations; however, no decline in winter nutrient concentrations (as estimated here) is reported. The conclusion is based on literature considerations and the “lack(ing) of other explanations”. I think this pattern is quite interesting and an alternative explanation may be supported by the results here presented. The authors report a shift in peak nutrient concentration to earlier dates and a strong correlation between winter nutrient concentration and bloom peak magnitude. Earlier increases in nutrient concentrations mean that nutrient limitation is alleviated earlier during the year, when light limitation might still be strong. As the year progresses and light limitation is alleviated, a fraction of the nutrients has been already consumed. The nutrient concentration “available for blooming” would then not be equal to the winter maximum, but lower than it. That would produce a decrease in the bloom peak magnitude, an apparent extend in bloom duration, but no change in total chlorophyll during the bloom (also reported). This is just a quick idea and might be better captured by , which are mentioned in the introduction, but not used in the analysis. As I mentioned before, I think it is important to include the nutrient concentrations results in the manuscript to better support its conclusions. I would also suggest including the actual time series (environmental factors and fluorescence) as part of supplementary material.

Authors’ response: Unfortunately the temporal resolution of the nutrient concentration data is not sufficient to quantify the timing of nutrient uptake onset and light limitation alleviation – especially in winter when only few cruises are sampled for nutrients. Nevertheless, judging from the few transects sampled for nutrients in December and early January, nutrient limitation is alleviated well before light availability increases, and this has been our understanding of nutrient dynamics in the high latitude, semi-enclosed Baltic Sea. We looked into several metrics for nutrient uptake rates but could not link inter-annual nutrient variability to bloom phenology parameters. This may be due to the relatively sparse collection of bottle samples for laboratory analysis (on average every 6th transect). While at present this result does not prompt further discussion in the manuscript, future studies may benefit from additional data so as suggested, we added nutrient metric results to the appendix. In addition we added the following to the manuscript:

'Several times ship service had not commenced early enough in the year to record bloom onset, which implies that trends in bloom start and nutrient peak timing could not be derived with the same accuracy and precision as the other phenological parameters. Nutrient metrics are provided in the appendix to aid future work, if additional data or longer time-series become available.'
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. 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 along a gradient 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 except for shorter bloom duration in the Bay of Mecklenburg. Bloom peak chlorophyll-a concentrations were highest and most variable) in the Gulf of Finland and the Bay of Mecklenburg.

Bloom peak chlorophyll-a concentration showed a negative trend of \SI{-0.31(10)}{\milli\gram\per\cubic\meter\per\year}. Trend-agnostic distribution-based (Weibull-type) bloom metrics showed a positive trend in bloom duration of \SI{1.04(20)}{\day\per\year}, 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.
Phytoplankton bloom intensity and timing (bloom phenology) are indicators for ecosystem health at the base of the food web \citep{Hays2005, Adrian2009, Vargas2009}. Phenological studies are increasingly used to inspect regional ecosystem response to nutrient reduction efforts \citep{Helcom2007, Voss2011, Fleming-Lehtinen2015} and changing climatic conditions \citep{Sommer2008, Paerl2009}. The Baltic Sea is a coastal ecosystem affected by eutrophication \citep{Korpinen2012}, which intensifies naturally occurring spring- and summer bloom \citep{Bianchi2000, Helcom2007}. The Helsinki Commission formulated a nutrient reduction scheme aimed at improving ecosystem health in 1992 \citep{Helcom1992}, 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 \citep{Leppanen1995, Ainsworth2008}. Phytoplankton pigment fluorometers are included in most so-called ferryboxes. In the Baltic sea, such systems have recorded phytoplankton \citep{31mbloom} on the route from Helsinki to Travemünde (v.v.) since 1992 \citep{Rantajarvi2003}. 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 \citep{Neumann2002, Tamminen2007} and light availability \citep{Sverdrup1953, Smetacek1990, Nelson1991, Siegel2002}, mixing-status \citep{Ueyama2005a, Sharples2006}, temperature \citep{Grayek2011}, ice cover \citep{Karhu1999, Sommer2008}, and salinity \citep{Fennel1999, Tamminen2007}. In addition, the quantum yield of fluorescence is influenced by solar irradiance \citep{Kiefer1973, Dandonneau1997, Marra1997, Sackmann2008a}, species composition, and physiology \citep{Kiefer1989}. Hence, interpretation of unattended pigment fluorescence measurements in terms of phytoplankton biomass presents a number of challenges \citep{Roesler2013}. Firstly, phytoplankton distribution exhibits high spatial and temporal variability, while ferryboxes measure pigment fluorescence at fixed depth \citep{Ruokanen2003}. Therefore, stratified conditions may not be well represented in the data \citep{Groetsch2014}. Secondly, in a typical ferrybox setup fluorescence yield is at best determined as a daily sea area-average, regional average, which disregards variability on smaller spatio-temporal scales. Despite these challenges, \citep{Fleming2006} demonstrated that ferrybox observations in the Baltic Sea can be used to derive bloom timing and intensity for biomass-rich sea areas. The authors reported 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 \citep{Klais2011, Wasmund2011, Wasmund2013}, but \citep{Klais2011, Wasmund2011, Wasmund2013}. \citep{Kahru2014} reported that the timing of cyanobacterial surface accumulations has advanced approximately 20 days from 1979 to 2013. However, in formation about shifts in

