General comments

The manuscript provides the first long time remote sensing chlorophyll data for the Baltic Sea. There is a good reason why such analysis have not been undertaken before– simply the chlorophyll-a algorithms do not provide accurate enough concentrations to make such analysis reasonable. It has been shown by several authors (cited in the manuscript) that the blue-green ratios (like the OC4) do not work in the Baltic Sea. Mainly because the reflectance signal in the blue band is not determined by chlorophyll-a, but CDOM. The results of this study confirm these findings as the highest correlation coefficient found for all tested algorithms is R2=0.44. Therefore, the whole following analysis seems a bit artificial to me.

An important element that characterizes the new version of the manuscript is the clear focus on the operational CMEMS context in which the paper builds and develops. One of the main issues with previous CHL validation works over the Baltic (not many and all cited within the manuscript) is linked to the used datasets that were limited in both the geographic extent and in their seasonal representativeness, casting doubts on their overall relevance: most of these only use data from a few cruises. As it is clearly mentioned in the revised manuscript this constitutes one of the main motivations for this work, which uses the largest in situ dataset ever used for CHL algorithm calibration and validation over the Baltic region. Moreover, this is also the first work providing statistics for the entire Baltic, using the merged multi-sensor CCI and GlobColour products both covering a timeframe of fifteen years.

As for the blue-to-green ratio algorithms, it is true that this approach has demonstrated to perform poorly into the Baltic, though tested with limited datasets. It is also true that the ocean colour Baltic product currently provided in the context of CMEMS is far from being optimal (it is the worst of the proposed algorithms in all areas), as clearly shown by this analysis for the first time. All the Reviewer's concerns about the reliability of blue-to-green ratio algorithms (like the OC4) over the Baltic Sea are now included into the manuscript to increase the readers' awareness on this topic.

Moreover, the results presented here about the phytoplankton phenology in the Baltic are in line with what is expected from the literature with the Central Baltic blooming twice a year, in spring and summer (Reissmann et al., 2009), the Gulf of Bothnia only during spring (Carstensen et al., 2015) and the Skagerrak and Kattegat showing a minimum in Summer (Edelvang et al., 2005).

The current version of the manuscript takes account of all the comments from the Reviewer. Moreover, the comments from both Reviewers stimulated a thorough and careful re-reading of the entire manuscript. We have more strongly stressed the focus of the paper, which is worth reminding develops in the operational context of CMEMS. Main limitations of this work are now mentioned and discussed in the manuscript.

The entire manuscript went through a general re-editing:

- *The abstract and Introduction have been nearly totally rewritten.*
- *Section Data (now Data and methods) now includes a new section with the description of the statistical analysis (formerly present at the beginning of the results).*
- *Similarly, the structure of section Results (now Results and discussion) has been improved by including a forth subsection (Algorithm regional calibration) to better ease the reading.*
- *Conclusions should now read less cumbersomely as more general statements did replace lots of the many summarizing and purely technical and repetitive sentences.*

All figures have been edited.

- *Figure 1 (formerly showing the matchup spatial distribution only) now includes the space-temporal matchup distribution along with the in situ CHL frequency distribution.*

- *Figures 2 and 3 about the scatterplot of satellite versus in-situ CHL now include a qualitative colour legend and regression lines for each plot.*
- *Figure 4 now includes a panel to show more clearly the variability and robustness of the regression lines obtained within the bootstrapping exercise.*
- *In Figure 5 the zero is now part of one bin of the histogram, while formerly was the boundary between two bins.*
- *Former figure 6 on the full time series of CHL has been removed because the discussion was and still is only focused on the climatological values (current figure 6) and the supplementary material already contained detailed figures of each CHL yearly time series. On the other hand former figure 8 on the full time series of SST have been moved to the supplementary material. Both figures 6 and 8 were not adding any specific information while potentially distracting the reader.*
- *A brand new figure 7 has been added to show that the adopted algorithms can also be used to monitor the space-time variability of the cyanobacteria in the Central Baltic, consistent with previous findings (Kahru and Elmgren, 2014).*
- *In figure 8 only the years with corresponding SST data are shown.*

The specific issues raised by the Reviewer are addressed below (in italic).

Detailed comments

Being familiar with unpublished (yet) results from different countries around the Baltic Sea it is hard to say whether universal chlorophyll algorithm for the Baltic Sea is feasible. Studies on the specific inherent optical water properties show that optical properties of the spring bloom assemblages are very different compared to cyanobacteria in summer. Therefore, two sets of chlorophyll retrieval algorithms may be needed. These are results for the open parts of the Baltic Sea. Other studies show that absorption and backscattering coefficients (determining the reflectance) differ by order of magnitude between rocky granite shores (Sweden, Finland) and sandy shores (Russia, Estonia, Latvia, Lithuania, Lithuania, Poland, Denmark). Third study (also unpublished) shows that the correlation between the OC4v6 and chlorophyll is close to zero even if reflectances created by HydroLight model (i.e. free from atmospheric correction problems, glint, sensor noise, etc.) are used. So, it is really hard to believe that the algorithm used in this study will ever produce reasonable results for the Baltic Sea. In the conclusions the Authors state themselves (like tens of authors before) that green to red bands have to be used in order to get reasonable chlorophyll-a estimates.

It is surely true that one single algorithm for the entire domain (this in theory applies to the global ocean as well) might provide unrealistic observations given the phenological heterogeneity of the area. It is also true, however, that different algorithms meant to capture both the space and time variability (as the Reviewer is suggesting) would be extremely difficult to merge operationally. Here, as it should now be more clear, we aim at evaluating algorithms for the whole Baltic area, while still trying to provide region-specific metrics for the Skagerrak-Kattegat, the Central Baltic and for the gulf of Bothnia. Obviously, this fragmentation could be continued ad infinitum, but in the context of the operational oceanography, it is very unlikely that provided CHL products will reach such level of geographic and seasonal distinction in the near future.

On page 8 the Authors discuss problems related with vertical distribution of phytoplankton biomass. During most of the year this should not be an issue as top 10-20 m is mixed. However, vertical distribution becomes a huge issue during the period of cyanobacterial dominance. Unlike other phytoplankton cyanobacteria can regulate their buoyancy and (in calm weather) tend to be at the depth most optimal for them. It has been shown before (Kutser et al. 2008) that vertical distribution of cyanobacteria has significant impact on the reflectance i.e. the same biomass distributed differently in the water column produces very different reflectance.

The discussion on the phytoplankton vertical distribution is meant to provide details on the method used to convert a vertical profile into a "surface" observation. And the purpose of the method is exactly aimed to address the issue raised by the Reviewer. We now added a sentence to better stress the purpose of this particular bit of work.

Findings in the page 15 are contradictory to what was proposed by Kahru et al. (1993). Not the elevated temperature causes blooms (how it can be elevated?) but bloom absorbs solar radiation and heats the water.

We expanded a bit more on the motivations for this part of the analysis. The Reviewer's comment has been included into the main text and helped us expanding a little more on the motivations for this analysis. Moreover, as explicitly asked by the other Reviewer, we added a cross-correlation analysis between CHL and SST anomalies over the Central Baltic during summer to better support what was in the previous version a more qualitative discussion.

I obviously cannot agree with the last conclusion that the analysis provides a good confidence level about ocean colour retrieval over the Baltic Sea.

The sentence has been removed.

I recommend the Authors to read biological literature about the phenology of the Baltic Sea phytoplankton, what kind of concentrations of chlorophyll-a have been actually observed in the Baltic Sea during different bloom periods and how this matches/contradicts with their findings. There is plenty of literature available for the Baltic Sea.

We did follow the Reviewer's suggestion and incorporated findings from the literature into the manuscript, and found that the ranges of CHL variability described in this work are fully consistent with those from the literature. Analogously, as mentioned above, the tentative description of the phytoplankton phenology sketched in this work is coherent with that available from the literature for the three regions.

Cyanobacterial blooms in the Baltic Sea have been known for large mats (scum) floating on the water surface. These mats may be several centimetres thick and cover areas of 200 000 km2 (Kahru and Elmgren 2014). These issues have not been discussed at all. If standard processing chains are used then the scum pixels are masked out as errors whereas the algorithms cannot cope with "terrestrial" reflectance (high NIR) in the middle of the sea. If these pixels are masked out then what kind of chlorophyll dynamics we discuss here at all?

This comment has been included into the main text and motivated the new figure 7, which confirms (after Kahru and Elmgren, 2014) that even standard algorithms can be used for the monitoring of phytoplankton biomass into the Baltic. Moreover, Kahru and Elmgren (2014) in their in depth analysis of the cyanobacteria blooms explicitly mention that only a small (<5%) fraction of the bloom pixels is affected by atmospheric correction failure, due to surface scum.

Remote sensing of chlorophyll in the Baltic Sea at basin scale from 1997 to 2012 using merged multi-sensor data

Jaime Pitarch(1),*, *Gianluca Volpe*(1), *Simone Colella*(1), *Hajo Krasemann*(2), *Rosalia Santoleri*(1)

(1) Institute for Climate and Atmospheric Sciences, Italian National Research Council. Via del Fosso del Cavaliere 100, I-00133 Rome, Italy.

(2) Helmholtz-Zentrum Geesthacht, Centre for Materials and Coastal Research GmbH. Max-Planck-Strasse 1, D-21502 Geesthacht, Germany.

*Corresponding author: jaime.pitarchportero@artov.isac.cnr.it, phone +39 06 4993 4313

Keywords: Baltic Sea; chlorophyll; remote sensing; ocean colour; multisensor; algorithms; in-situ data; calibration; validation; time series; phytoplankton phenology

Abstract

A fifteen-year (1997-2012) time series of chlorophyll-a (CHL) in the Baltic Sea, based on merged multi-sensor satellite data were analysed. Several available CHL algorithms were sea-truthed against the largest in-situ publicly available CHL dataset ever used for calibration and validation over the Baltic region. To account for the known biogeochemical heterogeneity of the Baltic, matchups were calculated for three separate areas: (1) the Skagerrak and Kattegat, (2) the Central Baltic, including the Baltic Proper and the gulfs of Riga and Finland, and (3) the Gulf of Bothnia. Similarly, within the operational context of the Copernicus Marine Environment Monitoring Service (CMEMS) the three areas were also considered as a whole in the analysis. In general, statistics showed low linearity. However, a bootstrapping-like assessment did provide the means for removing the bias from the satellite observations, which were then used to compute basin average time series. Resulting climatologies confirmed the three regions to display completely different CHL seasonal dynamics. The Gulf of Bothnia displays a single CHL peak during spring, whereas in the Skagerrak and Kattegat the dynamics is less regular and made of highs and lows during winter towards a small bloom in spring and a minimum in summer. In the Central Baltic, CHL follows a dynamics of a mild spring bloom followed by a much stronger bloom in summer. Surface temperature data are able to explain a variable (with years) fraction of the intensity of the summer bloom, in the Central Baltic.

1. Introduction

Global to regional monitoring of the surface ocean is believed to be an essential element for the sustainability of the ocean resources. In Europe, the Ocean Colour (OC) Thematic Assembly Centre (TAC) is the entity devoted to produce and provide ocean colour remote sensing data; and this is performed in the context of the Copernicus Marine Environment Monitoring Service (CMEMS). OC data are currently provided at both global and regional scales. These two scales refer to both the geographical limits and the algorithms used to process the data. The OCTAC is thus meant to provide an added value by not only zooming the data from the global domain to the single regional European seas, but also and especially for the application of tailored *ad hoc* regional algorithms for chlorophyll (CHL) retrieval. The present work aims at assessing the performance of existing CHL algorithms for operational applications over the Baltic Sea. CHL is routinely measured over the world oceans with two main kinds of algorithms: i) those using the blue-to-green reflectance ratio (e.g., empirical) and ii) the semi-analytical, e.g., those using the inherent optical properties to infer the chlorophyll concentration. The former builds on the common experience that water colour spans from blue to green as CHL increases, in open ocean (Case I waters). The latter are mathematically more complex and thus based on a larger number of assumptions; nevertheless, they are believed to be more suited for optically complex waters (known as Case II waters) where the colour of the ocean is determined by several non-covarying constituents, such as CHL, Coloured Dissolved Organic Matter (CDOM) and non-algal particles. Both types of algorithms are very sensitive to the in-situ observations used to calibrate them, thus providing the motivation of the regionalization approach adopted within CMEMS. Those based on neural network constitute a third kind of algorithms for CHL retrieval whose limitations are, in theory, the same as the first two: strong dependency on the training datasets that limit their overall applicability. Here, all three kinds of algorithms are tested.

The Baltic Sea is a semi-enclosed basin bordering the North Sea in correspondence of the Danish archipelago. Skagerrak and Kattegat are generally not associated with the Baltic Sea. However, the Baltic domain that is defined within CMEMS extends the eastern limit to the meridian 9.24 °E, thus including most of the Skagerrak and the entire Kattegat basins. The Baltic is characterized by significant CDOM concentration due to high river runoff. It is known that high CDOM concentration reduces the water-leaving radiance making the seawater darker (Berthon and Zibordi, 2010), and this constitutes one of the main challenges for ocean colour algorithms to work properly over the Baltic Sea (Mélin and Vantrepotte, 2015). Despite the fact that the Baltic Sea is widely recognized as a challenging test bed for remote sensing, literature on calibration and validation of CHL algorithms is not abundant. Standard algorithms are those provided by the space agencies for global and operational applications. The application of these algorithms to the in-situ Remote Sensing Reflectance (R_{rs}) collected in 707 stations off Poland between 1993 and 2001 revealed uncertainties exceeding 100% when the output was compared against collocated CHL measurements (Darecki and Stramski, 2004). Even less encouraging results were obtained when four standard CHL algorithms were applied to Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) images between 2000 and 2001 (HELCOM, 2004). Matchup with 75 CHL profiles across all the Baltic Sea, with predominance of Swedish coastal waters, gave virtually null correlation, with satellite CHL underestimating the in-situ CHL by 180% to 500%, in contradiction with Darecki and Stramski (2004). More recently, the Case II Regional, Boreal, and Eutrophic MERIS processors were applied to images between 2006 and 2009 (Attila et al., 2013). Matchup with 312 stations in the Gulf of Finland and in the central Baltic Sea showed large CHL overestimation. However, when the standard bio-optical relationships of these processors were tuned with the local in-situ CHL, the bias did reduce significantly (Fig. 6 in Attila et al., 2013). The heterogeneity of results combined with the limited spatial and temporal representativeness of the in-situ observations used in the mentioned data comparisons prompts for further investigation. This work aims at filling this gap by using the largest and publicly available in-situ dataset ever used over the Baltic Sea for validation activities.

There is an extensive literature on the biogeochemistry of the Baltic Sea and its relation to physics. River outflows release large amounts of organic matter, which sinks to the bottom and lowers the oxygen concentration leading to large amounts of phosphate to be released by the sediment and upwelled through complex mixing processes (Reissmann et al., 2009). In spring, a nutrient-enriched *hypolimnion* and warmer temperatures trigger diatom and dinoflagellate blooms, depleting nitrogen. In summer, nitrogen-fixing cyanobacteria take advantage of the relatively phosphate-rich, calm and warm surface waters, producing another bloom (Reissmann et al., 2009). The Central Baltic Sea is characterized by summer blooms of cyanobacteria that are known to have buoyancy regulation ability (e.g., N. Spumigena and Aphanizomenon sp., Ibelings et al., 1991) and that, under calm conditions, can accumulate at the sea surface (Ploug, 2008). Cyanobacteria blooms are commonly observed in the central Baltic Proper but not in the Skagerrak and Kattegat nor in the Gulf of Bothnia (Wasmund and Uhlig, 2003). Skagerrak and Kattegat are subject to much higher influence from the North Sea, so that the phytoplankton dynamics, here, is expected to be different than that of the Baltic Sea. Thus there is a strong need for the calibration and validation of the proposed algorithms to take account of the complex morphology and biogeochemistry of the basin. Algorithms are then tested in four geographical areas: (1) Skagerrak and Kattegat, (2) the Baltic Proper and the gulfs of Riga and Finland, here referred to as “Central Baltic”, (3) the Gulf of Bothnia, and (4) the entire area (1 to 3).

Ocean colour has cloud cover as additional problem, which is particularly high over northern Europe. To increase the spatial coverage of daily products, the International Ocean-Colour

Coordinating Group (IOCCG) recommended the merging of ocean colour data from multiple missions (IOCCG, 2007). At European level, the Climate Change Initiative (CCI) program (www.esa-oceancolour-cci.org) and the Globcolour (www.globcolour.info) project followed this recommendation and reprocessed archived data from various medium-resolution sensors. Here, the CCI-derived R_{rs} are used as input to the CHL algorithms for the comparison exercise (see section 2.1 for their description). One of the limitations of this approach is given by the fact that the CCI does not include any bands in the near-infrared, which are known to be better suited than the blue-green for Case II waters (Odermatt et al., 2012). On the other hand, merged data spans for longer time periods (1997-2012) than any of the individual sensors alone and provide higher coverage on a daily basis.

Applications of remote sensing in the Baltic Sea have been mainly focused on a few main topics: cyanobacteria blooms (Reinart and Kutser, 2006), light penetration (Pierson et al., 2008) and management of various coastal areas (Kratzer et al., 2008). A good overview of such different applications can be found in Siegel and Gerth (2008). Long-term multisensor satellite data were recently used to develop an indicator of surface cyanobacteria accumulation over defined Baltic regions for trend analysis (Kahru et al., 2007;Kahru and Elmgren, 2014). However, long-term phytoplankton dynamics over the entire Baltic region is still lacking, despite the fact that this is required by the European Water Framework Directive for coastal and inland waters and by the Marine Strategy Framework Directive for open ocean waters. In this article, we aim to partially fill this gap by focusing on long-term remote sensing of CHL at basin scale.

2. Data and methods

2.1 Satellite CHL data

Table **1** summarizes the four satellite CHL products evaluated in this article with their respective references.

Acronym	Input	CHL Algorithm	Reference
GLC	Rrs from single sensors	GSM	(Maritorena and Siegel, 2005)
OC4v6	ESA-CCI Rrs	OC4v6	(Werdell, 2010)
OC5	ESA-CCI Rrs	OC5	(Gohin et al., 2002)
MLP	ESA-CCI Rrs	MLP	(D'Alimonte et al., 2011)

Table 1: summary of the algorithms used in the validation analysis with the acronym used in this work along with the required input for each of them. GLC stands for GlobColour, OC4v6 for Ocean Colour four bands algorithm (version 6), OC5 for Ocean Colour five bands, and MLP for Multi-Layer Perceptron.

The GlobColour dataset (GLC hereafter) was developed in the framework of the European Space Agency Data User Element program to support global carbon cycle research. Daily GlobColour data were downloaded from the project web site (www.globcolour.info). Products are obtained by merging MERIS, MODIS, SeaWiFS and VIIRS data. Validation at global scale was carried out by Maritorena et al. (2010). Downloaded data are 2nd reprocessing Level 3 binned images (L3b), having a resolution of 1/24 of degree at the equator (i.e., around 4.63 km) and consisting of the accumulated data of all merged level 2 products, corresponding to periods of one day. Merged data are generated by the GSM model (Maritorena and Siegel, 2005), which also produces the CHL parameter, delivered as product named CHL1. CHL1 parameter is meant to provide best performances over case I waters and thus it is not recommended for use over optically complex waters, but no alternative is given for the Baltic Sea (further details in the Product User Guide, GlobColour, 2015).

The ESA Ocean Colour CCI program has the goal to provide stable, long-term, multisensor satellite products. The dataset consists of the merged SeaWiFS, MODIS, and MERIS data, by shifting MODIS and MERIS R_{rs} to the SeaWiFS wavebands, before merging (ESA-OC-CCI, 2014). Data are mapped at 4 km resolution and are available through the OC-CCI (www.oceancolour.org) and the CMEMS portals (marine.copernicus.eu). Standard CHL products are global-ocean daily mean sea surface CHL. ESA-CCI retrieves CHL through the application of the OC4v6 algorithm (O'Reilly et al., 2000;Werdell, 2010) to the merged R_{rs}. The dataset available from CMEMS also includes an additional CHL product by applying the OC5 algorithm (Gohin et al., 2002), developed as an adaptation of the OC4 to French Atlantic coastal waters (further details in the Product User Manual, CMEMS, 2015). Calibrated R_{rs} are also available for the application of custom algorithms. We used these R_{rs} to test a Baltic Sea-specific CHL algorithm, available for the SeaWiFS bands, developed by D'Alimonte et al. (2011). This algorithm is based on a Multi-layer perceptron (MLP) and was trained with in-situ R_{rs} and CHL. MLP was only validated with in-situ R_{rs} and CHL (D'Alimonte et al., 2012), thus not taking into account all the known issues linked to the atmospheric correction over the basin.

An image pre-analysis revealed ~15 % more flagged (invalid) pixels for MLP than for OC4v6 and OC5, despite all are derived from the same CCI reflectances. The cause is the frequent occurrence of negative $R_{rs}(412)$ most likely due to aerosol optical thickness overestimation in the blue together with high CDOM. In contrast, OC4v6 does not use $R_{rs}(412)$, the most sensible band to the atmospheric correction procedure, thus allowing for problematic pixels (those with $R_{rs}(412)<0$) to be retrieved as well. Similarly, OC5 is insensitive to negative $R_{rs}(412)$ values, thus allowing CHL to be retrieved also under the extreme conditions of atmospheric correction failure.

2.2 In-situ CHL data

We downloaded publicly available in-situ CHL data, contained in the repositories at Seadatanet (www.seadatanet.org), the Baltic Marine Environment Protection Commission (www.helcom.fi) and the NOAA World Ocean Database (www.nodc.noaa.gov/OC5/WOD/pr_wod.html). CHL is computed from bottle samples using standard laboratory techniques. The technique used to collect and measure CHL spans from fluorimetry to spectrophotometry and HPLC. The amount of information provided depends upon each environmental agency or research institution that collected and uploaded the data. For their part, data repositories have additional quality control criteria based on outlier estimation. All data collected in the Baltic region during the period covered by the satellite observations (1997-2012) were merged and duplicates were eliminated.

Since the remote sensing signal can be fairly considered as a weighted average within the first optical depth, in-situ observations must be treated accordingly. In-situ CHL consisted either of a single sub-surface reading or CHL profiles derived from a few depth readings. In this latter case, a proper vertical average of a CHL profile is needed for comparison to remote-sensing data. The vertical weighting function depends on the inherent optical properties (IOPs) that cannot be inferred solely from CHL in case II waters. In rigor, coincident IOP measurements are needed to perform the vertical averaging, but such measurements are scarce and not publicly available. In case I waters, vertical averaging can be performed with the method by Morel and Berthon (1989) with input CHL profile data. The remaining applicable options to our in-situ data were either to select only the sub-surface CHL value or to average the profiles with the method by Morel and Berthon (1989), despite the theoretical inconsistencies. Calculations (not shown) revealed that satellite-in-situ correlations did improve (even if only slightly) if available profiles were vertically averaged (and this is the approach used in this work) instead of taking only the uppermost reading. To avoid bottom contribution to the water-leaving radiance, only stations with a bottom depth of at least 10 m were selected.