Choosing an adequate bloom metric is not trivial as no strict clear guidelines exist that unambiguously recommend conclusively supporting one metric over the other. 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 \citep{Siegel2002, Fle
1) threshold- and growth-rate based metrics (Rolinski2007, Wiltshire2008), and 2) distribution-based metrics (Rolinski2007, Platt2009, Vargas2009, Zhai2011). 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. Threshold-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 (Ji2010). Interpretation based on several, conceptually different bloom metrics can be used to obtain uncertainty estimates (Ho2015), and it also allows to [3] imqualitatively indicate 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 5 illustrates how a gradual decline (negative trend) in peak concentration causes a threshold-based metric based on fixed thresholds (e.g. derived from climatology or expert-judgement) to introduce an artificial negative trend in bloom duration. In contrast, growth-rate and distribution-based metrics based on growth rate, distribution, or annually derived thresholds yield a constant single bloom duration for the given in this example because they are sensitive to concentration distributions, rather than absolute concentrations. 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. This paper describes a methodology to derive quality controlled time-series of chlorophyll-\textit{a} concentrations from observations collected by the Baltic Sea Alg@line program and its predecessors 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 represented, and explored for long-term trends. Inter-annual variability of bloom phenology 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.

\section{Materials and Methods}

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. This paper describes a methodology to derive quality controlled time-series of chlorophyll-\textit{a} concentrations from observations collected by the Baltic Sea Alg@line program and its predecessors 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 represented, and explored for long-term trends. Inter-annual variability of bloom phenology 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.

\section{Materials and Methods}

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. This paper describes a methodology to derive quality controlled time-series of chlorophyll-\textit{a} concentrations from observations collected by the Baltic Sea Alg@line program and its predecessors 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 represented, and explored for long-term trends. Inter-annual variability of bloom phenology 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.
Details on the instrumentation of the Alg@line ferrybox systems can be found in \cite{Leppanen1994, Rantajarvi2003, Ruokanen2003, Seppala2007}. In summary, the systems recorded in vivo fluorescence of chlorophyll-$\alpha$ (chl$a$), 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 with a sampling interval of 20 seconds resulted in a nominal spatial resolution of 200 meters.

Quality control flags were defined from (1) sensor reading thresholds on speed, flow rate, hull and line temperatures, sampling water temperature, and (2) data variability, expressed as lower and upper bounds for standard deviation between neighbouring measurements, as described below. Measurements at low speed ($<=$ 5 knots) or zero ship speed are typically collected in harbour and were disregarded. Erroneous records, e.g. caused by instrument communication errors, were removed using a moving window mean filter. A window length of 25 observations (approximately 8.3 minutes) was used for records of ship speed, and a window length of 100 observations (33.3 minutes) 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 temperature and in-line temperature. Flow rates $<=$ 0.3 liter per minute or a temperature difference $>$ 2° Celsius 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 ($\sigma$) of chla fluorescence measurements or GPS-derived latitude were omitted. GPS-derived latitude was additionally filtered for exceptionally high short-term variability ($\sigma$0.5°, window size 50 samples), caused by poor satellite reception or serial communication errors. Table \ref{tab:qc} 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, 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 microscopy counts of phytoplankton species. Laboratory chla concentration results were used to convert transect-normalized chla fluorescence to units of chla concentration (in $\text{milli gram per cubic meter}$). 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 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 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 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 parameter of the two neighbouring sampled transects. These two solutions were then interpolated linearly, weighted by their temporal distance to the respective transect. Negative 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).