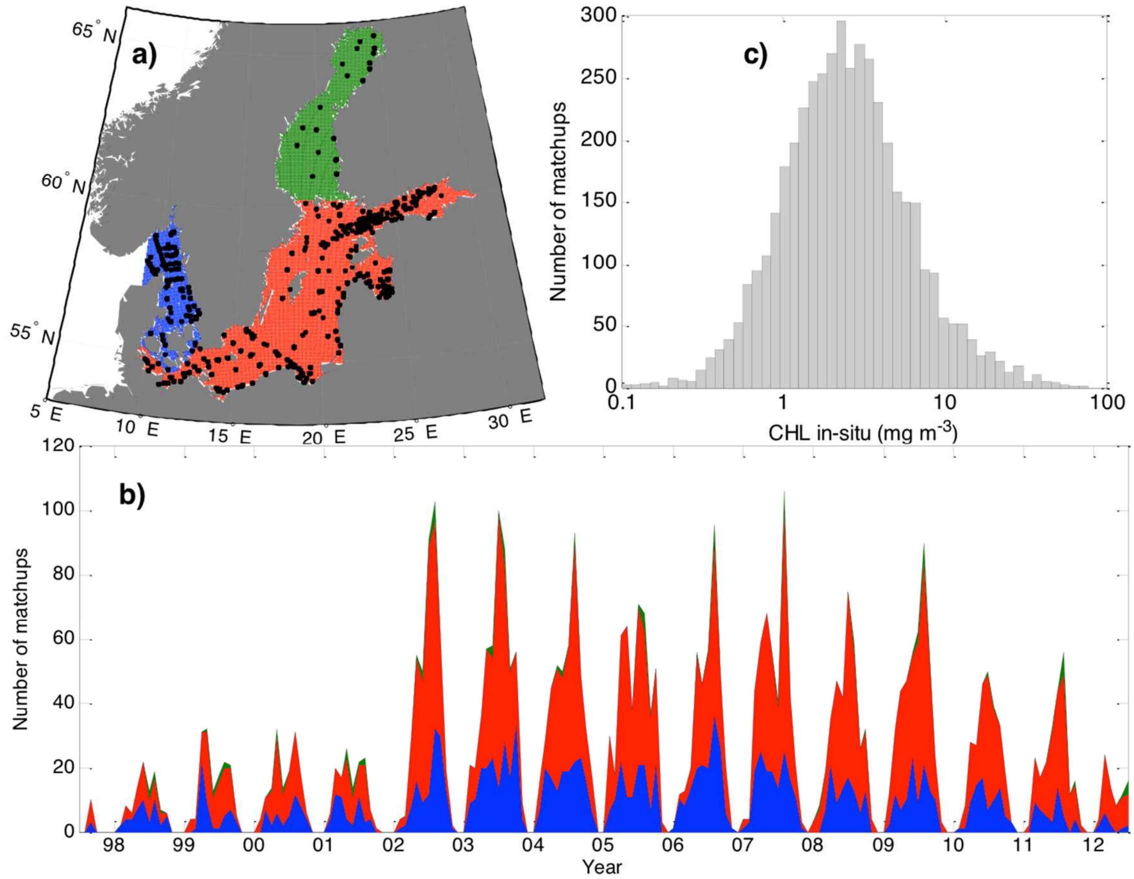

Fig. 1: a) Spatial distribution of the 4492 in-situ stations used in the matchup analysis (see section 3.1) along with the partition of the area of study. Skagerrak and Kattegat is highlighted in blue with 1456 matchup points. Central Baltic is highlighted in red with 2922 matchup points, and the Gulf of Bothnia is green with 114 stations. Temporal station distribution is also shown using the same colour code (b). The frequency distribution of the entire in-situ CHL is shown in panel c.

Similarly, to ensure representativeness of the data in the case of CHL stratification, only stations with the uppermost reading shallower than 2 m were retained for the analysis. The spatial location of matchup stations is shown in Fig. 1a. Although covering the entire Baltic region, data are not uniformly distributed, as the dataset is built from different sources, in which individual institutions and agencies are interested in specific zones. The number of matchups increases significantly when both MODIS-Aqua and MERIS started to be operational in 2002 (Fig. 1b), thus providing further evidence of the utility of merging different sensors for oceanographic research. The CHL in-situ dataset used in the following of this work is log-normally distributed around the mean value of ~ 2.46 mg m^{-3} and spanning from 0.1 mg m^{-3} to 77 mg m^{-3} (Fig. 1c). Fleming and Kaitala (2006) reported CHL values 7 to 12 mg m^{-3} in the northern Baltic Proper during the spring bloom. Our gathered in-situ matchup dataset during April in the northern Baltic Proper (35 samples) shows CHL to range from 1.39 to 14.7 mg m^{-3}, consistent with these previous findings.

2.3 Statistical evaluation

Satellite CHL was extracted from single pixels without further spatial windowing. To calculate the mean bias and the RMS we applied decimal logarithm-transformation to the CHL data, and returned to percentage linear scale:

$$\mathrm{BIAS} = \left[10^{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)} - 1 \right] \cdot 100 \tag{1}$$

$$\mathrm{RMS} = \left[10^{\frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_i - x_i)^2}} - 1 \right] \cdot 100 \tag{2}$$

where x_i and y_i are the $\log_{10}$-transformed in-situ and satellite CHL, respectively. N is the number of matchups. The best linear fits were found using the log-transformed CHL. The corresponding coefficient of determination (R^2), slope (m) and intercept (n) are also presented. The whole area was divided into regions with expected bio-optical differences (see Fig. 1a). The number of observations available from the Gulf of Bothnia is very limited, so the statistical information that can be derived from the regressions must be interpreted with caution. Nevertheless, results are presented for completeness. The p-value of the regressions was zero for all regions except for the Gulf of Bothnia, where it was $p<10^{-3}$.

Outliers were defined as data in which any of the four algorithms gave CHL outside the range within one twentieth and twenty times the in-situ CHL. In applying this criterion, roughly 3.5 % of the data were discarded and led N to become 1873. Most of these discarded matchups were rejected because of the GLC underestimation, together with the high scattering (Fig. 2a). The discarded data were evenly distributed over the entire range of CHL variability and without any specific temporal or spatial patterns. For comparison issues among algorithms, only matchups with coincident valid pixels for all four satellite products within the same day were considered, but once the best performing algorithm was identified, all available matchup stations for this algorithm were used to provide its full record of statistics (N = 4492).

3. Results and discussion

3.1 Matchups

In general, satellite and in-situ data show modest agreement in the Baltic. This can be intuitively associated with both the non-full traceability of the methods used to assemble the in-situ dataset and the satellite algorithms. MLP and GLC provide poor R^2 and negative BIAS with respect to the in-situ data. Results of OC4v6 (R^2=0.43) are consistent with findings by Darecki and Stramski (2004). The positive bias of 44 % here (Fig. 2b) is smaller than 119 %, as found by Darecki and Stramski (2004), but still confirms the OC4v6 to overestimate CHL in the Baltic Sea. OC4v6 matches better the in-situ data for high CHL, whereas tends to saturate for low values. OC5 has similar linearity (R^2=0.44) but significantly improves in terms of bias (-14 %) with respect to OC4v6. Besides the similar R^2, we appreciated graphical similarities between the scatter plots of OC4v6 and OC5. Guided by this hint, we performed a linear regression in log form between OC4v6 and OC5 satellite derived CHL (not shown). Regression analysis revealed a very high linear dependence (R^2=0.97), although the relationship is more complex in theory (Gohin et al., 2002), and this will have implications for the rest of this work (see below).

Geographical partition of the matchup dataset highlighted significant differences in the statistical behaviour of algorithms. For instance, the performance of MLP strongly degrades in Skagerrak and Kattegat (Fig. 2h) with respect to the central Baltic Sea (Fig. 2l). MLP was calibrated with data only inside the Baltic Sea, and not in the Skagerrak and Kattegat (D'Alimonte et al., 2012, Fig. 2d). It appears then that such algorithm design is highly dependent on the calibration data. GLC performs always worst in all regions, and the scatter plots seem like undefined clouds, which is best highlighted by the large RMS errors. OC4v6 displays similar statistics at both sides of the Danish Strait, although the slope of the regression line decreases towards Skagerrak and Kattegat. In each region, OC4v6 overestimates CHL by more than 40 %. The behaviour of OC5 is always in accordance with OC4v6, with a shifted BIAS, given the very high correlation between the two. Due to the much simpler applicability of OC4v6 and its wider diffusion in the community, the following analysis will be based on the OC4v6.

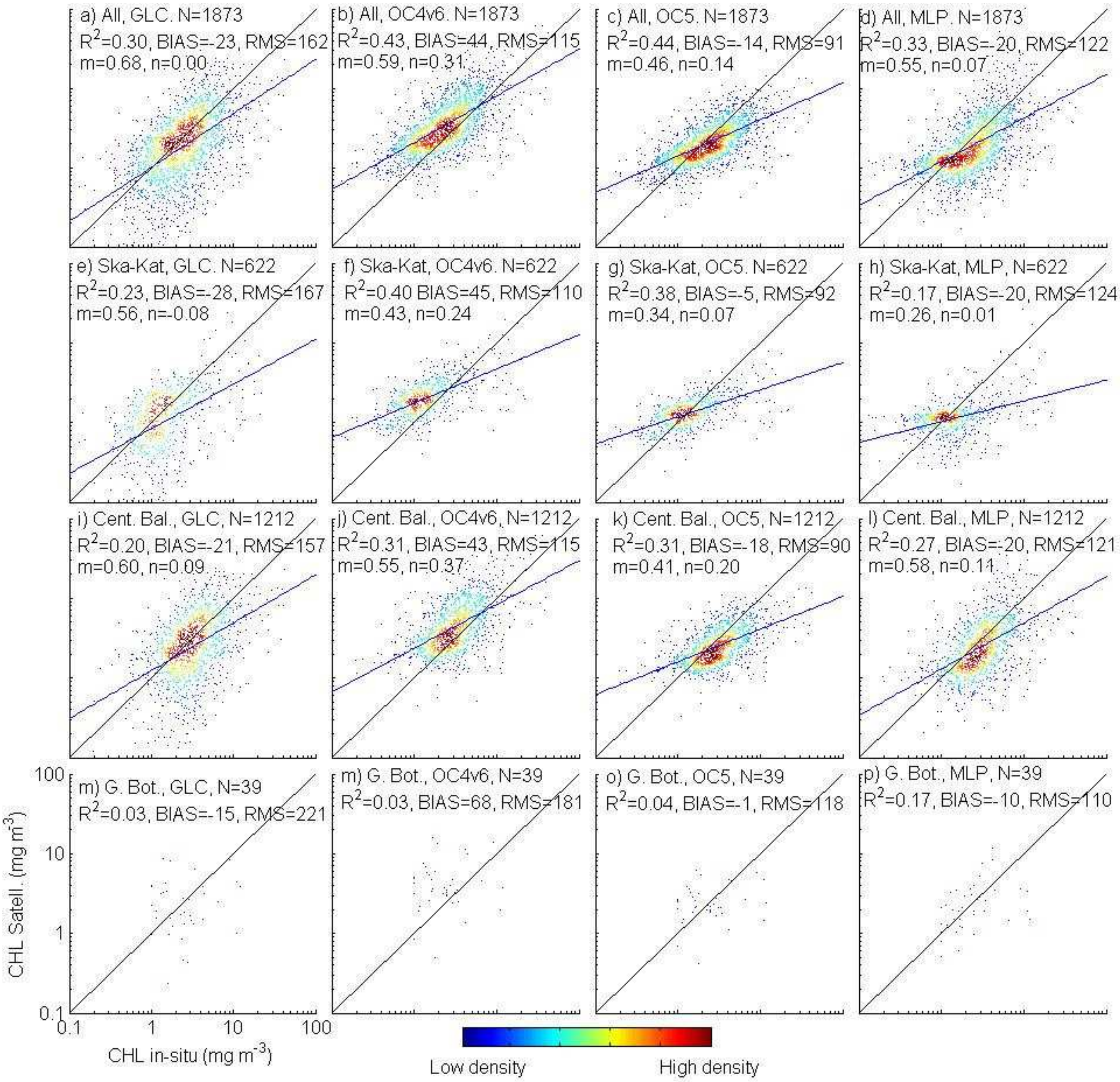

Fig. 2 Density scatter plots of in-situ versus satellite-retrieved CHL for all algorithms providing meaningful values. The line of best fit (blue) and that of equal value (black) are superimposed together with relevant statistics.

The matchup analysis is here repeated with the same conditions, including the definition and removal of the outliers, but now for OC4v6 alone. Only 22 matchups were discarded (< 0.5 % of the data), of which 17 due to overestimation (i.e., higher than twenty times the in-situ counterpart). As mentioned, when the coincidence with the other algorithms is removed, the number of matchups increases to 4492, distributed as 1456 in Skagerrak and Kattegat, 2922 in the Central Baltic and 114 in the Gulf of Bothnia. Fig. 3 shows the corresponding density scatter plots and statistics. In general, the interpretation from Fig. 2 still holds, with the bigger size of the matchup dataset providing increased confidence level of the derived statistics. However, since the additional data were previously discarded (not used to produce Fig. 2), it is not surprising that the latter statistics did degrade (R^2 = 0.43, BIAS = 72%, RMSE = 151%, m = 0.57, n = 0.41, N = 2619). The orders of magnitude of the uncertainties found here (Fig. 3) are in line with those available from the literature (Darecki and Stramski, 2004) even considering the wider space and time distribution of the data (both in-situ and satellite) used here.

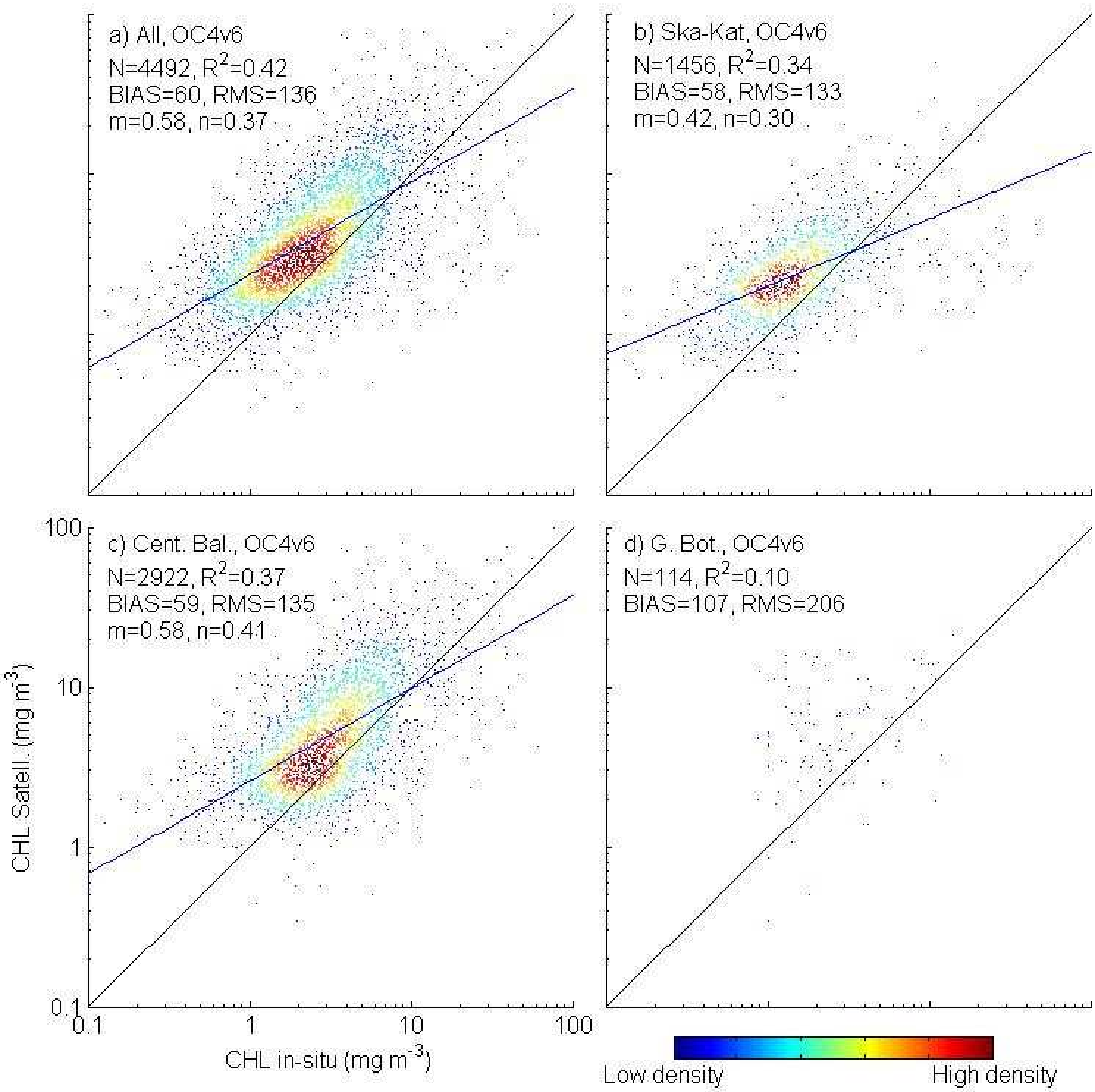

Fig. 3 Density scatter plots of in-situ versus satellite-retrieved CHL for OC4v6 algorithm. The best linear regression (blue) and the line of equal value (black) are superimposed along with relevant statistics.

3.2 Validation

When the regression coefficients are used to re-calibrate existing algorithms, the validity and robustness of the matchup statistics needs to be validated against independent data. Starting from the matchups for OC4v6 alone (Fig. 3a), we performed a sensitivity study to test the dataset homogeneity by a bootstrapping-like assessment (Efron, 1979) as used in recent validation exercises (Brewin et al., 2013). The whole dataset (N = 4492) was partitioned a thousand times into two randomly chosen halves: calibration (N_{cal} = 2246) and validation (N_{val} = 2246). Each calibration dataset is used to compute the linear regression coefficients (m,n) which are then applied to the corresponding complementary validation half to compute the associated statistics. The obtained series of coefficients and statistics are shown in Fig. 4. Results are remarkably robust: the averages of the regressions found (m=0.5843, n=0.3657, green dashed line in Fig. 4) are almost equal to those when the whole dataset is used (m=0.5845 and n=0.3656, red line in Fig. 4). Moreover, the dispersion is very small with the coefficient of variation being 2.07 % when computed over the slopes and 1.38 % over the intercepts. As for the validation statistics, their mean values (graphically shown in green in Fig. 4) R^2 = 0.4236, BIAS = 59.55 %, RMS = 136.13 % are very similar to those obtained for the whole dataset (Fig. 3a, R^2 = 0.4241, BIAS = 59.53 %, RMS = 136.19 %).

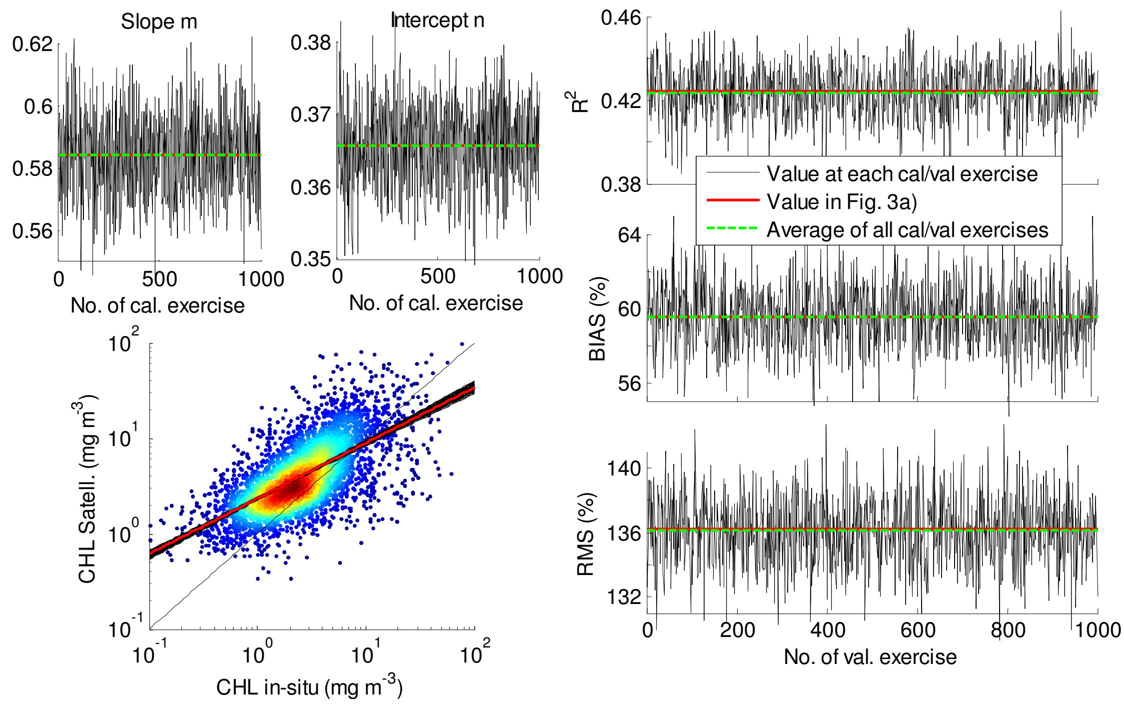

Fig. 4 Left up, in black: best linear fits (slope m and intercept n) of 1000 randomly chosen calibration datasets (N_{cal} = 2246, X axes) of $\log_{10}(CHL_{in\text{-}situ})$ versus $\log_{10}(CHL_{OC4v6})$. Left down:

application of all 1000 (m,n) pairs to the OC4v6 vs. in-situ scatter cloud. In red, slope and intercept for the whole dataset, as shown in Fig. 3a. In green, average of the 1000 calibration results. Right, in black: statistics when applying each m and n pairs from the left side to the complementary validation datasets (N_{val} = 2246, X axes). These are: coefficient of determination, BIAS (eq (1) and RMS (eq. (2)). In red: same statistics found for the whole dataset, as shown in Fig. 3a. In green, average of the 1000 validation results.