The diurnal variability of the fluorescence signal was estimated from quality-controlled photosynthetically active radiation (PAR), sea surface temperature (SST), and wind speed (WIND) derived from the ECMWF ERA-Interim reanalysis data set \cite{Dee2011}. 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 A* 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.

\subsection*{Nutrient Concentration and Depletion Timing}
A single term for nutrient availability was adopted from \cite{Fleming2006}, 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 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 \ref{ssec:algaline}). $\text{nut}$ 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 A* line processing of A* line observations (see section \ref{ssec:bloom_timing}) 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 ($\text{nut-peakvalue}$) 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 to as $\text{nut-peakday}$, $\text{nut-deplday-50}$, and $\text{nut-deplday-25}$. The day and value of the lowest nutrient concentration index are referred to as $\text{nut-minday}$ and $\text{nut-minvalue}$. The rate of nutrient depletion between 75 \% and 25 \% of the peak value ($\text{nut-slope}$) was determined through linear regression.

\subsection*{Extraction of Bloom Timing and Intensity}
Extraction of bloom timing and intensity was carried out for five Baltic Sea areas as described above for the joint data set. Nutrient concentrations and depletion timing are described using the following metrics. The nutrient concentration prior to bloom start ($\text{nut-peakvalue}$) 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 to as $\text{nut-peakday}$, $\text{nut-deplday-50}$, and $\text{nut-deplday-25}$. The day and value of the lowest nutrient concentration index are referred to as $\text{nut-minday}$ and $\text{nut-minvalue}$. The rate of nutrient depletion between 75 \% and 25 \% of the peak value ($\text{nut-slope}$) was determined through linear regression.
Alg@line chla concentrations (see section \ref{ssec:algalineline}) were resampled to daily sea area averages, using linear interpolation, and subsequently smoothed with a 21-day centred-running-mean filter \citep{Ferreira2014,Racault2015} to fill in gaps and reduce short-term variability. We derive several metrics, all of which have in common that the bloom peak concentration \textsc{peakheight}, see Table \ref{tab:params} for explanations of acronyms) and timing \textsc{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 $\SI{5}{\milli\gram\per\cubic\meter}$ was defined as bloom by \citep{Fleming2006}, further referred to as \texttt{const5}. A 21-day centred-running-mean filter was used to keep results comparable to the other metrics \citep{Fleming2006}: considered, whereas \citep{Fleming2006} used a 7-day centred-running-median filter.

2) \citep{Siegel2002} proposed a 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 \texttt{median5}.

3) Distributions proposed to describe bloom phenology include shifted-Gaussian \citep{Platt2009}, Gamma \citep{Vargas2009}, and Weibull distributions \citep{Rolinski2007}. While 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 \cite{Rolinski2007}, 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 $10^\text{th}$ and $90^\text{th}$ percentiles before and after the bloom peak, respectively. This metric is further referred to as \texttt{weibull}.

For each metric, bloom initiation, peak, and end dates \texttt{(startday)}, \texttt{peakday}, and \texttt{endday}) were extracted from the data set. Based on the se dates, bloom duration \texttt{(duration)}, concentration average \texttt{(concavg)}, and the sum of daily chla concentrations \texttt{(bloomidx)} were calculated. The latter was proposed by \citep{Fleming2006} to characterize bloom intensity. We assumed the bloom to have started prior to Alg@line service commen ce if the first data point already satisfied the bloom criterion for a given met ric. Such cases were identified for 30 out of 225 combinations of sea region, ye ar, and bloom metric (9 times for bloom metric \texttt{const5}, 16 times for \texttt{median5}, and 5 times for \texttt{weibull}). Corresponding bloom start days were replaced by the median value for the region over the 15 years studied in a ll subsequent calculations.

\subsection{Principal Component Analysis}
Principal component analysis (PCA) was carried out to attribute seasonally dete nd meteorological conditions \texttt{(sst)}, \texttt{(par)}, \texttt{(wind)} and nutrient concentrations \texttt{(nut-peakvalue)}, \texttt{(nut-minvalue)} to the inter-annual variability in bloom intensity \texttt{(bloomidx)}, \texttt{(concavg)}, \texttt{peakheight}) and timing \texttt{(startday}) and \texttt{peakday}), \texttt{duration}). Outliers were defined for each parameter as departure by more than 3 standard deviations from the parameter mean, and replaced with the region-medi an. Z-score normalization (subtraction of mean, division by standard deviation) was carried out on a per-region basis.\textit{For 30 out of 225 combinations of sea}}
region, year, and bloom metric, ferrybox records started after blooms had initiated. Such cases were identified 9 times for bloom metric \texttt{const5}, 15 times for \texttt{median5}, and 5 times for \texttt{weibull}. Corresponding bloom start days were replaced by the median value for the region over the 15 years studied during subsequent calculation of bloom phenology trends. 