3.3 Algorithm regional calibration

Efficient and useful algorithm regionalization needs appropriate bio-optical in-situ data. Unfortunately, the Baltic lacks of such publicly available in-situ dataset that therefore prevents a canonical regionalization. This, together with the high confidence level associated with the described statistics and in view of obtaining an unbiased proxy for CHL, with the available data, prompt for using the computed coefficients (m and n in Fig. 4) for recalibrating the OC4v6, as follows:

$$\log_{10}\left(\mathrm{CHL}_{\mathrm{OC4v6corr}}\right)=\frac{\log_{10}\left(\mathrm{CHL}_{\mathrm{OC4v6}}\right)-n}{m} \qquad (3)$$

Errors between eq. (3) and the complementary in-situ validation matchups were calculated. Each of the 1000 chosen combinations generated a vector of errors with length N_{val} = 2246. Their accumulation led to a total of 2246000 error estimates, whose distribution is shown in Fig. 5, together with the fitted Gaussian curve. The recalibration via eq. (3) removed the bias, resulting in a zero-centred error distribution. It is worth reminding that within the calibration and validation exercises the two datasets are independent. The standard deviation (σ = 0.4582) includes all errors not taken into account by the system, i.e. atmospheric noise, errors in the in-situ measurements and, most of all, the limited suitability of algorithms as the OC4v6.

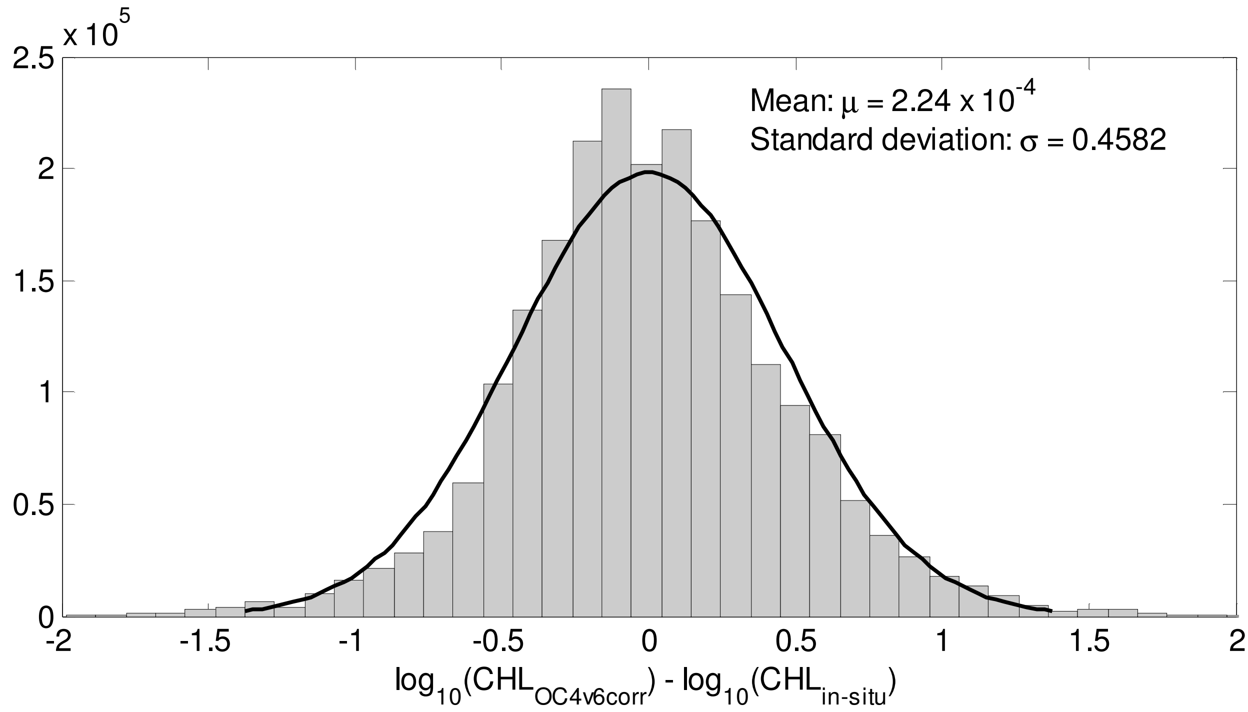

Fig. 5 Histogram of the absolute error between OC4v6$_{corr}$ and in-situ CHL, both in logarithmic form. Associated mean and standard deviation are also shown and used to compute relevant fitted Gaussian distribution (black line).

The symmetric and zero-centred error distribution (Fig. 5) obtained with the application of eq. (3) within the bootstrapping-like assessment warrants a high level of confidence when basin averages are calculated; all the errors at the level of individual pixels are expected to cancel out when a horizontal (pixel-wise) average is performed over a large region. Although the former statement implies that the statistical properties of the matchup dataset can be extrapolated to the whole Baltic area, the good spatial and temporal coverage of the former (see Fig. 1) is a good asset to support this argument. From this point, we defined the algorithm OC4v6$_{corr}$ through eq. (3), with the coefficients (m = 0.5884, n = 0.3751) of Fig. 3a. This enabled the bias to be removed. Nevertheless, RMS was altered, rising up to 187 %, in agreement with σ = 0.4582 in Fig. 5 through eq. (2). The mathematical explanation of the latter relationship is that the RMS and the standard deviation of a zero-mean distribution are equal.

Among all regions in which the Baltic Area has been divided, Fig. 3 highlights different best linear fits. Given the coefficients of variation 2.07 % and 1.38 % for the slope and intercept respectively found in the bootstrapping assessment, the coefficients in Fig. 3 are significantly different. If OC4v6 is linearly adjusted with eq. (3), the coefficients must be different for each region, in particular, equal to those found in Fig. 3. Therefore, for Skagerrak and Kattegat, they were set to 0.4212 and 0.3027, respectively for m and n. Due to the lack of enough data, the stations in the Gulf of Bothnia were aggregated to those of the Central Baltic. Resulting statistics for these two regions were almost equal to those of the Central Baltic alone: R^2 = 0.35, BIAS = 60.45 %, RMS = 138.64 %, m = 0.5632, and n = 0.4206. These linear coefficients were applied to recalibrate OC4v6 for the Central Baltic and the Gulf of Bothnia. Even if the same algorithm was used results are presented separately for the two basins.

3.4 Satellite derived basin averages

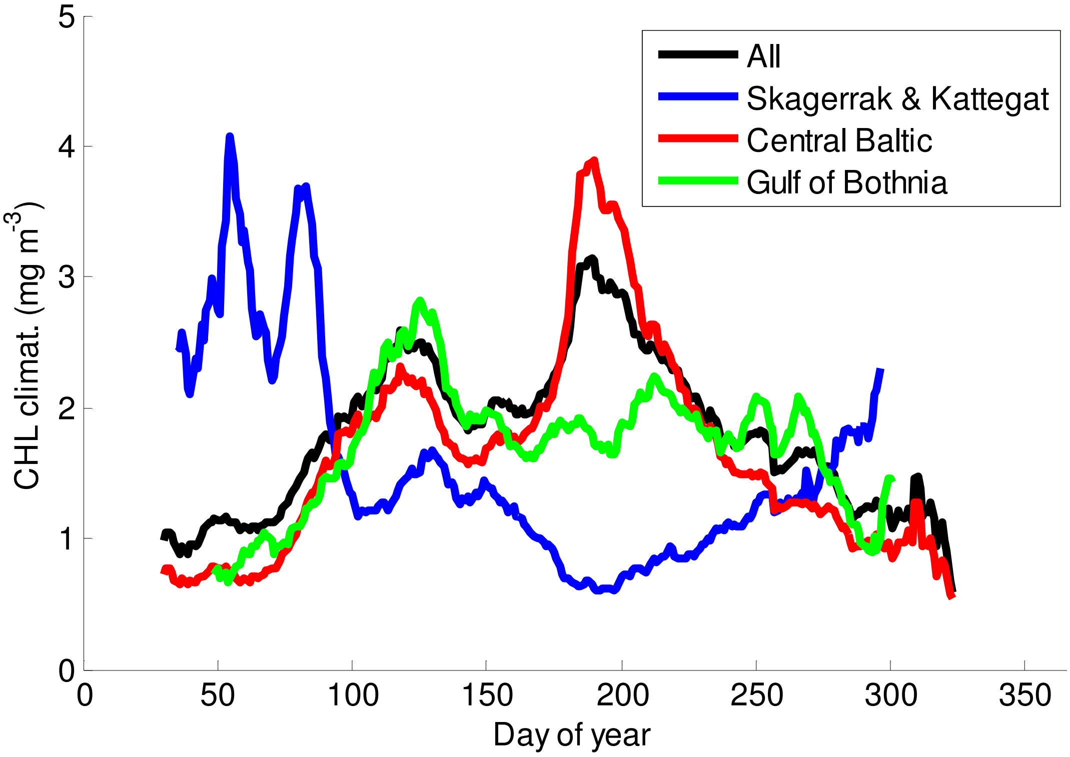

Fig. 6 CHL daily Climatology. For any given day of the year, the average was computed only if data for a minimum of six years were available. Plots of individual time series with their associated standard deviation bars can be found in the supplementary material. To improve the plot readability, all time series were smoothed with a one-week moving average.

Horizontally-averaged CHL for $OC4v6_{corr}$ were computed only for images with a minimum number of 1000 valid pixels. The entire Baltic has 21424 pixels, with the Gulf of Bothnia contributing with 5750 pixels, Skagerrak and Kattegat with 2625 pixels and the Central Baltic with 13049 pixels. One thousand pixels correspond to 5 %, 17 %, 38 % and 7 % of their respective surfaces. CHL dynamics strongly varies among regions at both seasonal (Fig. 6) and interannual time scales (supplementary material). In Skagerrak and Kattegat, the dynamics consists of intermittent growth periods in late winter (up to ~ 4 mg m^{-3}) and a small bloom in spring, reaching a minimum in summer (~ 0.5 mg m^{-3}), consistent with other works (Edelvang et al., 2005) . In the Gulf of Bothnia, the overall range of CHL variability is limited to ~ 2 mg m^{-3} (0.7 – 2.8 mg m^{-3}) with minima in winter and a series of bloom-like pulses from spring to fall. The spring bloom is the most intense and lasts longer than the others (Carstensen et al., 2015). Given the prolonged winter darkness, the length of this data time series is shorter than those from the other regions. Moreover, in winter the Gulf of Bothnia is normally ice-covered and some ice remains in the northern part until May, thus not the entire domain contributed to the displayed CHL. A non-trivial point is that this time series has to be interpreted with caution due to lack of a significant number of data for specific calibration in this area. In the Central Baltic, the dynamics is completely different. Two distinct CHL maxima are appreciable (Reissmann *et al*., 2009): the first one peaks at the end of April, reaching ~ 2.5 mg m^{-3}, which is at the lower end of the variability previously observed by Schneider et al. (2006); the intensity of the second peak, in mid-July, (~ 4.6 mg m^{-3}) is consistent with previous observations in the area (Schneider et al., 2006), and from which it steadily decreases and reaches a minimum in winter. The dynamics of the entire domain (black line in Fig. 6) is clearly dominated by the Central Baltic due to its major weight in terms of area coverage. The summer bloom that occurs in the Central Baltic is known to be due to cyanobacteria taking advantage of the milder weather conditions and of the increased water temperature. As cyanobacteria can form surface scum, it is worth questioning whether such data would be masked during the operational image processing. A previously documented mild cyanobacteria bloom on the 11^{th} of July, 2010 was visible from space via qualitative RGB image, and for which surface accumulation was not reported (SMHI, 2010). To assess whether the standard processing is able to provide reliable observations also in these conditions, MODIS-Aqua Level-1A was downloaded and processed to L2 using the same settings used to produce the CCI input data. Fig. 7a shows the Central Baltic blooming also in the areas identified as cyanobacteria by the SeaDAS Level-2 flag TURBIDW (Fig. 7b) used to discriminate the accumulation of cyanobacteria (Kahru and Elmgren, 2014). During summer 2005, the Baltic experienced the second largest cyanobacteria bloom (Kahru and Elmgren, 2014) that covered 25% of the entire domain (183000 km^2). As for the 2010 bloom and apart from the small area classified as too bright in the north Baltic Proper (in light grey in Fig. 7c and 7d), the standard processing demonstrated its ability to provide valid data also under these conditions. Therefore, the data used here appear suitable for the study of phytoplankton dynamics throughout the year, even during cyanobacteria bloom events, during which only a negligible percentage of pixels is affected by atmospheric correction failures (Kahru and Elmgren, 2014).

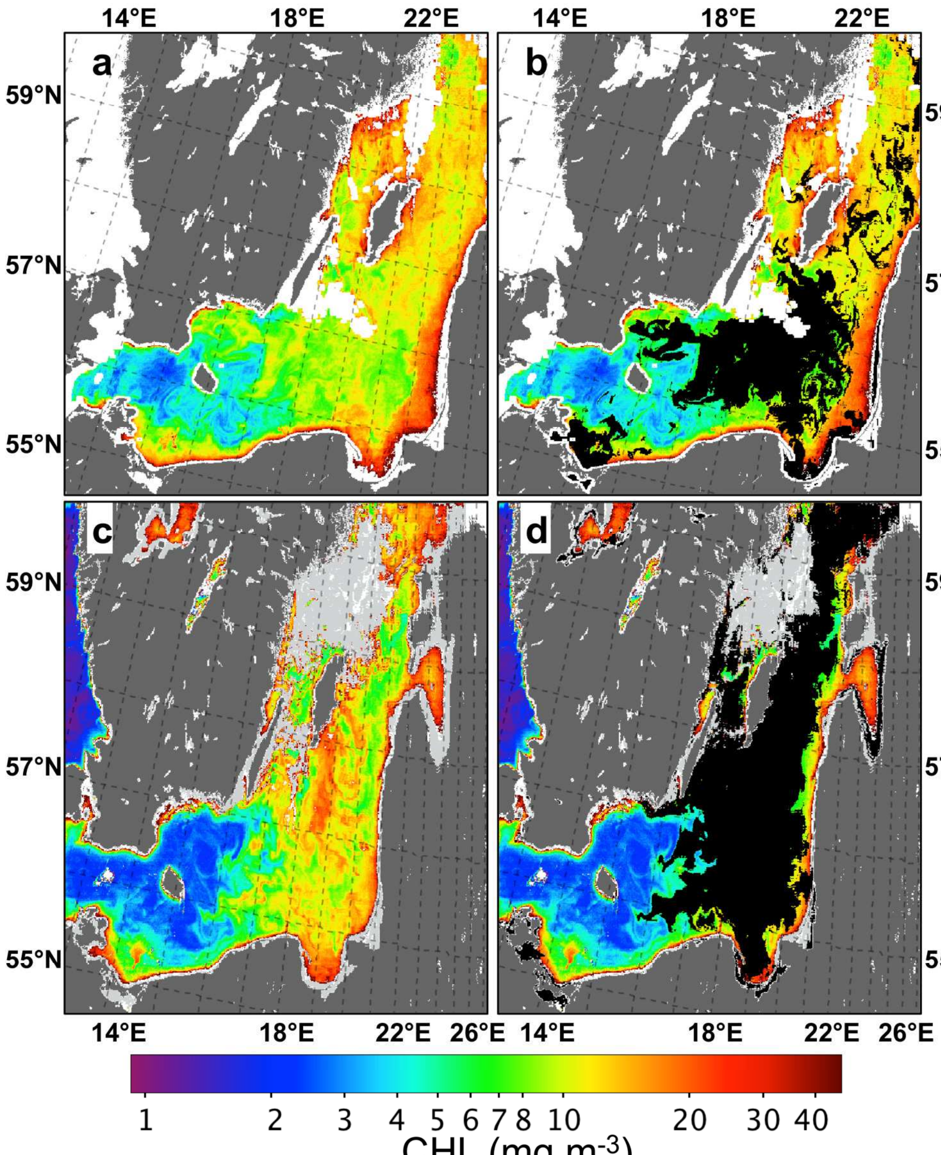

Fig. 7: MODIS Level-1A of the 11th of July, 2010 (a and b) and 2005 (c and d) were downloaded from the OBPG website (Ocean Biology Processing Group, oceancolor.gsfc.nasa.gov) and processed to Level-2 using the standard settings within SeaDAS version 7.3 (seadas.gsfc.nasa.gov). Kahru and Elmgren (2014) recently identified the presence of cyanobacteria accumulating on the sea surface using the SeaDAS Level-2 flag TURBIDW (“Turbid water”) when the flag MAXAERITER (“Maximum Aerosol Iterations”) is turned off within the Level-1 to Level-2 processing. Here, CHL images without (panels a and c) and with (panels b and d) the application of the TURBIDW flag is shown; pixels affected by TURBIDW are coloured black. As mentioned by Kahru and Elmgren

(2014), the MAXAERITER flag is, by default, turned on within the NASA standard processing (e.g., the same used here); light grey area (panels c and d) in the north-western Baltic Proper is perceived by the operational processing as too bright (i.e., masked as MAXAERITER) and as such not processed.

Fig. 6 shows that the strongest signal in the Central Baltic is given by the summer bloom. Cyanobacteria-like species are known to bloom under warm and calm weather conditions (Ploug, 2008). High sea surface temperature (SST) are known to enhance the growth of cyanobacteria, both directly through higher growth rates, and indirectly by increasing the stability of the water column to allow cyanobacteria to take advantage of their buoyancy regulation ability (Ibelings et al., 1991). Analogously, cyanobacteria were demonstrated to provide positive feedbacks to the surface temperaure by absorbing the incoming radiation (Kahru et al., 1993). It is then reasonable to investigate whether CHL and SST covary over the Central Baltic during summer. In the specific context of this cross-correlation analysis, we are implicitly assuming that both SST and CHL respond to the calm weather conditions with the same time lag. For this matter, daily-averages SST data (1998-2009) over the Baltic Sea were downloaded from the CMEMS website. The SST dataset is the merged product from the sensors AVHRRs (series 7, 9, 11, 14, 16, 17, 18), Envisat ATSR1 and ATSR2, and the AATSR (see CMEMS (2015) for details and Supplementary Material for their basin-average time series). Both CHL and SST data time series were deseasonalized by computing the anomalies with respect to their climatologies, which were used as input for the cross correlation analysis. Fig. 8 shows the two time series anomalies along with correlation values computed over the summer period (between the Julian days 150 and 250) for all years for which SST was available. Prior to the correlation analysis, the CHL anomaly time series was further smoothed with a one-week moving average. Here, the basic underlying assumption is that warm waters, as proxy of calm weather conditions, can explain the dynamics of cyanobacteria. Thus when cyanobacteria do represent a high fraction (in terms of their space and time presence) of the CHL signal, the correlation is expected to be high, and vice versa.

Fig. 8 shows quite a surprising relationship between both quantities with high-amplitude SST correlating with those of CHL. This related behaviour is somewhat unexpected, because we are comparing here not absolute CHL and temperature, but their differences with respect to their climatological values. Generally, during the second half of the time series, from 2003 on, the correlation appears to be tighter. The causes of the dynamics shown are undoubtedly complex involving considerations on the circulation and the peculiar biogeochemistry of the basin (Reissmann et al., 2009). Nevertheless, this article is focused on the remote sensing aspect and the intensity of the cyanobacteria bloom appears to depend on the timing of the summer temperature peak: although 2004 had a high SST peak, such peak happened late in the season (August 10th), which appeared not favourable for cyanobacteria growth. On the contrary, years 2002, 2003, 2005 and 2006 had SST peaks of similar or lower intensity, but much earlier in the season. Instead, 2001 displayed two marked positive SST anomalies that were only mildly followed by CHL anomalies. Despite the CHL and SST anomalies are poorly correlated during 1998 (Fig. 8), they are both negative suggesting that in that year the cyanobacteria bloom, generally dominating the summer signal in the Central Baltic, was only partially contributing to the overall dynamics. This is clearly documented in Kahru and Elmgren (2014), who found the Fraction of Cyanobacteria Accumulations of only 6%, in 1998; FCA being the ratio of the number of pixels classified as cyanobacteria to the number of cloud-free sea-surface views during the period July to August.

On the other hand, the year 2008 was completely anomalous with respect to both the climatology value and timing of the summer bloom, with a maximum at the beginning of May. This massive and early bloom has already been documented (Majaneva et al., 2012;Larsson et al., 2014), with the dominant species being *Prymnesium polylepis*. Responsible abiotic factors were exceptionally calm and sunny weather during October 2007, resulting in high light availability and low turbulence above the thermocline (Majaneva et al., 2012;Larsson et al., 2014). These conditions enabled *P. polylepis* to build up a considerable biomass. The following winter was the mildest since more than a century, which allowed *P. polylepis* to persist throughout the winter. Improving weather and plenty of nutrients allowed further growth until a maximum in spring.

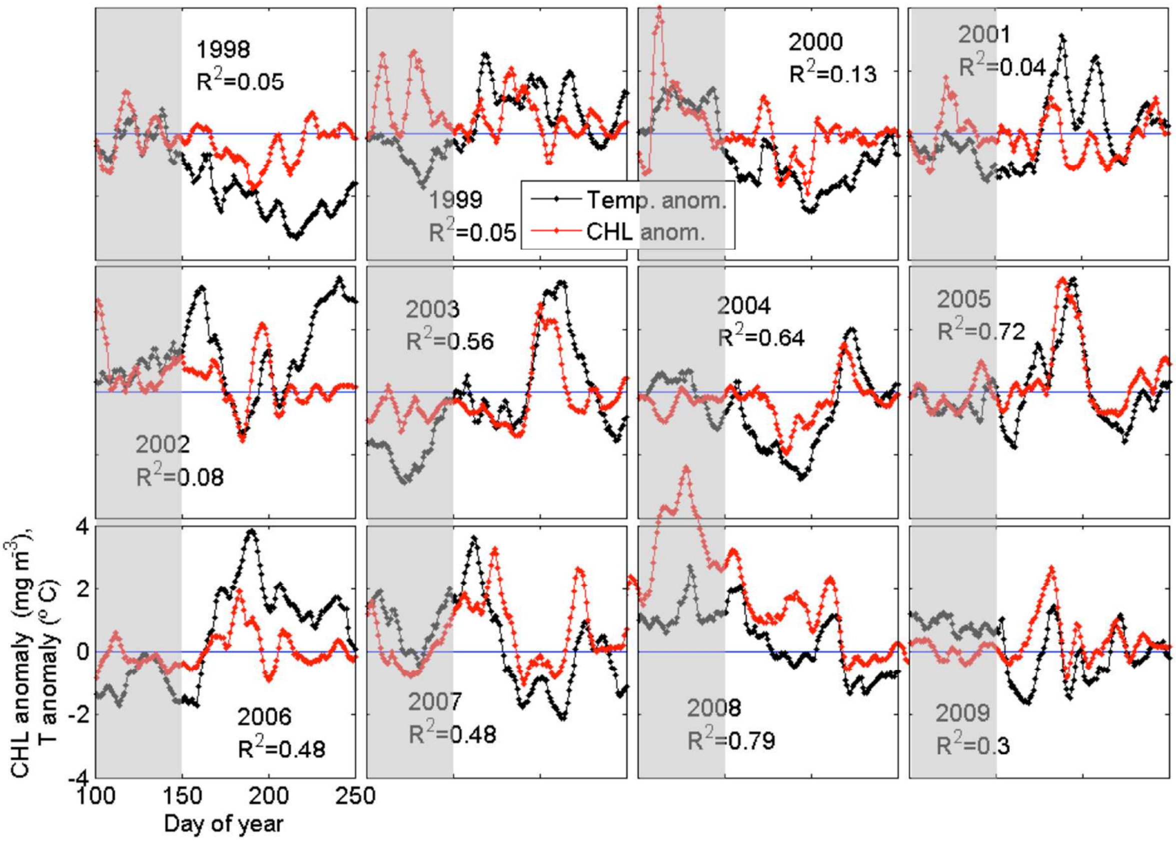

Fig. 8 Time series of the CHL and SST anomalies with respect to their climatologies, over the Central Baltic. The reference value 0 is also displayed. Shaded areas indicate the part of the time series not used for the computation of the cross-correlation coefficient, which is indicated on each year. Full size plots of individual years can be found in the supplementary material.

4. Conclusions

A fifteen-year merged-multi-sensor-daily dataset of satellite-derived CHL contains very valuable information for ecological studies if information is properly processed. Matchup analysis was undertaken with the largest in-situ database ever used for calibration and validation purposes over the Baltic region. Standard algorithms demonstrated easy to apply but, in the Baltic Sea, required further adjustments before an unbiased estimation of the basin-average CHL was obtained. Our derived time series take advantage of the independence of the error added by other water constituents and additional sources. The error distribution of the CHL estimates is such that, when averaging over a large number of observations, tends to zero, thus demonstrating that more accurate observations can be achieved when averaging over large areas.

The $OC4v6_{corr}$-derived climatology in Skagerrak and Kattegat revealed strong productivity in winter and a rather inactive summer. However, it should be noted that the blue-green CHL algorithms are not optimal for the coccolithophore detection (Gordon et al., 2001), commonly observed in this area. In the Gulf of Bothnia, CHL exhibits a single bloom during spring and experiences lower variability than the Skagerrak and Kattegat regions or the Central Baltic. In the latter region, the productivity in late fall, winter and early spring is severely inhibited. A first growth period with a maximum at the end of April is detected, followed by a stronger summer bloom peaking at the second week of July. The summer bloom in the Central Baltic constitutes the most intense signal found in this work, and attributed to cyanobacteria-like species. CHL and SST anomaly time series were cross-correlated to assess the cyanobacteria contribution to the overall CHL dynamics during the summer period of the Central Baltic. For example, the exceptionally warm winter 2007/2008 triggered an intense spring bloom in 2008 that also altered the normal dynamics throughout the year.

The Baltic region is widely recognized as a challenging test bed for ocean colour remote sensing. The interfering CDOM at blue wavelengths suggests that better CHL algorithms should use red and NIR bands, like the fluorescence line height or the maximum chlorophyll index algorithms (Odermatt et al., 2012, Fig. 1). Most of the Baltic CHL values range between ~ 1 and 10 mg m^{-3} and are at the lower part of the retrievable concentrations, via these algorithms (Odermatt et al., 2012, Fig. 1). These algorithms are only applicable to the archived MERIS data (2002-2012). The Ocean and Land Colour Instrument, on-board the Sentinel-3 will provide continuity with MERIS and algorithms will be adapted. The addition of the 400 nm band will expectedly aid in the separation of the CDOM contribution, given that proper atmospheric correction is achieved.

5. Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme through HORIZON 2020 under grant agreement no. 210129802 (Copernicus Marine environment monitoring service). Seadatanet, HELCOM and NOAA along with all single contributors are thanked for the in-situ data and CMEMS, ESA-CCI and GlobColour for the satellite data. Vega Forneris is thanked for technical support and Vittorio Brando for suggestions on the manuscript. Two anonymous reviewers are thanked for their comments and suggestions.

6. References

Attila, J., Koponen, S., Kallio, K., Lindfors, A., Kaitala, S., and Ylöstalo, P.: MERIS Case II water processor comparison on coastal sites of the northern Baltic Sea, Remote Sensing of Environment, 128, 138-149, http://dx.doi.org/10.1016/j.rse.2012.07.009, 2013.

Berthon, J.-F., and Zibordi, G.: Optically black waters in the northern Baltic Sea, Geophysical Research Letters, 37, n/a-n/a, 10.1029/2010GL043227, 2010.

Brewin, R. J. W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., Doerffer, R., Fomferra, N., Franz, B., Grant, M., Groom, S., Horseman, A., Hu, C., Krasemann, H., Lee, Z., Maritorena, S., Mélin, F., Peters, M., Platt, T., Regner, P., Smyth, T., Steinmetz, F., Swinton, J., Werdell, J., and White Iii, G. N.: The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms, Remote Sensing of Environment, In press, http://dx.doi.org/10.1016/j.rse.2013.09.016, 2013.

Carstensen, J., Klais, R., and Cloern, J. E.: Phytoplankton blooms in estuarine and coastal waters: Seasonal patterns and key species, Estuarine, Coastal and Shelf Science, 162, 98-109, http://dx.doi.org/10.1016/j.ecss.2015.05.005, 2015.

Product User Manual for Baltic Sea Physical Reanalysis Products: http://marine.copernicus.eu/documents/PUM/CMEMS-OC-PUM-009-ALL.pdf, 2015.

D'Alimonte, D., Zibordi, G., Berthon, J. F., Canuti, E., and Kajiyama, T.: Bio-optical algorithms for European seas: Performance and applicability of neural-net inversion schemes, Joint research Centre, IspraJRC66326, 2011.

D'Alimonte, D., Zibordi, G., Berthon, J.-F., Canuti, E., and Kajiyama, T.: Performance and applicability of bio-optical algorithms in different European seas, Remote Sensing of Environment, 124, 402-412, http://dx.doi.org/10.1016/j.rse.2012.05.022, 2012.

Darecki, M., and Stramski, D.: An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea, Remote Sensing of Environment, 89, 326-350, http://dx.doi.org/10.1016/j.rse.2003.10.012, 2004.

Edelvang, K., Kaas, H., Erichsen, A. C., Alvarez-Berastegui, D., Bundgaard, K., and Jørgensen, P. V.: Numerical modelling of phytoplankton biomass in coastal waters, Journal of Marine Systems, 57, 13-29, http://dx.doi.org/10.1016/j.jmarsys.2004.10.003, 2005.

Efron, B.: Bootstrap methods: another look at the jackknife, The annals of Statistics, 7, 1-26, 1979.

Product User Guide: http://www.esa-oceancolour-cci.org/?q=webfm_send/318, 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, 10.1007/s10750-005-1006-7, 2006.

Product User Guide: http://www.globcolour.info/CDR_Docs/GlobCOLOUR_PUG.pdf, 2015.

Gohin, F., Druon, J. N., and Lampert, L.: A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters, International Journal of Remote Sensing, 23, 1639-1661, 10.1080/01431160110071879, 2002.

Gordon, H. R., Boynton, G. C., Balch, W. M., Groom, S. B., Harbour, D. S., and Smyth, T. J.: Retrieval of coccolithophore calcite concentration from SeaWiFS Imagery, Geophysical Research Letters, 28, 1587-1590, 10.1029/2000GL012025, 2001.

HELCOM: Thematic Report on Validation of Algorithms for Chlorophyll a Retrieval from Satellite Data in the Baltic Sea Area, Helsinki Commission-HELCOM, Ispra94, 2004.

Ibelings, B. W., Mur, L. R., and Walsby, A. E.: Diurnal changes in buoyancy and vertical distribution in populations of Microcystisin two shallow lakes, Journal of Plankton Research, 13, 419-436, 10.1093/plankt/13.2.419, 1991.

IOCCG: Ocean-colour data merging, IOCCG, Dartmouth, Canada6, 2007.

Kahru, M., Savchuk, O. P., and Elmgren, R.: Satellite measurements of cyanobacterial bloom frequency in the Baltic Sea: interannual and spatial variability, Marine Ecology Progress Series, 343, 15-23, 10.3354/meps06943, 2007.

Kahru, M., and Elmgren, R.: Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea, Biogeosciences, 11, 3619-3633, 10.5194/bg-11-3619-2014, 2014.

Kratzer, S., Brockmann, C., and Moore, G.: Using MERIS full resolution data to monitor coastal waters — A case study from Himmerfjärden, a fjord-like bay in the northwestern Baltic Sea, Remote Sensing of Environment, 112, 2284-2300, http://dx.doi.org/10.1016/j.rse.2007.10.006, 2008.

Larsson, K., Hajdu, S., Kilpi, M., Larsson, R., Leito, A., and Lyngs, P.: Effects of an extensive Prymnesium polylepis bloom on breeding eiders in the Baltic Sea, Journal of Sea Research, 88, 21-28, http://dx.doi.org/10.1016/j.seares.2013.12.017, 2014.

Majaneva, M., Rintala, J.-M., Hajdu, S., Hällfors, S., Hällfors, G., Skjevik, A.-T., Gromisz, S., Kownacka, J., Busch, S., and Blomster, J.: The extensive bloom of alternate-stage Prymnesium polylepis (Haptophyta) in the Baltic Sea during autumn–spring 2007–2008, European Journal of Phycology, 47, 310-320, 10.1080/09670262.2012.713997, 2012.

Maritorena, S., and Siegel, D. A.: Consistent merging of satellite ocean color data sets using a bio-optical model, Remote Sensing of Environment, 94, 429-440, http://dx.doi.org/10.1016/j.rse.2004.08.014, 2005.

Maritorena, S., d'Andon, O. H. F., Mangin, A., and Siegel, D. A.: Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues, Remote Sensing of Environment, 114, 1791-1804, http://dx.doi.org/10.1016/j.rse.2010.04.002, 2010.

Mélin, F., and Vantrepotte, V.: How optically diverse is the coastal ocean?, Remote Sensing of Environment, 160, 235-251, http://dx.doi.org/10.1016/j.rse.2015.01.023, 2015.

Morel, A., and Berthon, J.-F.: Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications, Limnology and Oceanography, 34, 1545-1562, 1989.

Odermatt, D., Gitelson, A., Brando, V. E., and Schaepman, M.: Review of constituent retrieval in optically deep and complex waters from satellite imagery, Remote Sensing of Environment, 118, 116-126, http://dx.doi.org/10.1016/j.rse.2011.11.013, 2012.

Pierson, D. C., Kratzer, S., Strömbeck, N., and Håkansson, B.: Relationship between the attenuation of downwelling irradiance at 490 nm with the attenuation of PAR (400 nm–700 nm) in the Baltic Sea, Remote Sensing of Environment, 112, 668-680, http://dx.doi.org/10.1016/j.rse.2007.06.009, 2008.

Ploug, H.: Cyanobacterial surface blooms formed by Aphanizomenon sp. and Nodularia spumigena in the Baltic Sea: Small-scale fluxes, pH, and oxygen microenvironments, Limnology and Oceanography, 53, 914-921, 10.4319/lo.2008.53.3.0914, 2008.

Reinart, A., and Kutser, T.: Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea, Remote Sensing of Environment, 102, 74-85, http://dx.doi.org/10.1016/j.rse.2006.02.013, 2006.

Reissmann, J. H., Burchard, H., Feistel, R., Hagen, E., Lass, H. U., Mohrholz, V., Nausch, G., Umlauf, L., and Wieczorek, G.: Vertical mixing in the Baltic Sea and consequences for eutrophication – A review, Progress in Oceanography, 82, 47-80, http://dx.doi.org/10.1016/j.pocean.2007.10.004, 2009.

Schneider, B., Kaitala, S., and Maunula, P.: Identification and quantification of plankton bloom events in the Baltic Sea by continuous pCO2 and chlorophyll a measurements on a cargo ship, Journal of Marine Systems, 59, 238-248, http://dx.doi.org/10.1016/j.jmarsys.2005.11.003, 2006.

Siegel, H., and Gerth, M.: Optical Remote Sensing Applications in the Baltic Sea, in: Remote Sensing of the European Seas, edited by: Barale, V., and Gade, M., Springer Netherlands, Dordrecht, 91-102, 2008.

A mild algal bloom in 2010: http://www.smhi.se/en/news-archive/a-mild-algal-bloom-in-2010-1.12999, 2010.

Wasmund, N., and Uhlig, S.: Phytoplankton trends in the Baltic Sea, ICES Journal of Marine Science: Journal du Conseil, 60, 177-186, 10.1016/s1054-3139(02)00280-1, 2003.

Ocean color chlorophyll (OC) v6: http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/ocv6/, 2010.

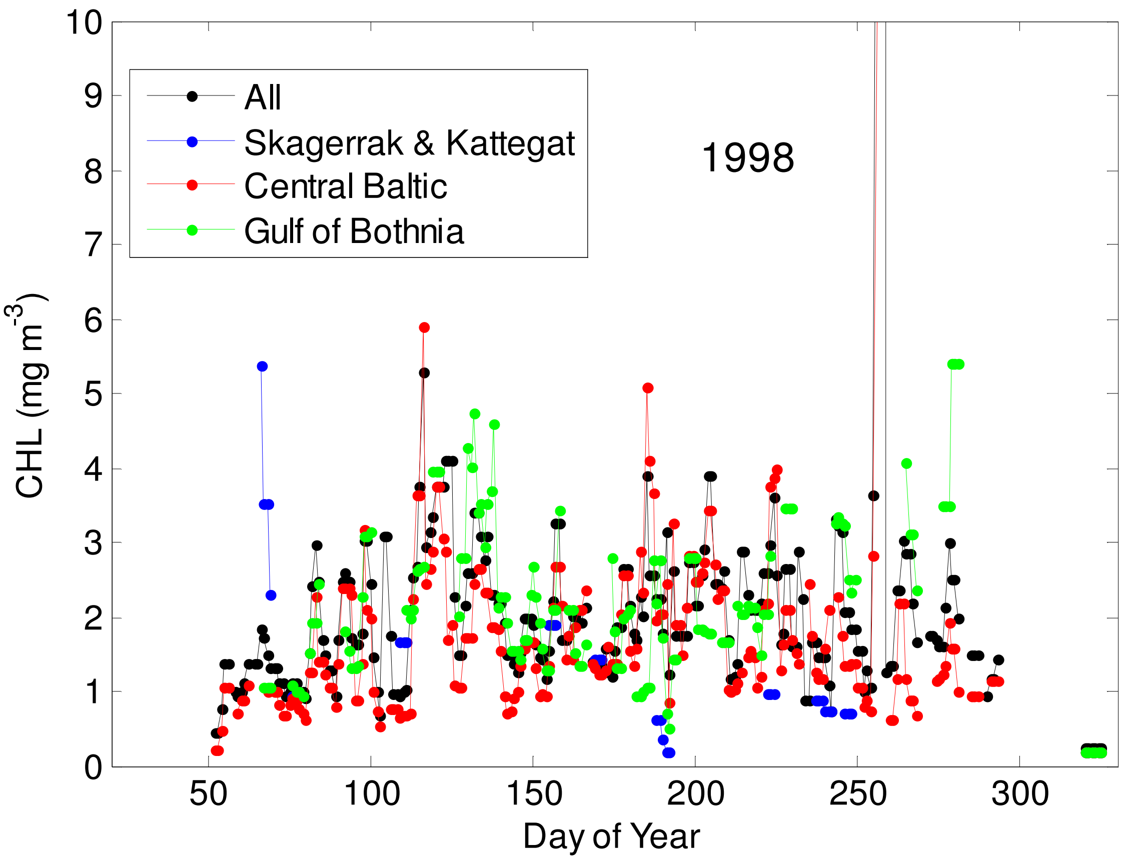

Fig. S1 CHL basin averages in 1998.

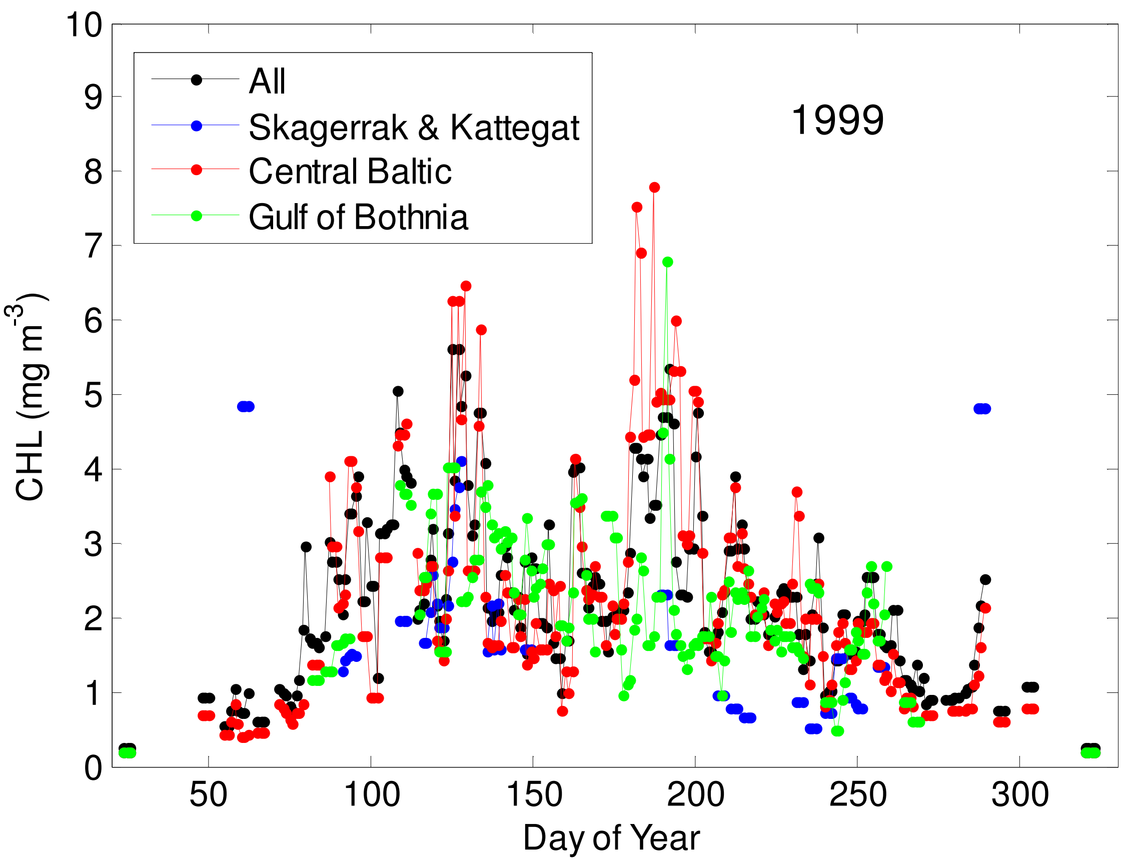

Fig. S2 CHL basin averages in 1999.

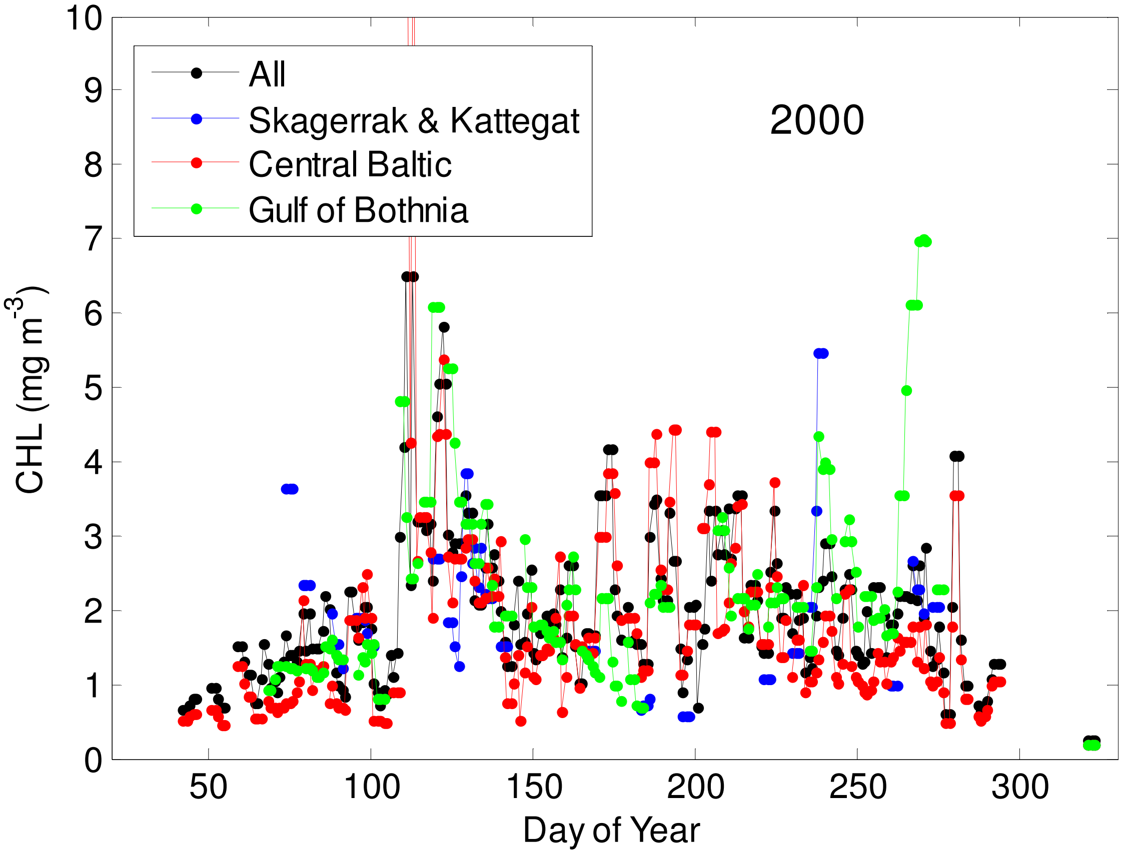

Fig. S3 CHL basin averages in 2000.