Region-equalized, zero-mean and unit-variance data were then subjected to the PCA function in the python framework scikit-learn \citep{Pedregosa2011}.

Principal component analysis (PCA) was carried out to attribute seasonally detrended, 

The Alg@line ferrybox systems collected over $9.5\times10^6$ observations between 2000 and 2014, of which $3.8\times10^6$ observations were sampled during spring (day-of-year 31 to 160). Availability and rejection rates for each quality control parameter are listed in Table \ref{tab:qc}. 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. This procedure allowed to successfully fit 318 (98 \%) out of the total 324 transects withfor 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 \ref{fig:diurnal_variability}A 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) the effect was less pronounced, although a diurnal variability of 35 and 38 \% is still contained in the respective 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 \ref{fig:diurnal_variability}B 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.

\subsection{Bloom Intensity and Timing}

Bloom generally developed first in the south and progressed towards the north (see Fig. \ref{fig:phenology_geo_timing} and Table \ref{tab:bloomstats}). 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 \texttt{const5} suggested longer blooms in high-biomass sea areas like the \texttt{gof}, compared to low-biomass areas such as the \texttt{sbs}. The variable-threshold metric \texttt{median5} applies area-specific bloom thresholds (\texttt{nbp}: $\SI{3.52}{\milli\gram\per\cubic\meter}$, \texttt{gof}: $\SI{4.95}{\milli\gram\per\cubic\meter}$, \texttt{got}: $\SI{2.51}{\milli\gram\per\cubic\meter}$, \texttt{sbs}: $\SI{2.62}{\milli\gram\per\cubic\meter}$, \texttt{bom}: $\SI{4.02}{\milli\gram\per\cubic\meter}$) and resulted in approximately stable bloom durations for all sea areas. The \texttt{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 ($\textsc{peakheight}$), the chla concentration average during bloom conditions ($\textsc{concavg}$), and the sum of daily chla concentrations over the bloom period ($\textsc{bloomidx}$). Similar patterns were observed for all these parameters and bloom metrics, as illustrated in Fig. \ref{fig:phenology_geo_intensity}. The highest bloom intensity was found in the \textsc{gof} and \textsc{nbp}, followed by the \textsc{bom}. Low-intensity blooms were observed in the \textsc{sbp} and the \textsc{got}. Variability was generally proportional to bloom intensity, highest in the high-biomass and coastal \textsc{gof} and \textsc{bom}. Variability in $\textsc{bloomidx}$ was comparable to that in $\textsc{peakheight}$, while $\textsc{concavg}$ was considerably more stable. All calculated bloom phenology parameters can be found in the supplementary material.

\subsection{Trends}

Figure \ref{fig:trends} shows mean-normalized (subtraction of area-average concentration) $\textsc{concavg}$ and $\textsc{peakheight}$ for all sea areas combined, as a function of bloom year. $\textsc{peakheight}$ is independent of bloom metric and shows a highly significant ($R^2 = 0.12$, $p \ll 0.01$) negative trend of $-0.30(10)\text{mg.chl.m^{-3}.yr^{-1}}$. $\textsc{concavg}$ is dependent on bloom start and end days and was therefore calculated for all applied metrics. Statistically significant, negative trends resulted from all metrics: $-0.12(4)\text{mg.chl.m^{-3}.yr^{-1}}$ for \texttt{const5} ($R^2 = 0.11$, $p \ll 0.01$), $-0.11(5)\text{mg.chl.m^{-3}.yr^{-1}}$ for \texttt{median5} ($R^2 = 0.12$, $p < 0.05$), and $-0.22(7)\text{mg.chl.m^{-3}.yr^{-1}}$ for \texttt{weibull} ($R^2 = 0.11$, $p \ll 0.01$).

No significant trends were found for $\textsc{bloomidx}$, $\textsc{startday}$, and $\textsc{peakday}$ with any of the applied metrics, while $\textsc{endday}$ showed weakly correlated but statistically significant ($R^2 = 0.06$, $p < 0.05$) positive trends for $\texttt{const5}$ and $\texttt{weibull}$ with slopes $0.6(3)$ and $0.17(3)$ day yr$^{-1}$, respectively.