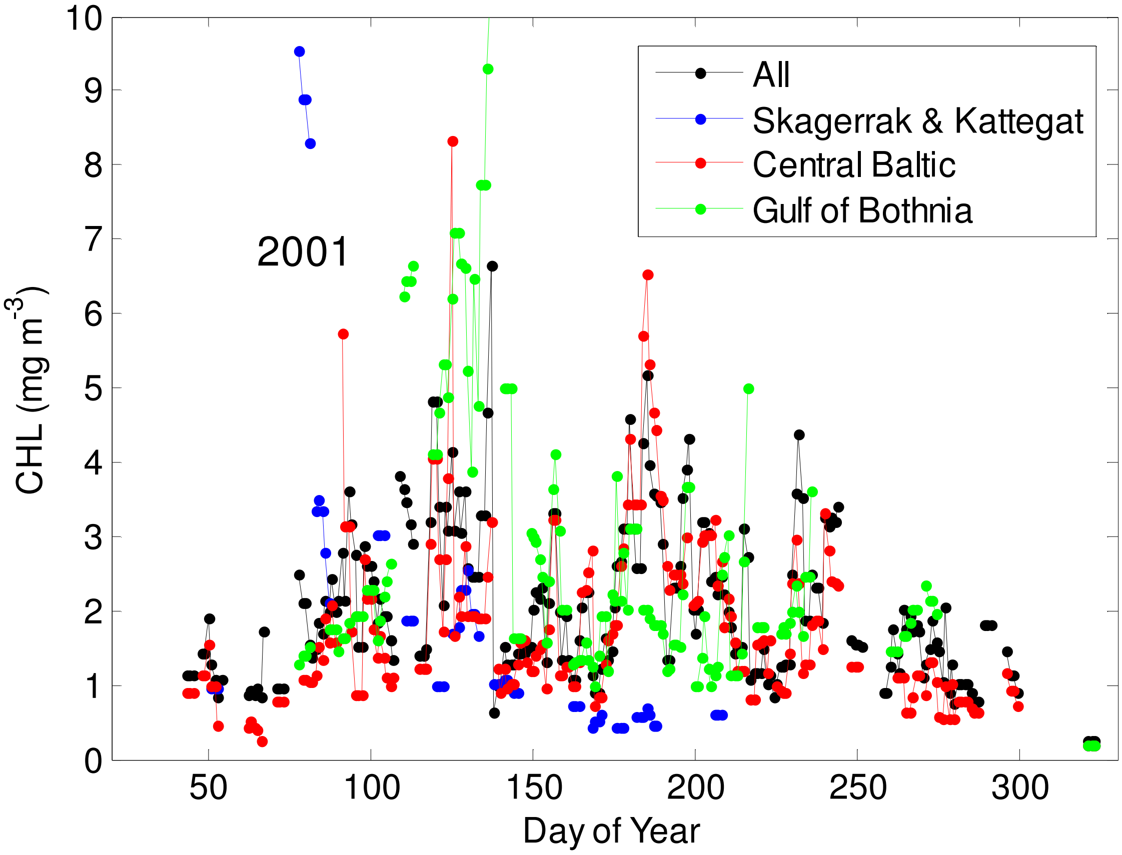

Fig. S4 CHL basin averages in 2001.

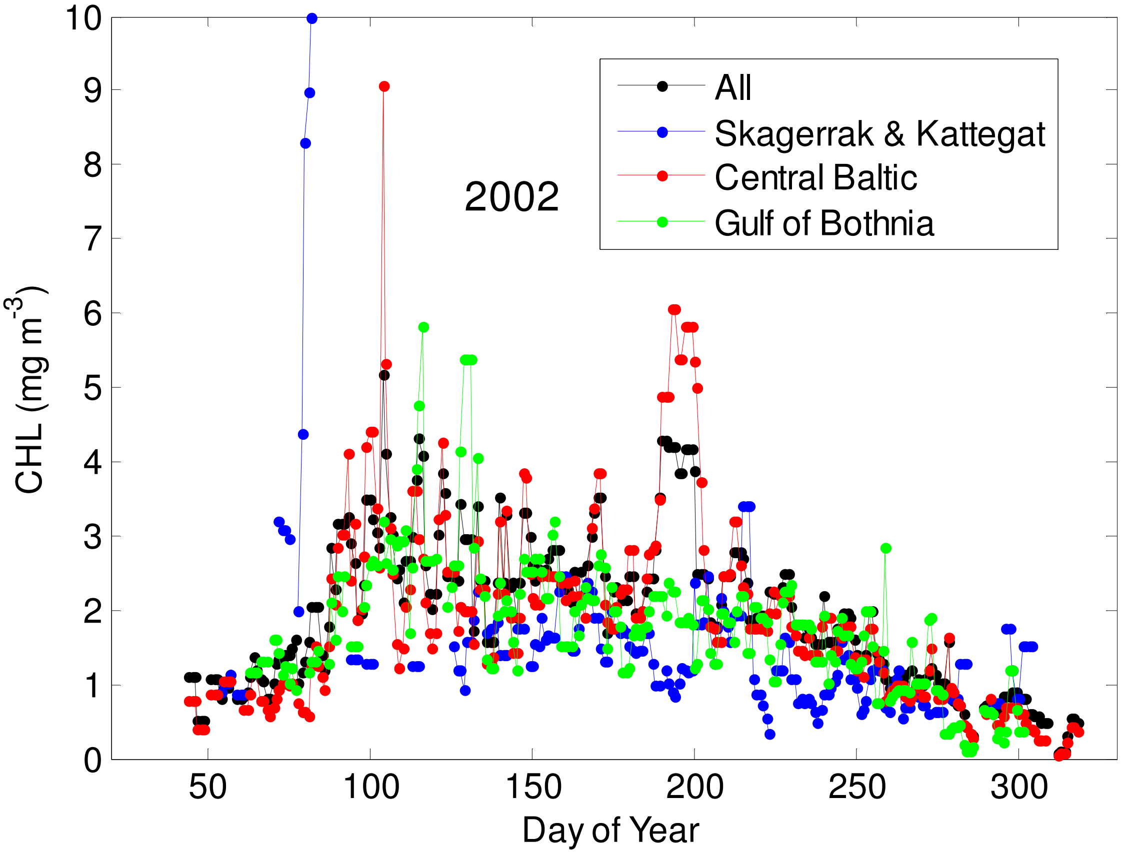

Fig. S5 CHL basin averages in 2002.

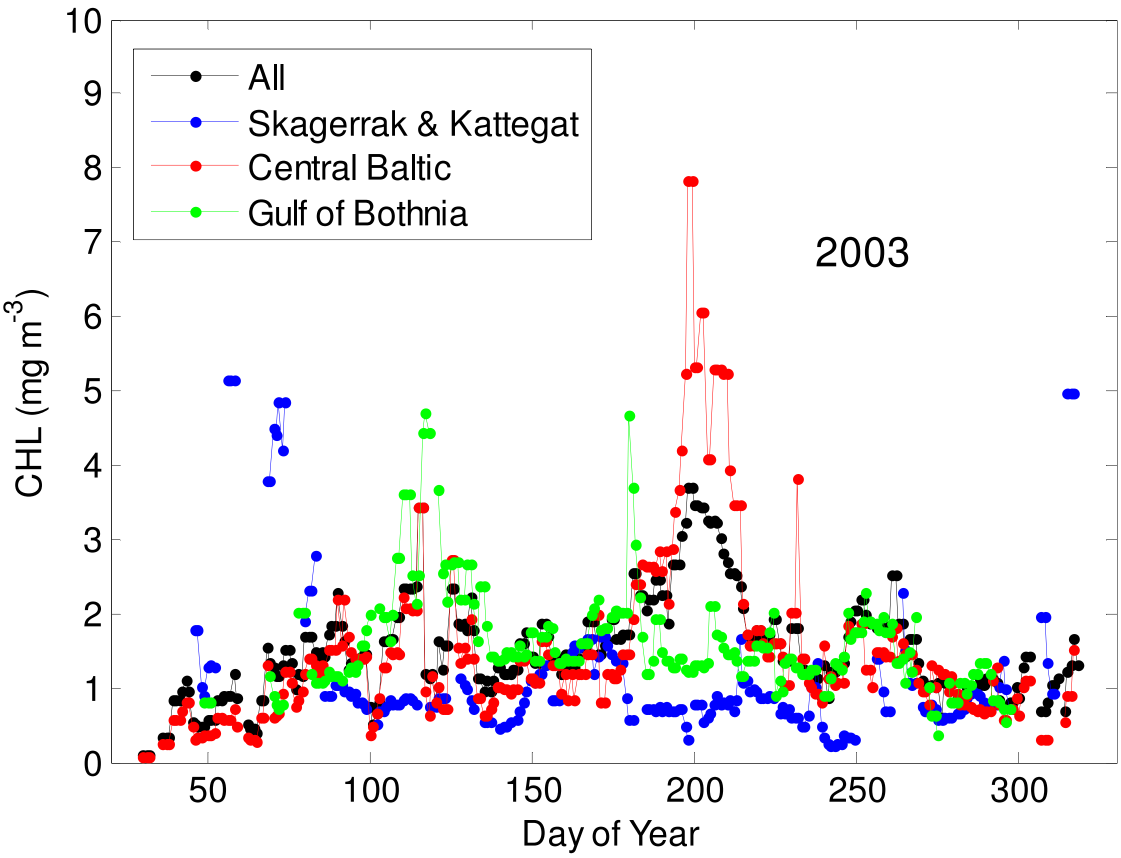

Fig. S6 CHL basin averages in 2003.

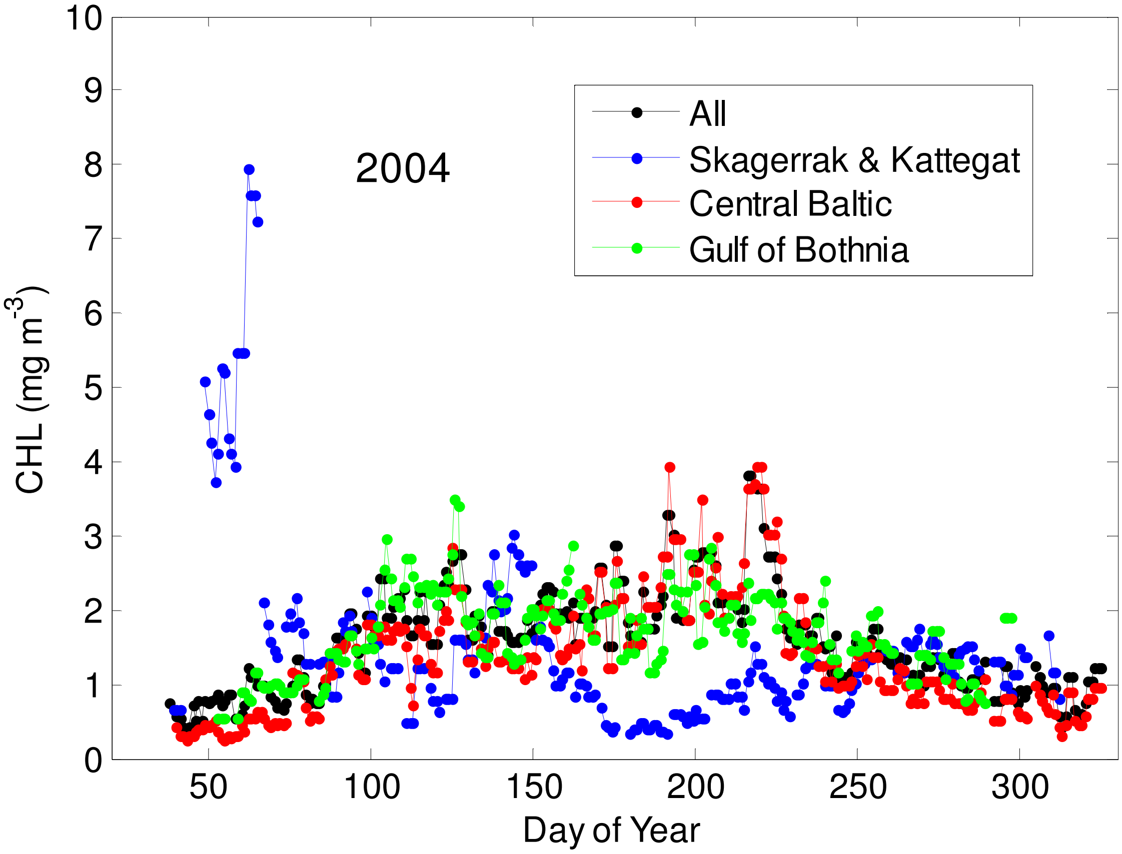

Fig. S7 CHL basin averages in 2004.

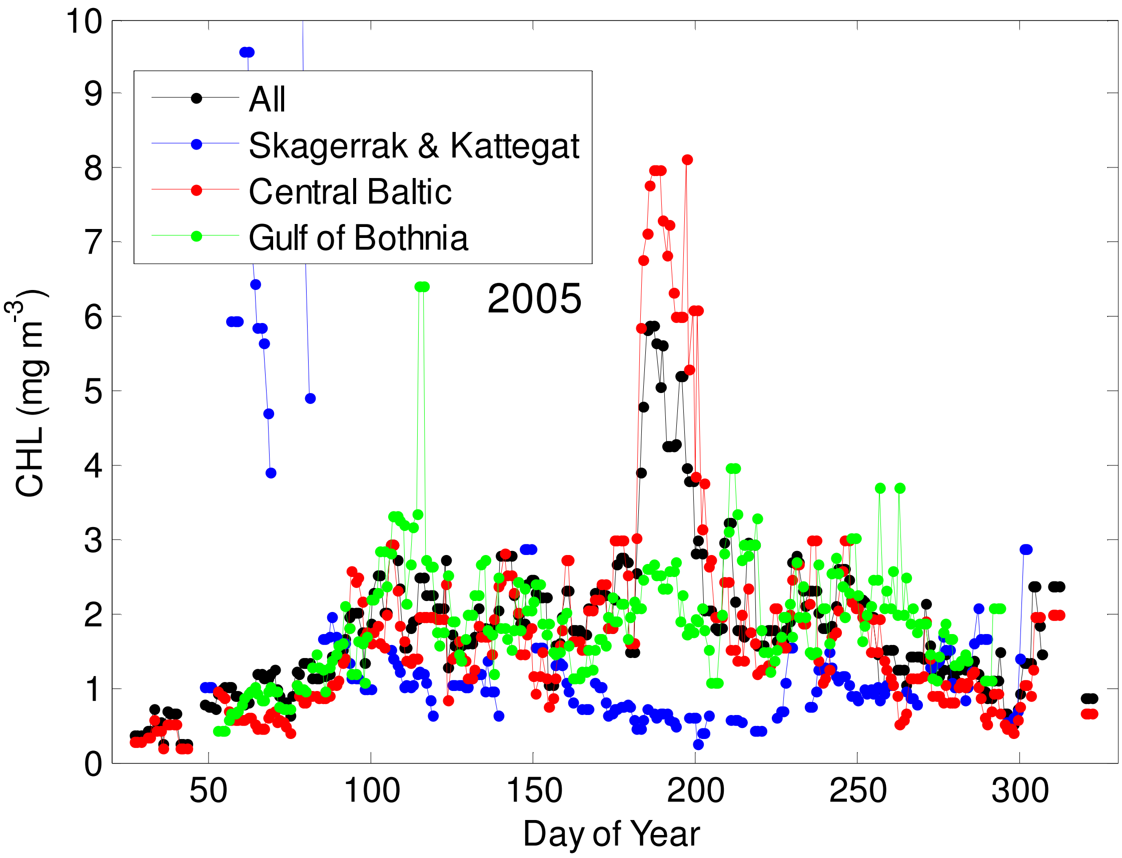

Fig. S8 CHL basin averages in 2005.

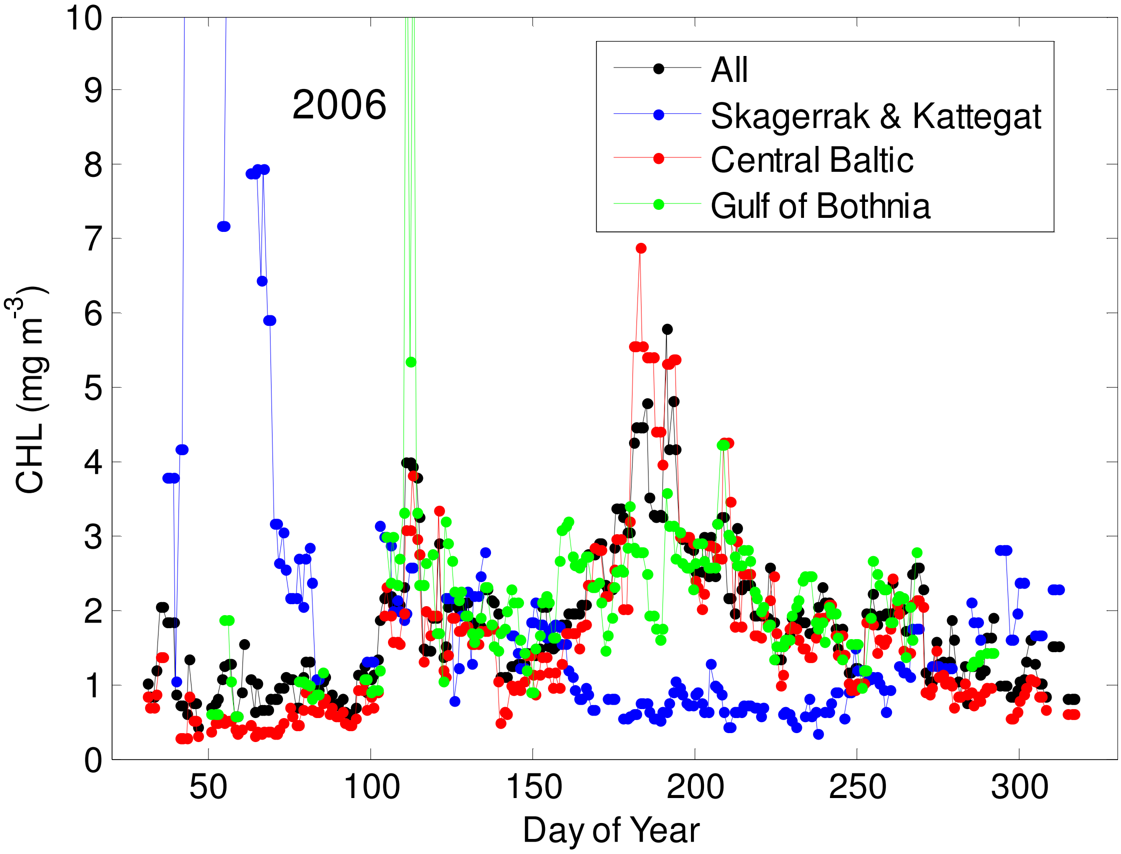

Fig. S9 CHL basin averages in 2006.

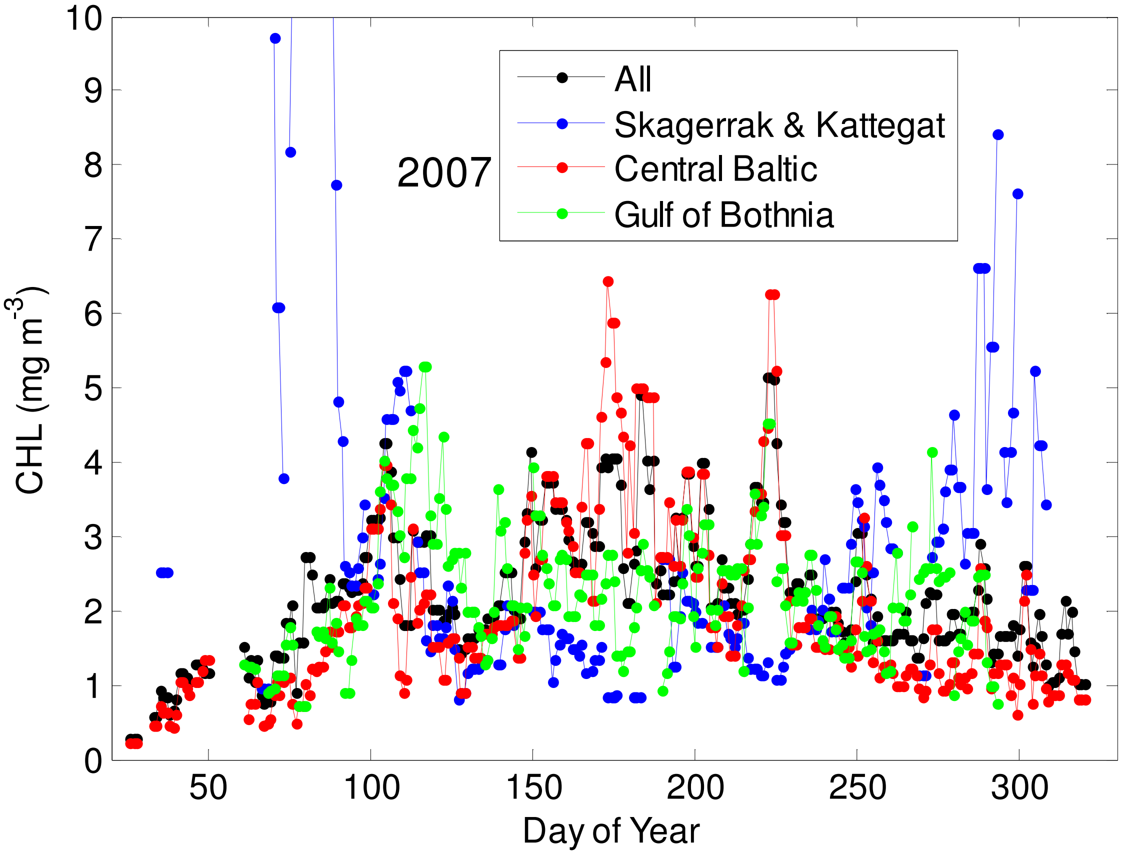

Fig. S10 CHL basin averages in 2007.

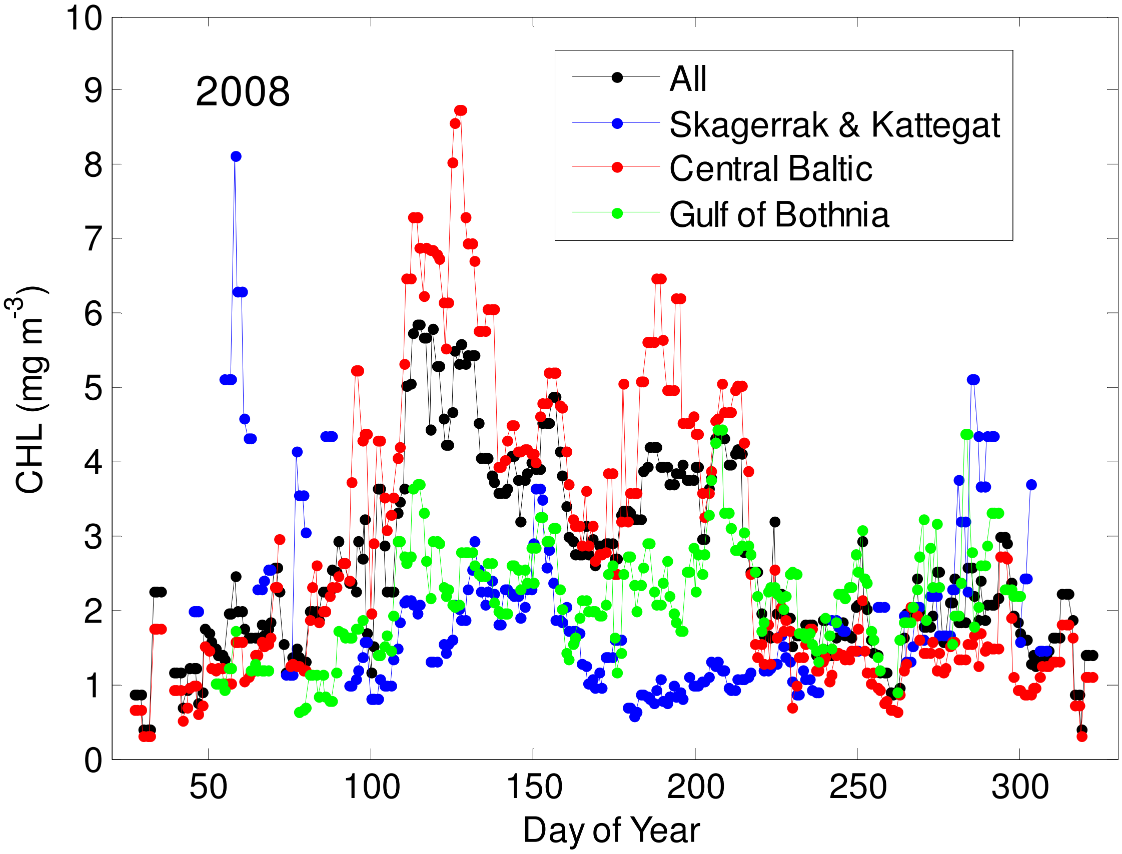

Fig. S11 CHL basin averages in 2008.

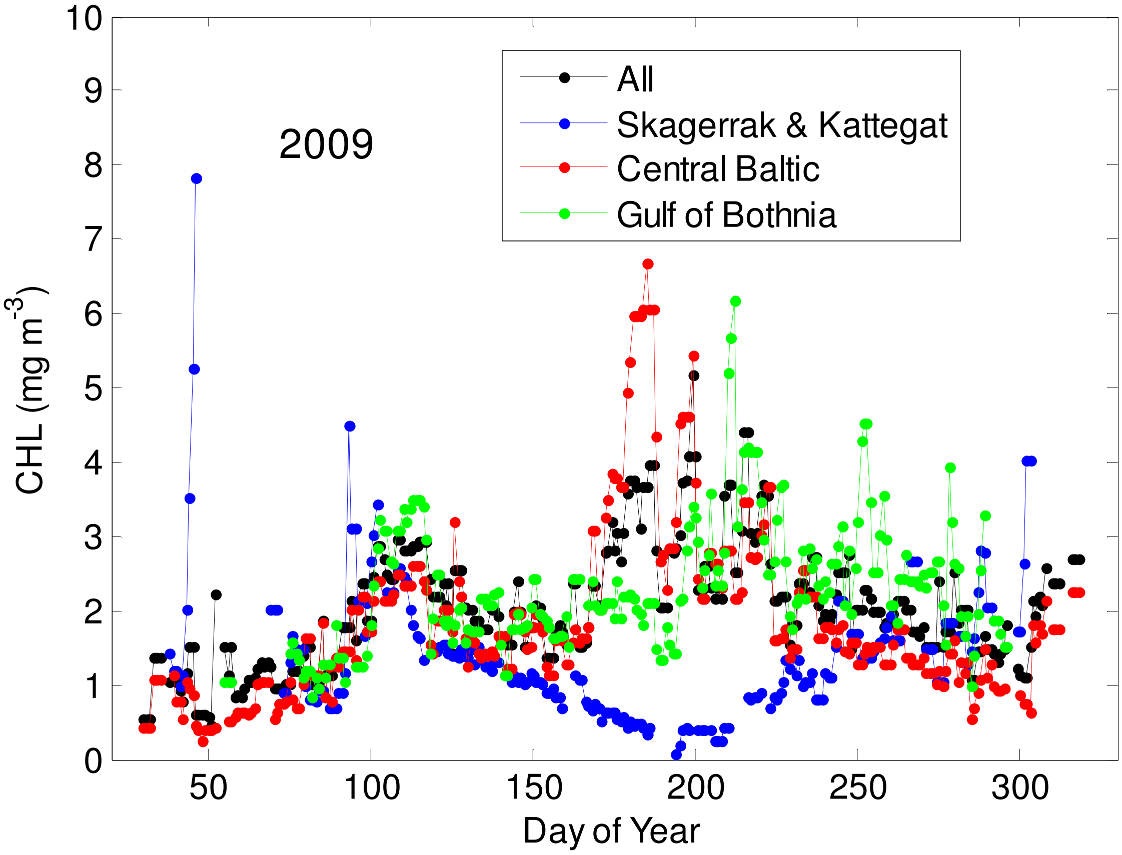

Fig. S12 CHL basin averages in 2009.

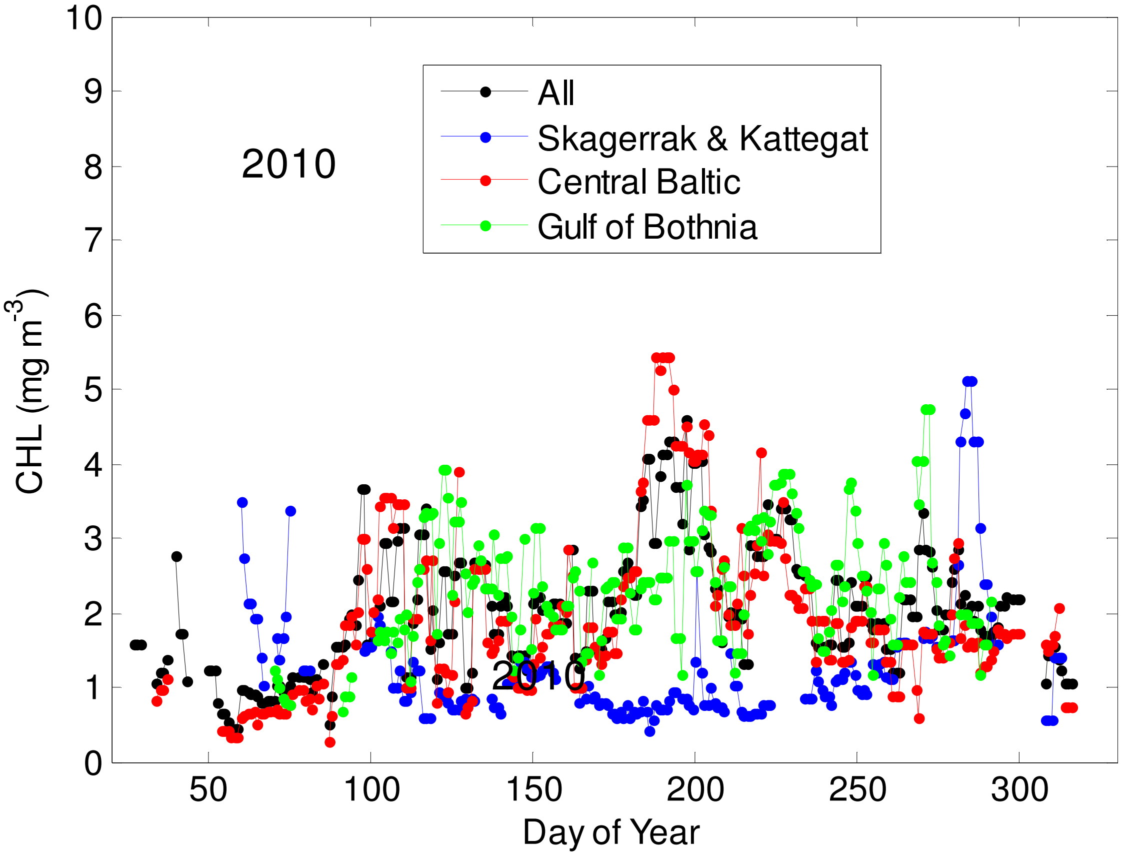

Fig. S13 CHL basin averages in 2010.

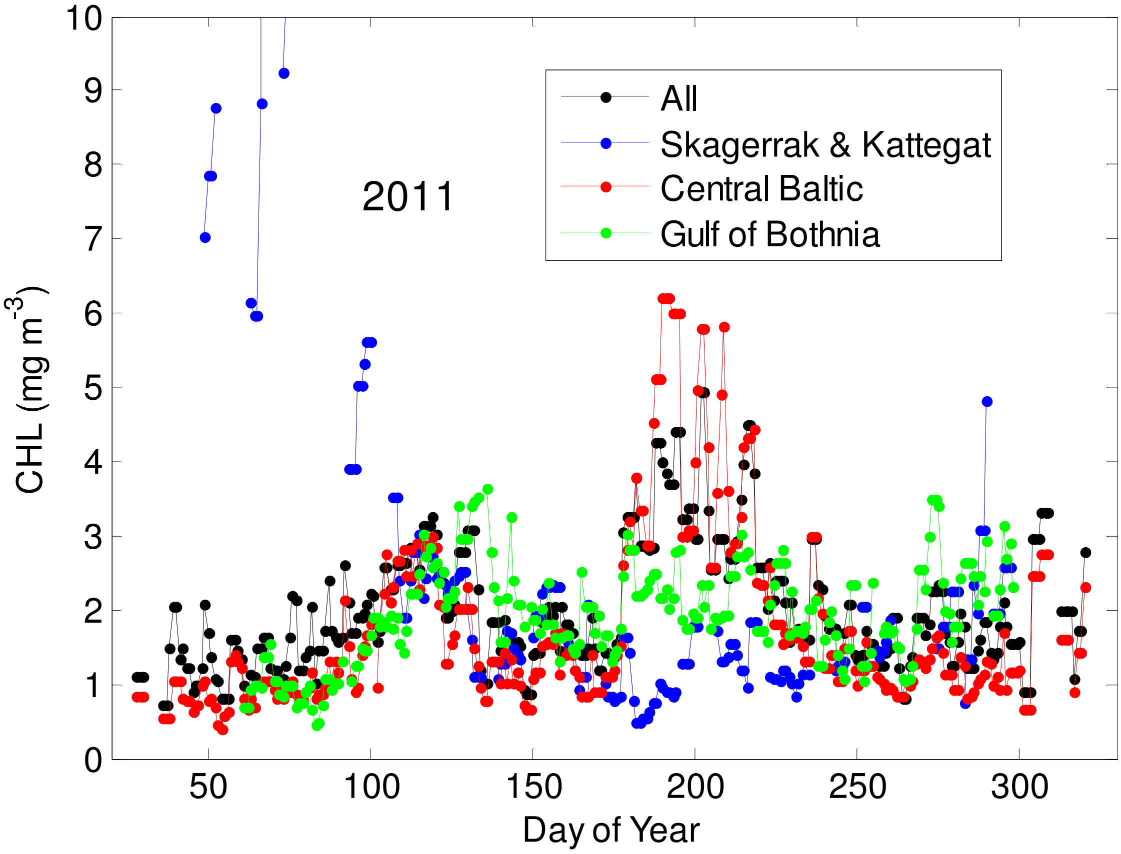

Fig. S14 CHL basin averages in 2011.

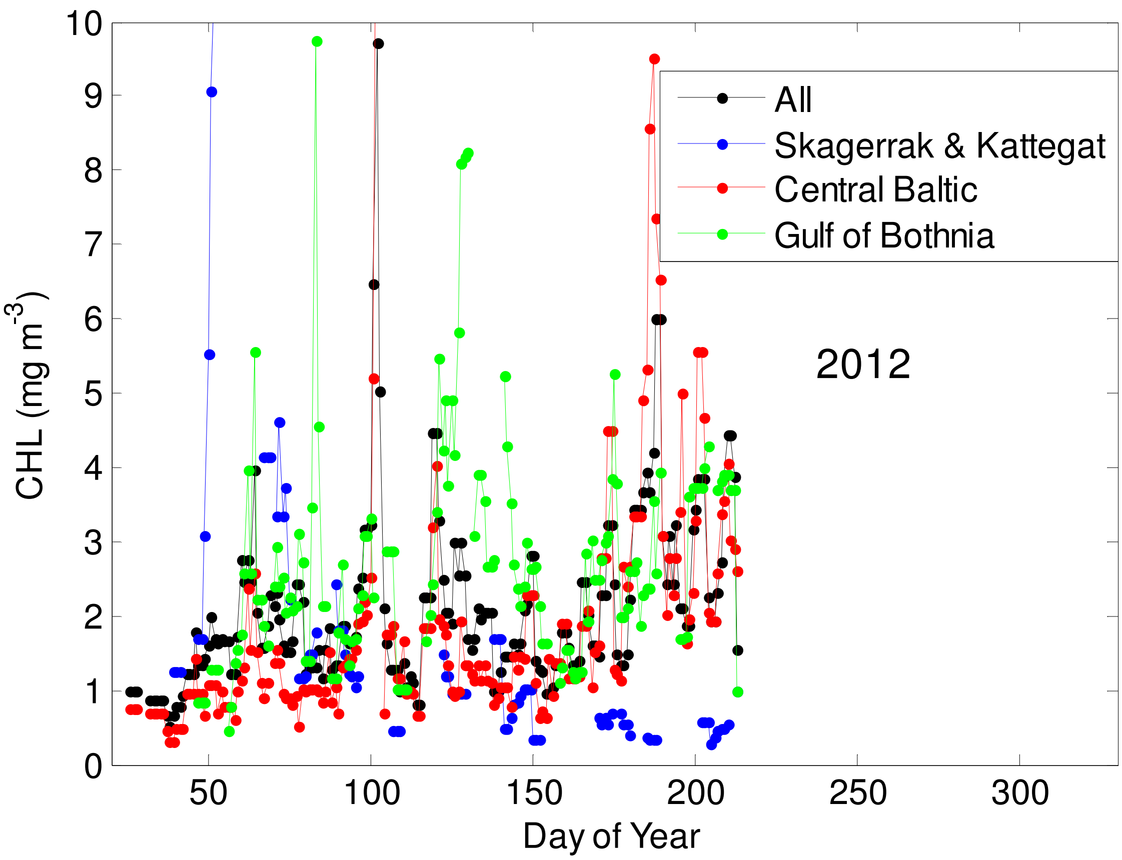

Fig. S15 CHL basin averages in 2012.

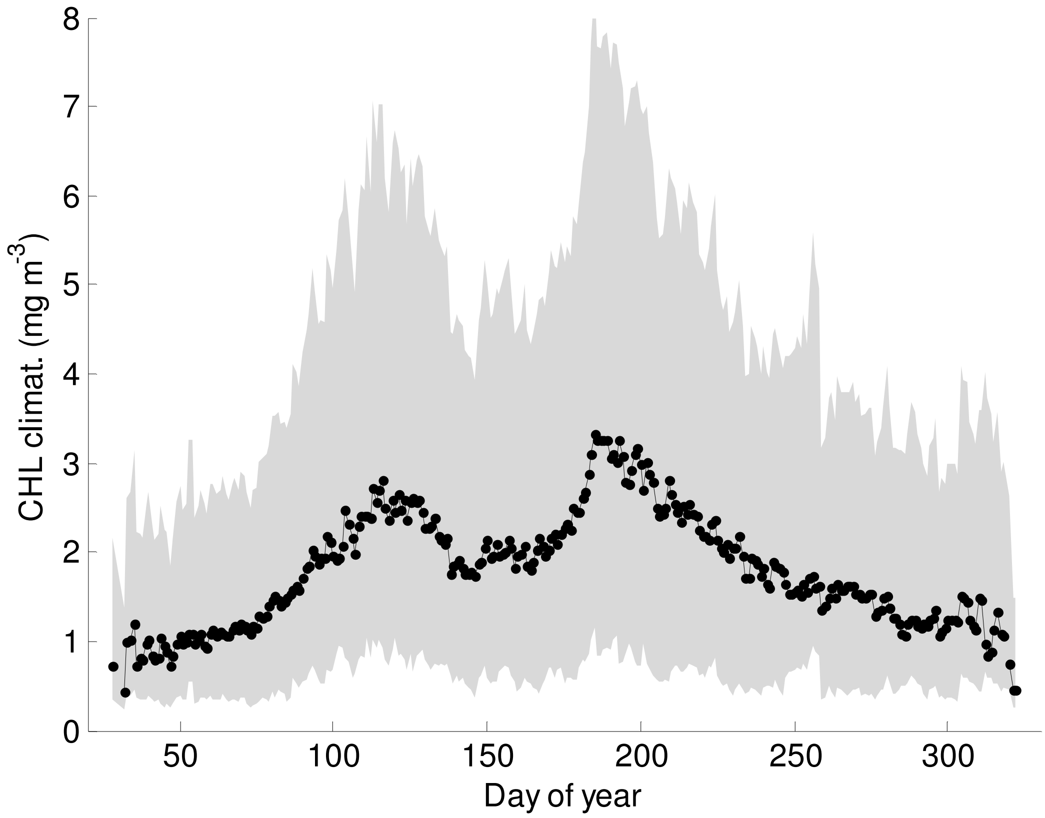

Fig. S16 Climatologic CHL of the whole domain, inside the plus-minus standard deviation band.

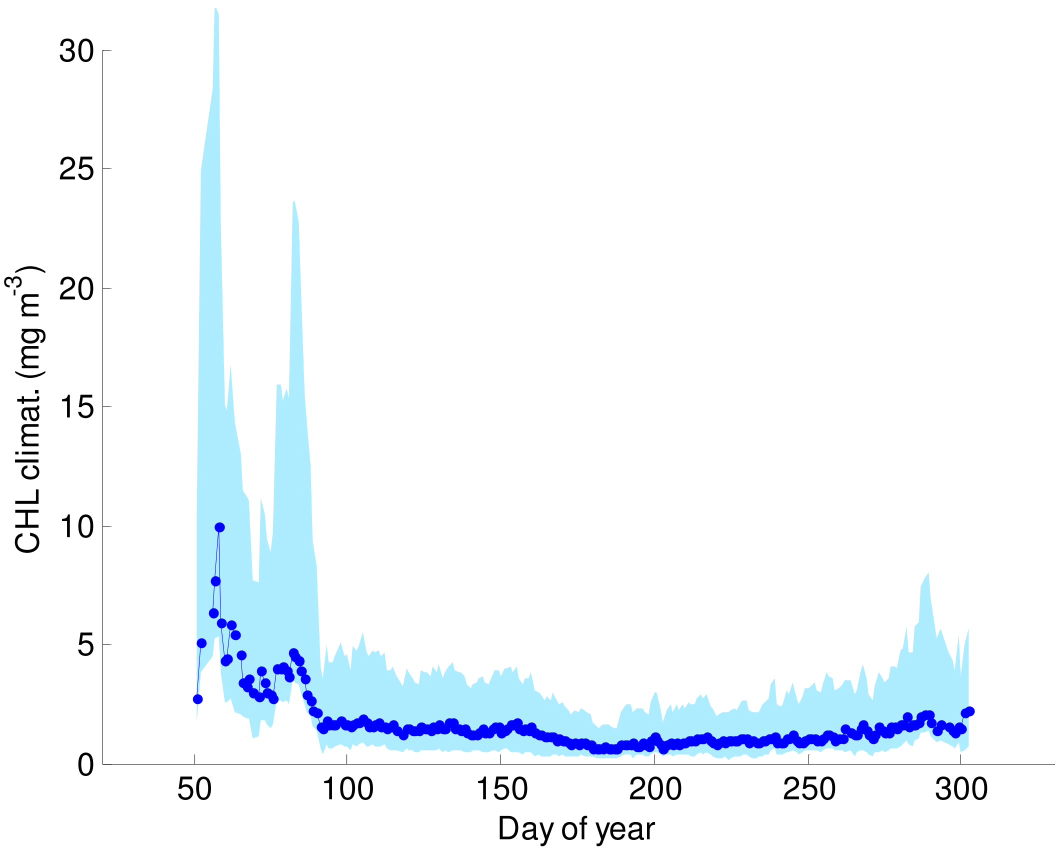

Fig. S17 Climatologic CHL of Skagerrak and Kattegat, inside the plus-minus standard deviation band.

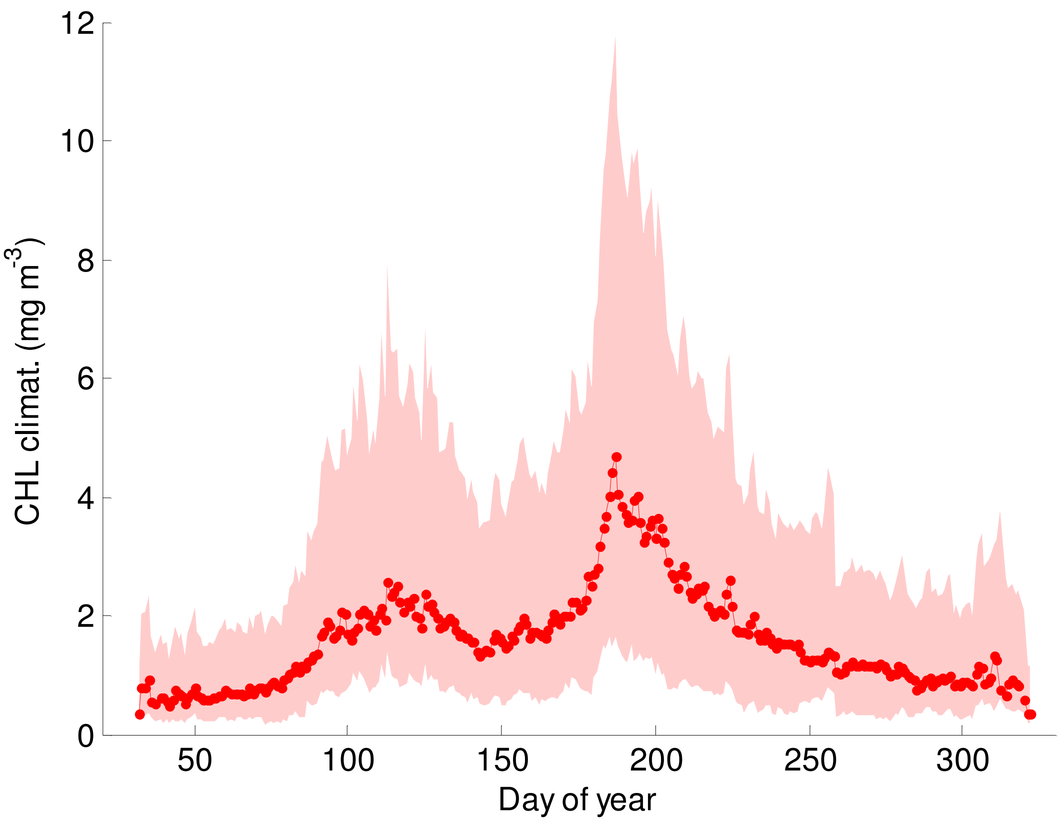

Fig. S18 Climatologic CHL of the Central Baltic, inside the plus-minus standard deviation band.