Bloom duration resulting from the $\texttt{weibull}$ metric stands out in the results. Peak nutrient concentrations showed no significant trend, in contrast to post-bloom nutrient concentrations with a highly significant, negative trend $\texttt{nut-deplday-50}$ ($R^2=0.23$, $p \ll 0.01$). Peak nutrient concentration timing shifted to earlier dates ($\texttt{day.per.year}$ ($R^2=0.06$, $p < 0.05$)), while the 25 \%-of-peak-value was reached progressively later ($\texttt{day.per.year}$, $R^2=0.06$, $p < 0.05$). No significant trends were found for nutrient depletion slope, 50 \%-of-peak-value-timing, or day of minimal nutrient concentrations.

\subsection{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). Surprisingly, after applying area-wise mean and variance (z-score) normalization, however, a negative correlation was found for $\textsc{peakheight}$ ($R^2=0.11$, $p \ll 0.01$, metric independent) and $\textsc{concavg}$ ($R^2=0.12$, 0.11, $p \ll 0.01$) for $\texttt{const5}$ and $\texttt{weibull}$, respectively.

Nutrient-depletion timing. The timing of nutrient depletion, specifically $\textsc{nut-deplday-50}$, was positively correlated to the bloom peak day ($R^2=0.47$, $p \ll 0.01$), as well as to bloom-averaged, detrended $\textsc{par}$-levels ($R^2=0.14 - 0.29$, $p \ll 0.01$). Average wind speed and $\textsc{par}$ were negatively correlated during bloom conditions ($R^2=0.10 - 0.23$, $p \ll 0.01$). The bloom timing parameters ($\textsc{startday}$, $\textsc{peakday}$, $\textsc{endday}$) were weakly but statistically significantly inter-correlated (results not shown).

PCA scores and loadings of the first three principal components (PC) are shown a
s biplots in Fig. \ref{fig:pca_biplots}. The first PC is dominated by negative correlations to bloom intensity parameters (\texttt{peakheight}, \texttt{concavg}, \texttt{bloomidx}). This component is positively correlated to pre-bloom nutrient concentration (\texttt{nut-peakvalue}) and bloom duration, illustrating that bloom intensity is affected driven by pre-bloom nutrient availability. The second PC is linked to bloom timing, with strong positive correlations to \texttt{startday} and \texttt{peakday}. Correlations to \texttt{par} (positive), \texttt{sst} (positive), and \texttt{wind} (negative) suggest that weather conditions affect bloom timing. Bloom duration is positively correlated to the third PC, as well as to \texttt{bloomidx}. Additional negative correlations to \texttt{nut-minvalue} and \texttt{wind}, as well as a positive correlation to \texttt{par}, suggest a link between favourable meteorological conditions (low wind mixing, high light level) and efficient nutrient depletion.

PCA scores and loadings of the first three principal components (PC) are shown in Fig. \ref{fig:pca_scores_loadings}. 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 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 conditions on inter-annual variability in bloom phenology.

\subsection*{Automated Processing of Ferrybox Observations}
Thresholds for speed, flow rate, and data variability were iteratively adjusted to the data set and might therefore not apply directly to other ferrybox implementations. Particularly flow [31] and rate, derived from differences in line and hull temperature [32], likely requires tuning to each ferrybox installation. However, 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 statistical 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 fluorescence observations to better represent natural variability.

\subsection*{Variability in Fluorescence Yield}
Fluorescence diurnal variability 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 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 [31] to cyanobacterial summer bloom [32], influenced fluorescence yield considerably less than diel cycles.

The diurnal variability in fluorescence response of 50 \% during an average summ
er day is within the range of earlier findings, e.g. 66 \% (33 \%) for near surface observations in upwelled waters of the equatorial Pacific reported by \cite{Dandonneau1997} or 30 \% for near-surface seaglider observations in Northeast Pacific waters off the Washington coast, USA \citep{Sackman2008a}, although differences in normalization impede direct comparison. The sampling depth of \SI{5}{\meter} for Alg@line systems and the high attenuation of the Baltic Sea in comparison to clear Pacific Ocean waters cause lower fluorescence yield along a transect, which is therefore of lesser relevance for the present study. However, if fluorescence measurements were to be quantitatively evaluated at a higher spatial resolution, variable locally varying fluorescence yield should be accounted for. Analysis of signal-coherence \citep{Grøetsh2014} 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.