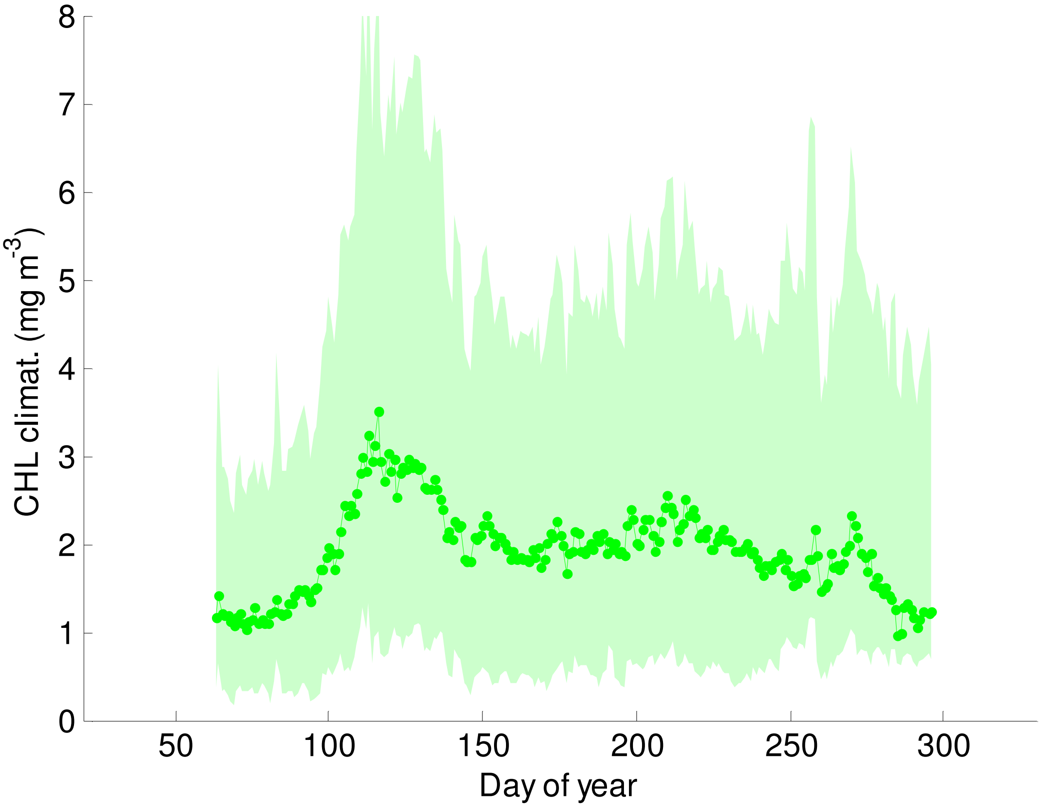

Fig. S19 Climatologic CHL of the Gulf of Bothnia, inside the plus-minus standard deviation band.

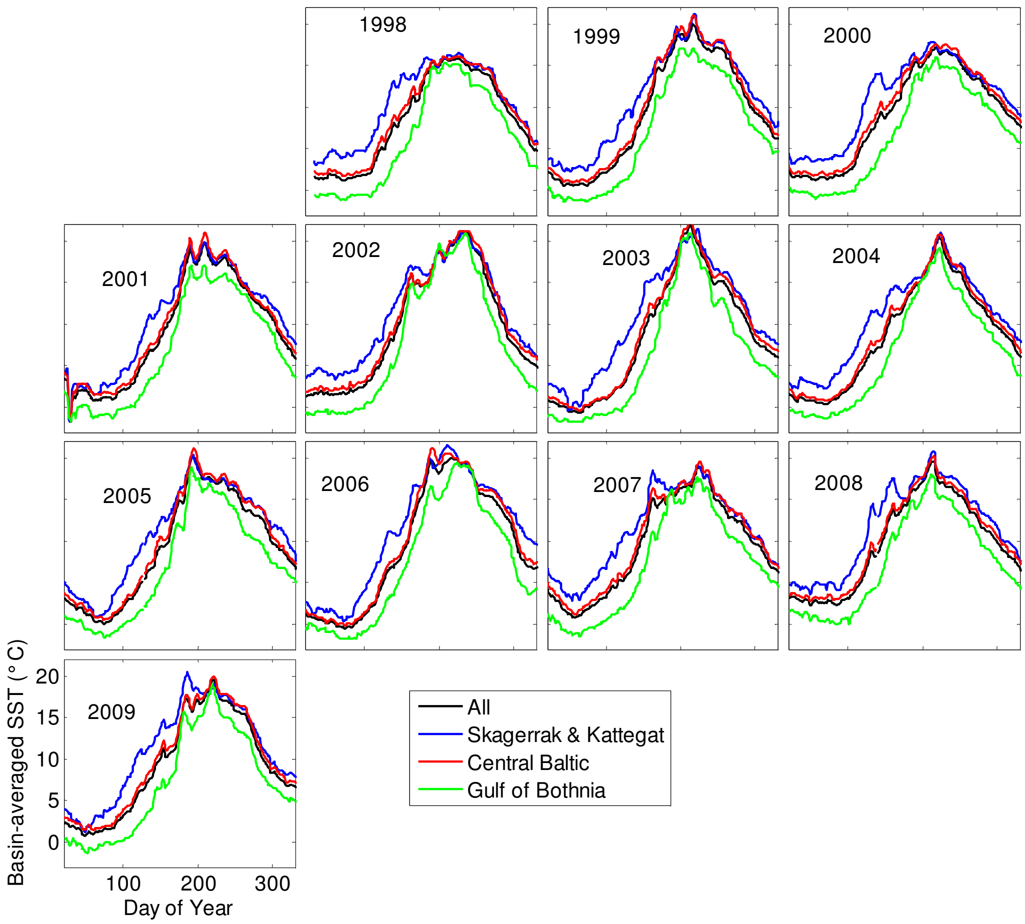

Fig. S19 Basin averages of the satellite-derived sea surface temperatures.

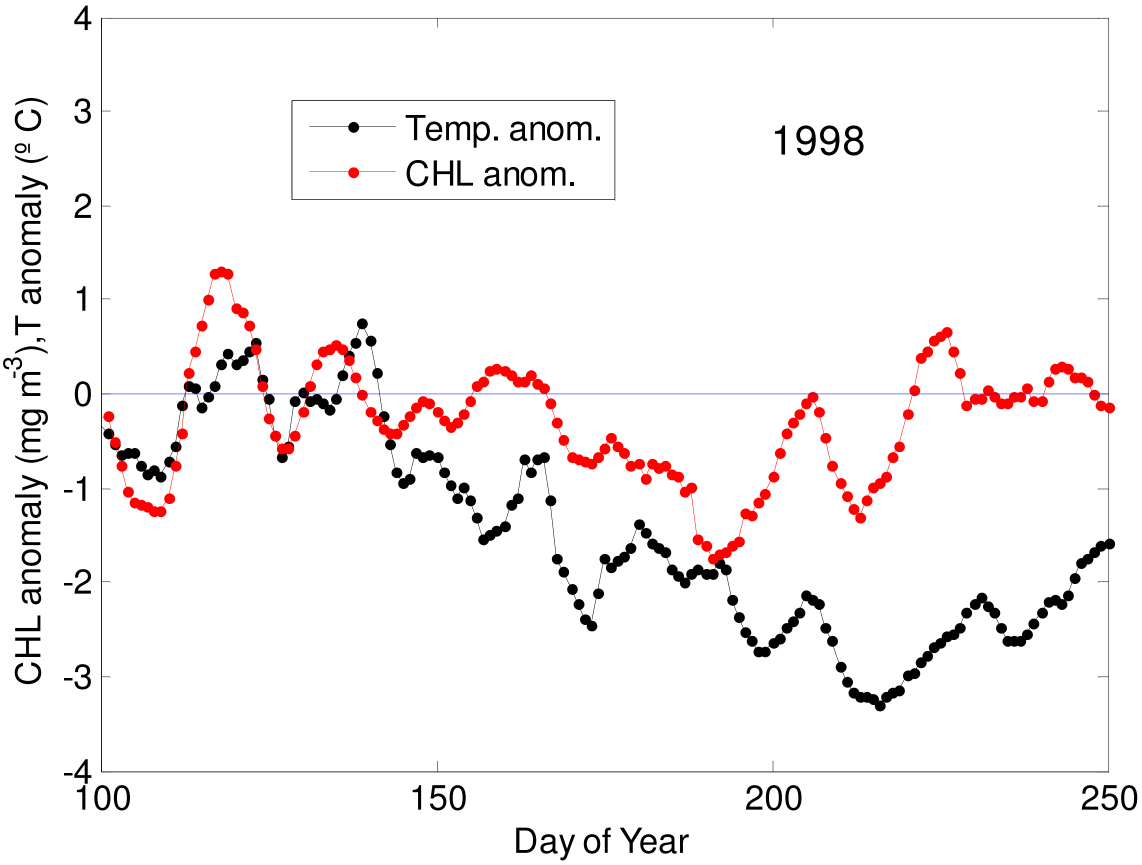

Fig. S20 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 1998.

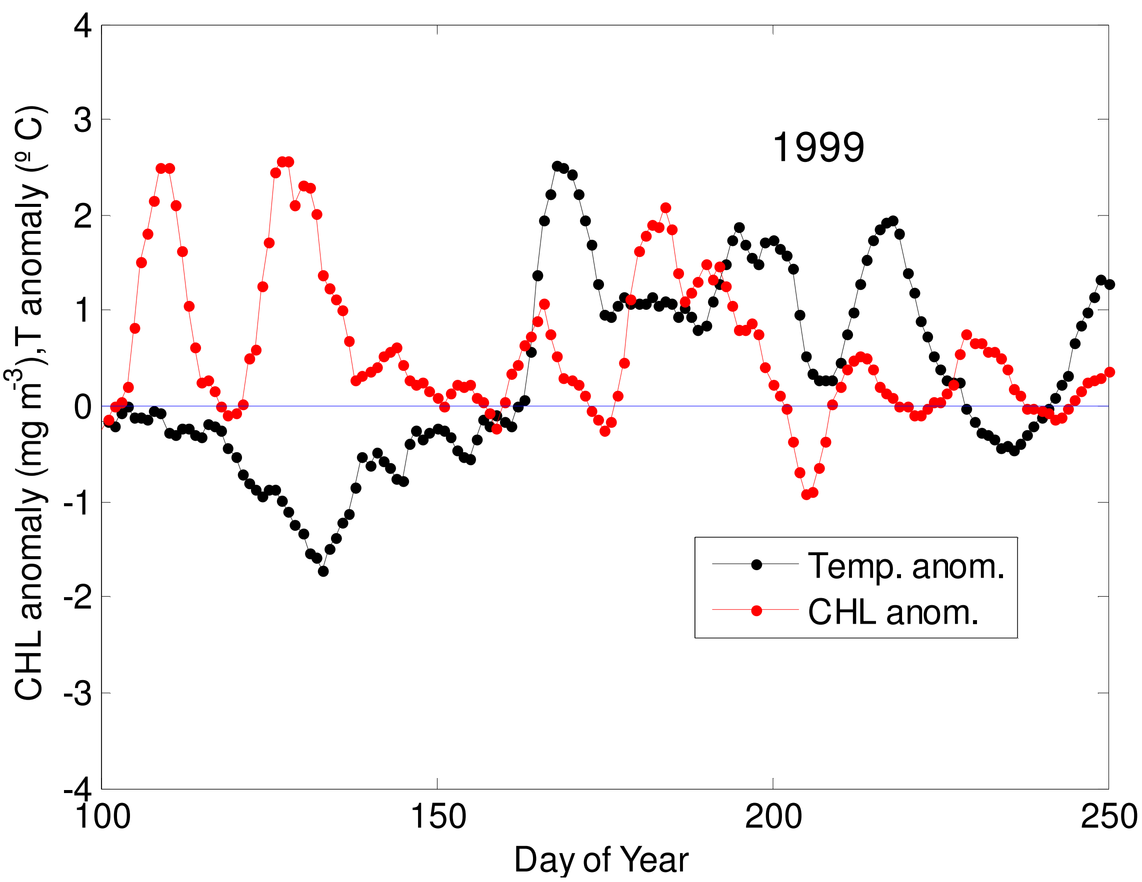

Fig. S21 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 1999.

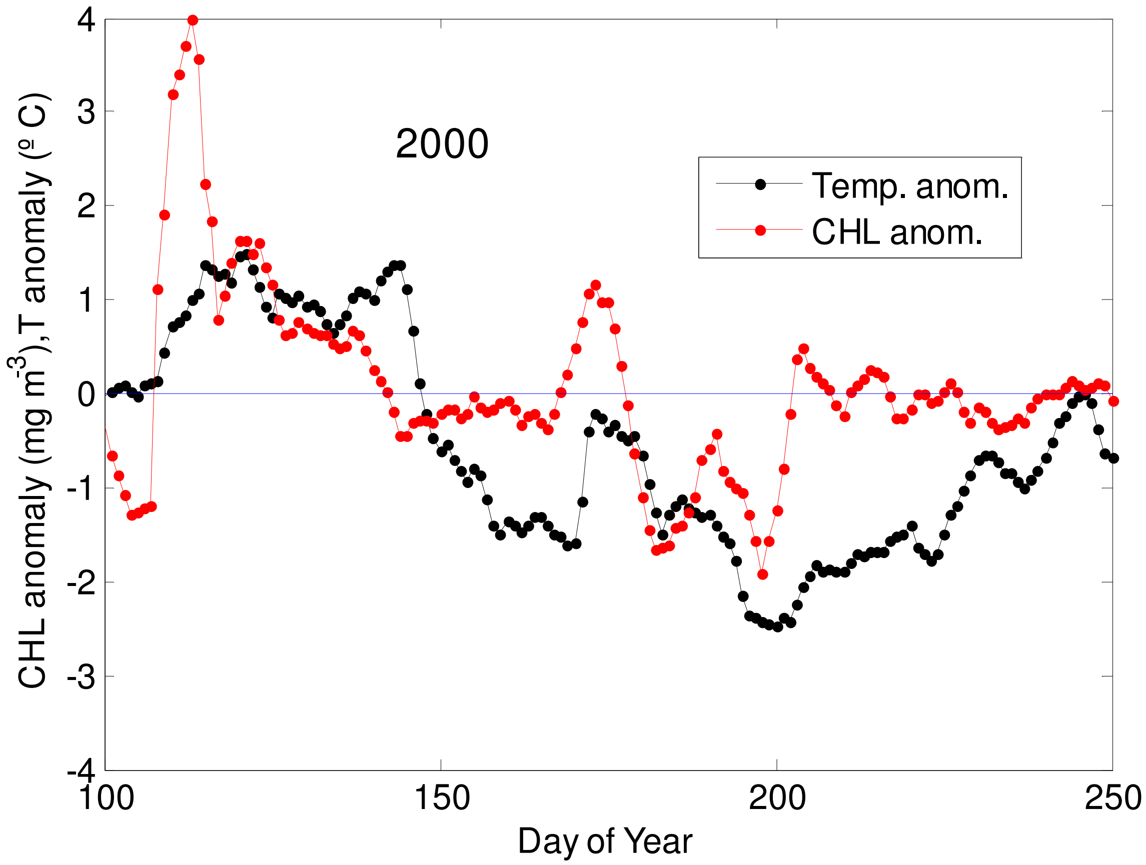

Fig. S22 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2000.

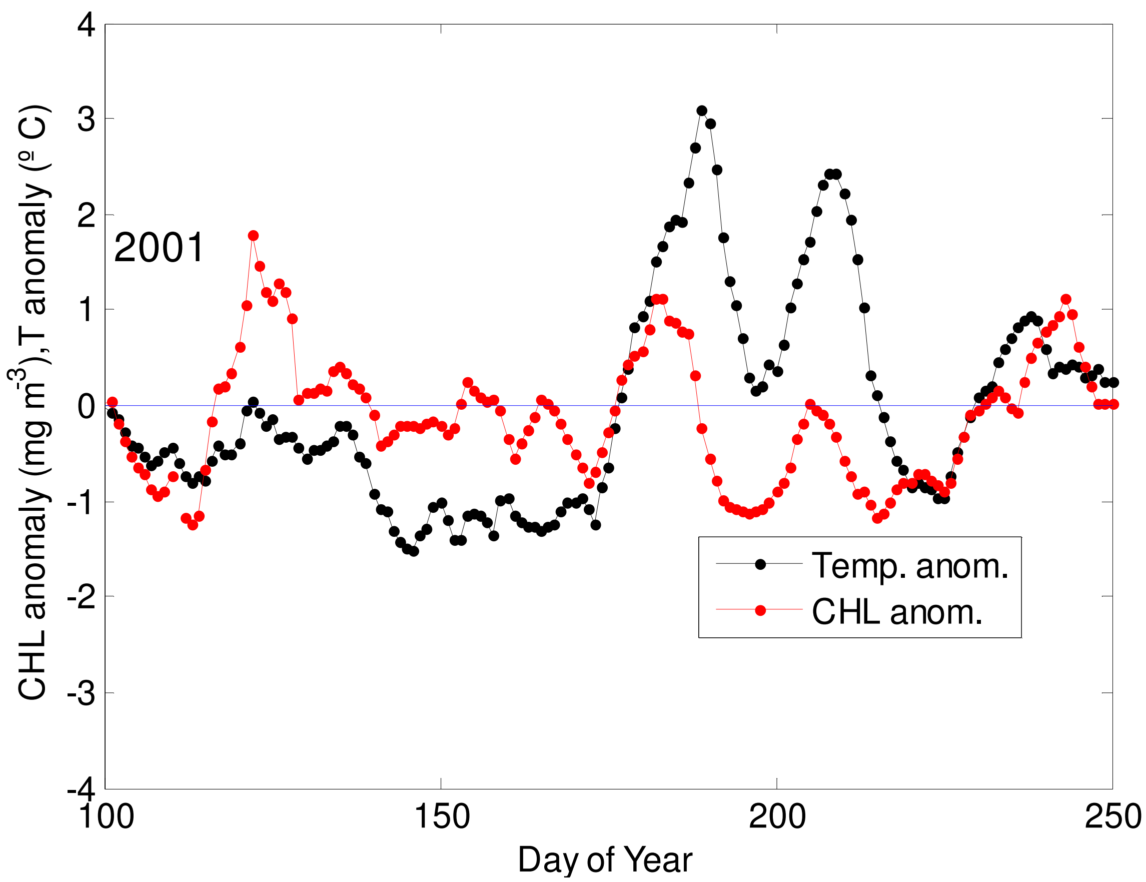

Fig. S23 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2001.

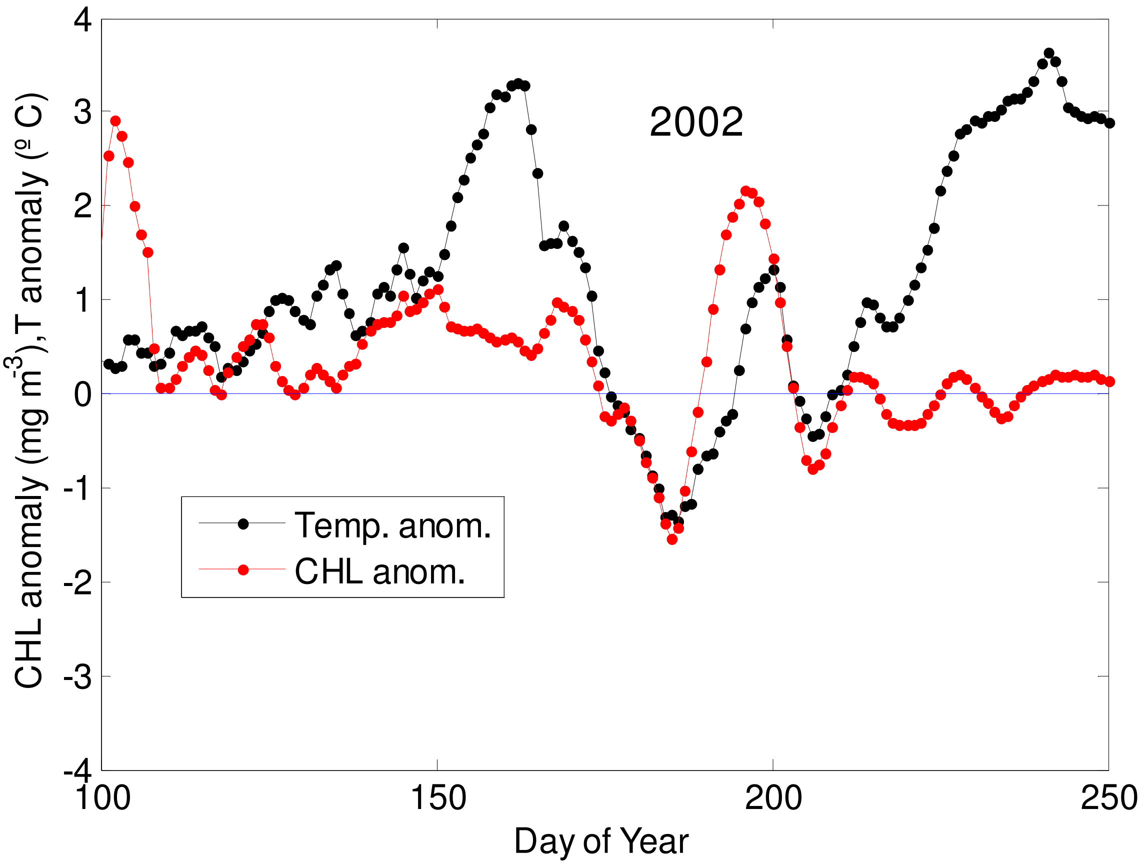

Fig. S24 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2002.

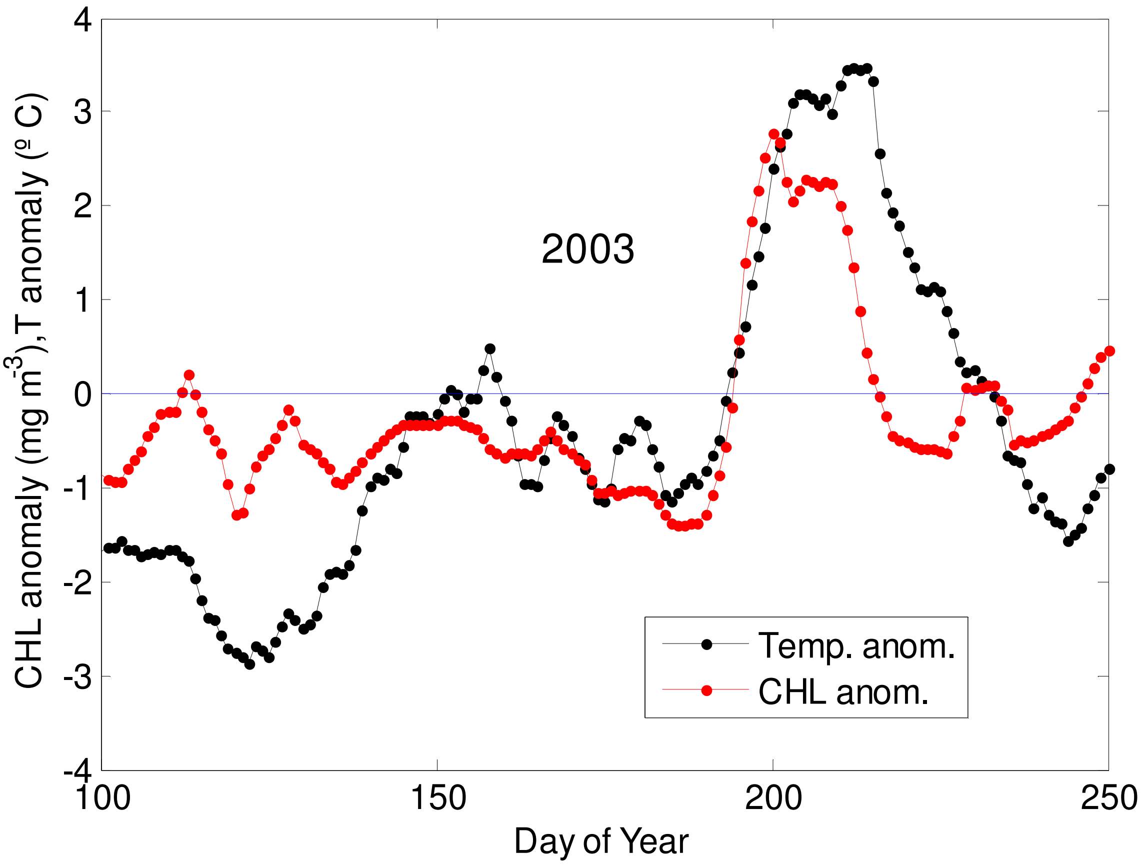

Fig. S25 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2003.

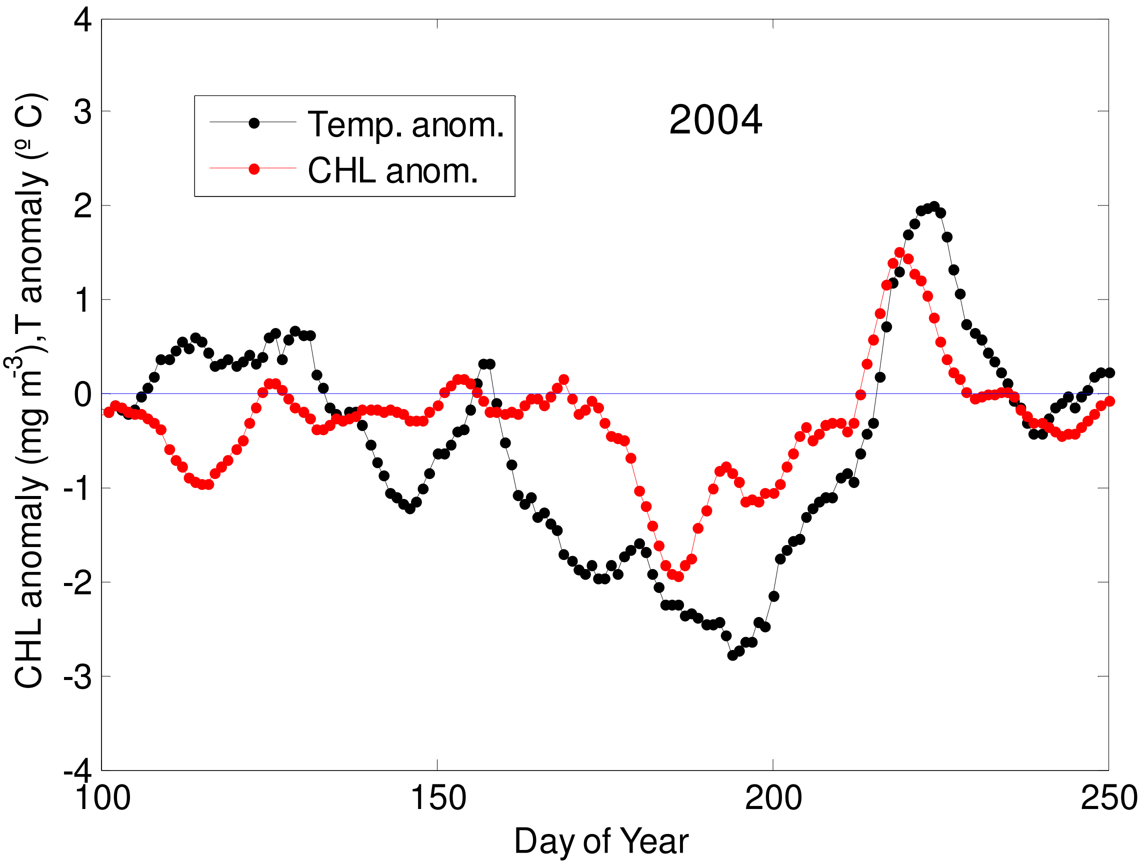

Fig. S26 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2004.

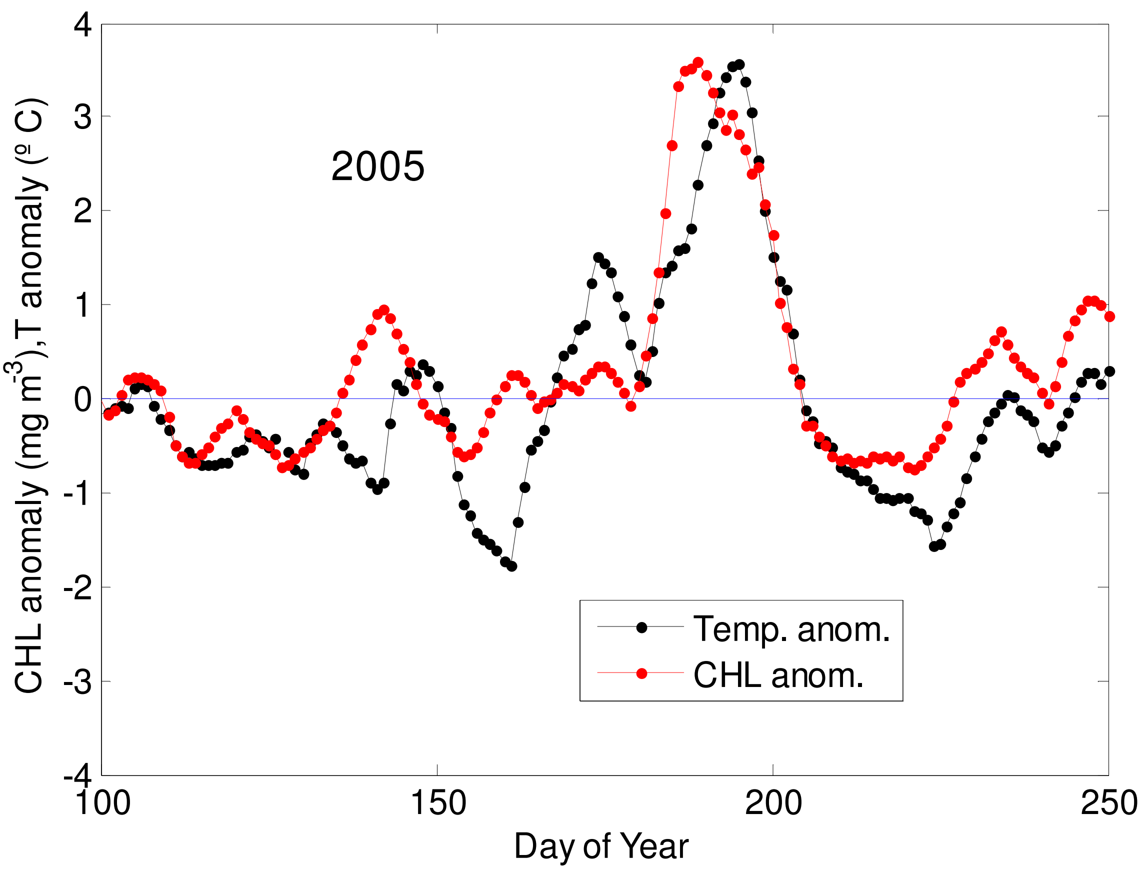

Fig. S27 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2005.

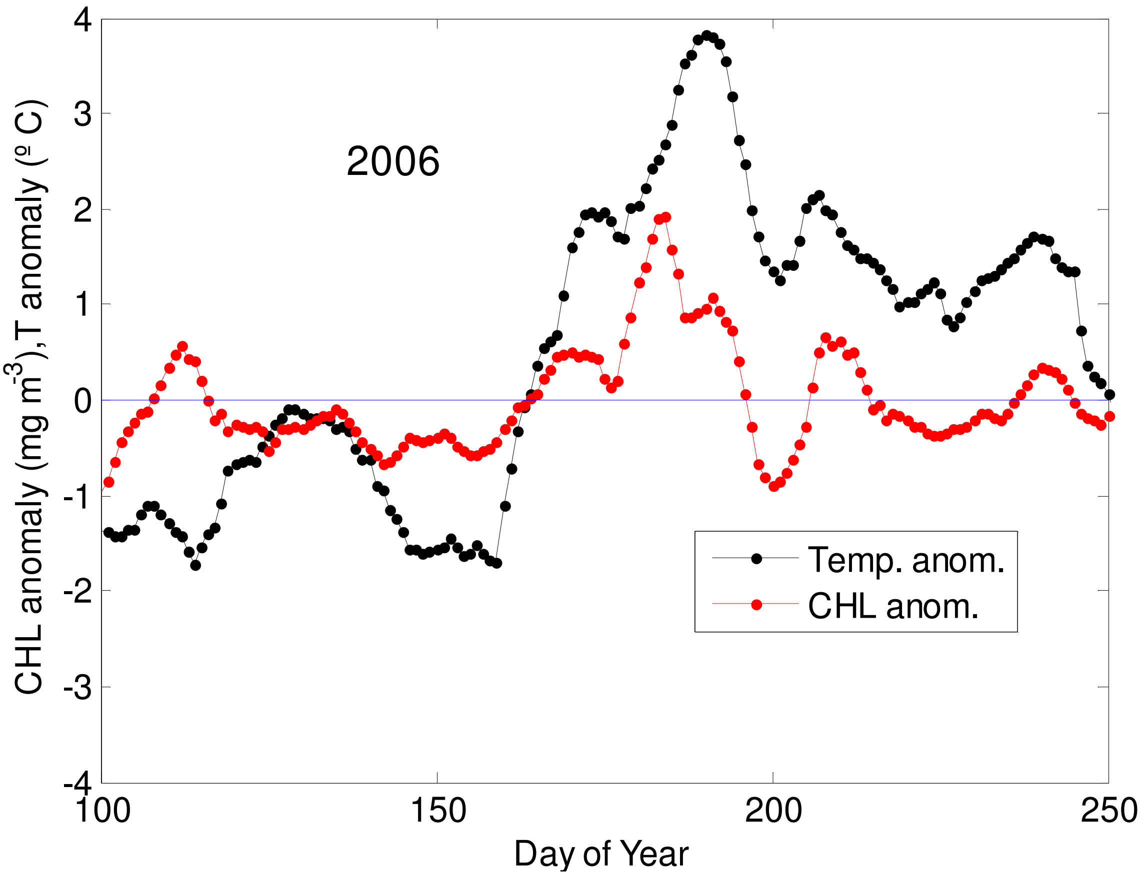

Fig. S28 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2006.

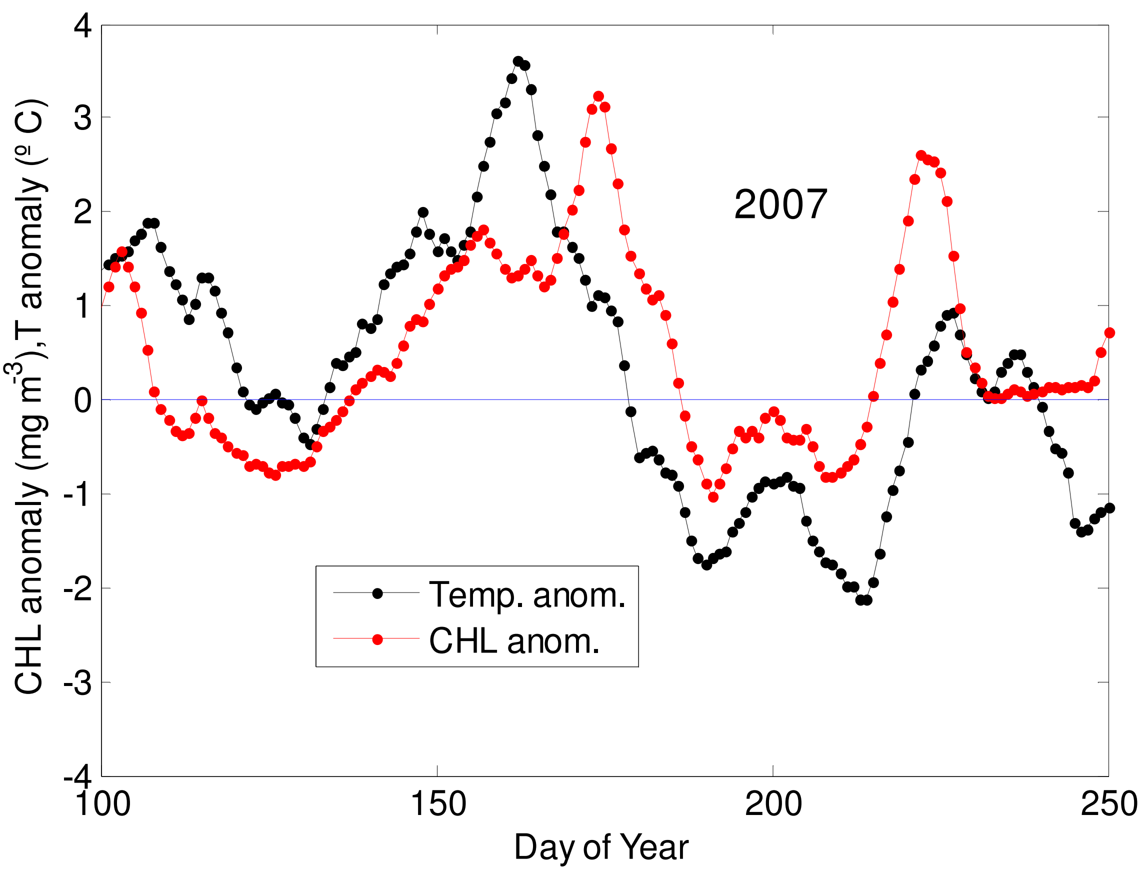

Fig. S29 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2007.

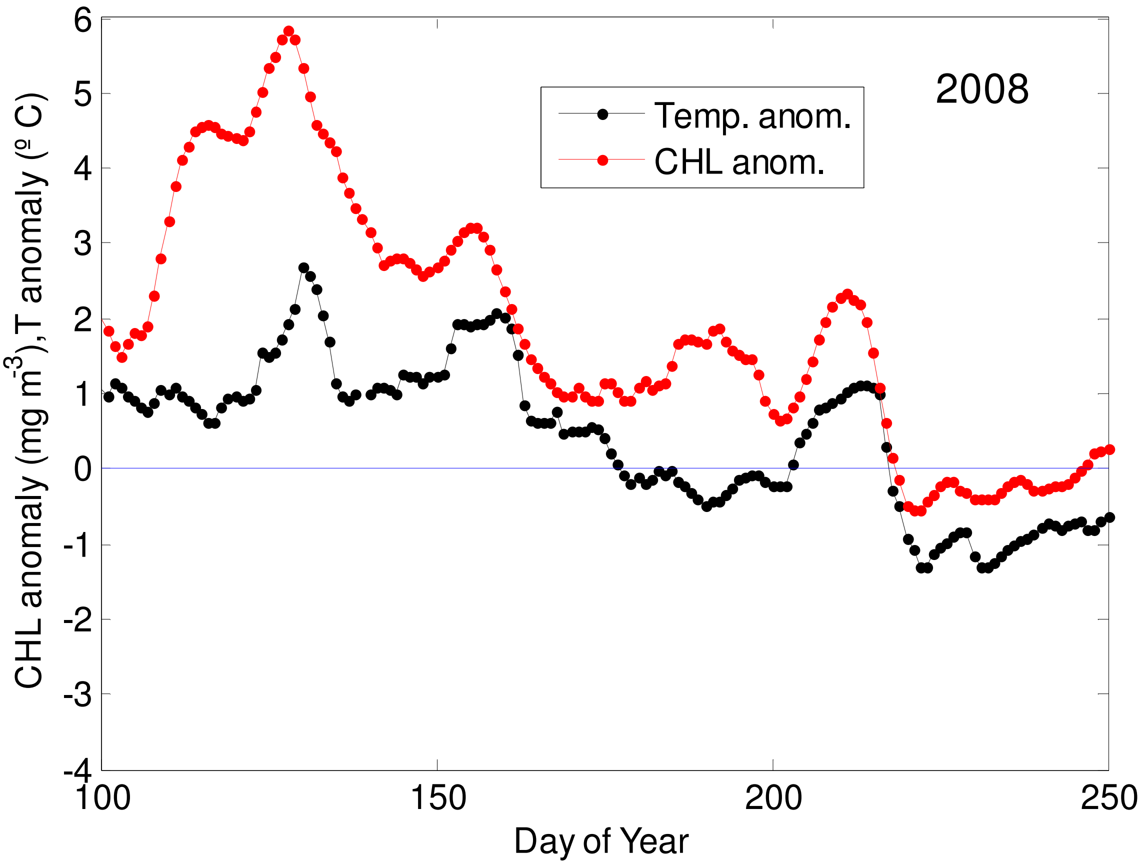

Fig. S30 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2008.

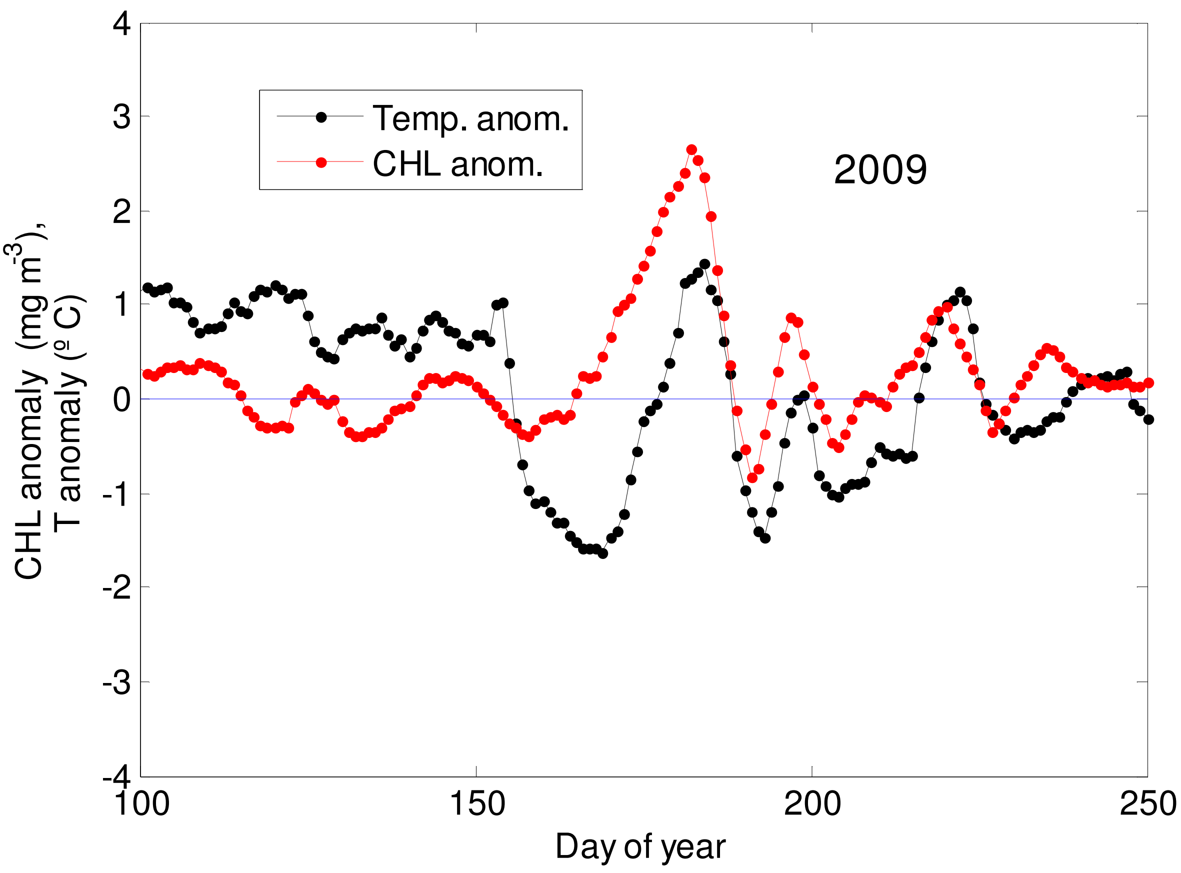

Fig. S31 Anomalies of CHL and SST over climatologies at the Central Baltic. Year 2009.