\subsection{Spring Bloom Timing and Intensity}

The presented bloom phenology expands the time series presented by \cite{Fleming2006} and is in good agreement for the overlapping period (2000 - 2004) when comparing the metric results. Remaining differences are likely due to quality-control and pre-processing procedures on the fluorescence records. In their work, the authors reported for \textsc{gof}, \textsc{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 \cite{Fleming2006}, e.g. the high-biomass \textsc{bom} and low-biomass \textsc{sbp} and \textsc{got}, followed the reported south-north trend in bloom development. Present results also support and expand the findings of \cite{Fennel1999}, 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.

\subsection{Trends}

Interannual variability in coastal systems exceeds long-term trends by orders of magnitude \citep{Cole2012}. Consequently, trends were observed at relatively low coefficients of correlation. The importance of appropriate data preprocessing and gap handling has been emphasized in literature and is further demonstrated by the present analysis. Robustness of the reported decadal trends is documented by high statistical significance levels ($p \ll 0.01$, Figs. \ref{fig:trend_duration} and \ref{fig:trends}), 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 \citep{Wasmund2003}.

\cite{Helcom2014} 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 \cite{Andersen2015}, based on analysis of 112 years of
f 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 found also for spring bloom, providing further evidence for their hypothesis of gradual nutrient load reduction.

The concentration distribution-ratios on which the Weibull-metric is based are calculated for each bloom individually, in contrast to the thresholds that are fixed for the complete time series (see Fig. Thresholds of \texttt{const5} and \texttt{median5} are calculated individually for each bloom. Therefore, Weibull-metric results for bloom duration are not sensitive to long-term trends in peak concentration. Threshold-based metrics Contrary to this expected behaviour, however, \texttt{const5} and \texttt{median5} revealed no significant trends in bloom duration, while 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 showed a highly significant, positive trend in bloom duration. These two contrasting sets of results nevertheless support 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 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. \cite{Kahru2014} 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 \SI{-0.6}{\day\per\year}. These results suggest that the gap has decreased between dinoflagellate- and diatom-dominated spring bloom and cyanobacterial summer bloom. Due to the shorter period covered here as compared to the time series presented by \cite{Kahru2014}, 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 \citep{Andersen2015} 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.
or marine ecosystem dynamics, and thus eutrophication status. Crucial for deriving bloom timing trends are observations at high temporal resolution and choice of bloom metrics. Eutrophication status assessment frameworks such as HEAT3.0 \cite{Andersen2015} may be adapted to embrace available high-frequency data sources to include bloom timing in their analysis. Ecosystem models of the Baltic Sea and other coastal or inland systems can also use the presented results for validation and to enhance their predictive capabilities.

\subsection*{Environmental Forcing}

Gradually decreasing nutrient concentrations \cite{Helcom2014, Andersen2015}, as well as rising average air- and sea-surface temperatures \cite{Omstedt2004, Helcom2013c} 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 \cite{Duarte2009} but extrapolation of the present results to climate scenarios is beyond the scope of this study. However, 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. Therefore, lacking other 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 hamper 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 \cite{Fleming2006,Meier2011,Neumann2012} for spring, and especially cyanobacterial summer bloom \cite{Wasmund1997,Kanoshina2003,Wynne2010,Wynne2011}.

\conclusions

A Baltic Sea spring bloom phenology was derived from 15 years of automated ferry box chla fluorescence observations. Procedures for automated quality control and processing were introduced and uncertainty due to diurnal variability in phytoplankton fluorescence response was quantified. 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 allow to better disentangle ecosystem response to changing nutrient availability and climate.
atic conditions.

\begin{acknowledgements}
The authors thank the Alg@line consortium, specifically scientists and technical personnel at SYKE (and formerly FIMR), 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 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.
\end{acknowledgements}

\begin{table}
\centering
\begin{tabular}{l||rrrrr}
\hline
& \textbf{Sign} & \textbf{Threshold} & \textbf{Availability [\%]} & \textbf{Rejection Rate [\%]} \\
\hline
Speed, $\si{\knot}$ & $<$ & 5 & 100 & 1.33 \\
\hline
\end{tabular}
\end{table}

\begin{figure}
\centering
\includegraphics[width=8.3cm]{/home/phil/Documents/work/dev/AlgalinePaper/spring_track/plot_scripts/plots/fig04.png}
\caption{Bloom timing (bloom start, peak, and end day) for each sea area along the routes in Figure \ref{fig:transect}, 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 thus plotted separately. The sea areas are ordered by latitude, from south to north.}
\label{fig:phenology_geo_timing}
\end{figure}