

## ***Interactive comment on “Remote sensing of chlorophyll in the Baltic Sea at basin scale from 1997 to 2012 using merged multisensor data” by J. Pitarch et al.***

### **Anonymous Referee #1**

Received and published: 17 November 2015

#### **General Comments**

This article provides a interesting insight into the regional dynamics of phytoplankton Chlorophyll at a regional scale and makes good use of a multi-sensor time series to provide new information on phenology. The justification of the dataset and algorithm used appears valid and the data processing is clearly explained. The linking of the chlorophyll dynamics to sea-surface temperature would benefit from a cross correlation analysis between the two time series to allow a numerical support to the statements made. A couple of minor corrections are required to the text (mostly for ease of reading) and a couple of the plots (such as missing legends or the addition of regression lines).

C1150

#### **Specific Comments**

page 2289: line 15: The authors state that there are “many more” invalid data in the MLP product. It might be of use to give a mean % valid pixels difference for a sample of representative images so that readers can quantitatively consider the difference.

pg2290 lines 5-20: There is no information provided on the methodology of the in situ chlorophyll a measurements. Were the data used from the database measured using HPLC, extracted fluorometric methods, both?

pg2291 line 20: The authors state that they discard outliers, where outliers are defined as possessing an error greater than a set criteria (ratio). I understand the need to remove outliers, especially in log normally distributed data, but are the points discarded evenly distributed across the range of the dataset? The distribution of the ‘very poor’ data is of interest in itself. Also, the authors state that 3.5% of the values are discarded using the ratio threshold. One assumes that the approximately normal distribution of satellite and in situ measurements also leads to a somewhat normal distribution of errors. Would it be possible to state that points where errors were outside the range 2nd-98th percentile of all errors were removed as outliers? If not then could the authors justify the error ratio chosen to dictate outliers?

pg 2295 line 12: The authors state that the horizontal-averaged CHL for OC4v6corr were computed for images with a minimum of 1000 valid pixels. I assume that this is referring to the mean CHL across the entire basin? If so then it might be worth stating the total number of pixels that make up the region of study. This would then allow a statement along the lines of “Horizontal-averaged (whole Baltic) CHL for OC4v6corr were computed only for images with at least 1000 valid pixels (x% of basin observed).” Obviously observing the whole basin in a single day is near extremely unlikely but it would be good to know if 1000 pixels corresponds to 10%, 20%, 30% etc when considering the following phenological discussion.

pg 2297: lines 2-8: The authors state that there is a relationship between SST anomaly

C1151

and CHL anomaly time series. It would be worth performing a brief cross correlation analysis of the two time series such that the link could be quantified beyond “high-amplitude temperature anomalies induce similar growth and decay in CHL.”. This may also provide information of the lag of a system response to forcing factors, which would be of interest to those cited in the discussion of lines 9-29.

#### Technical Comments

pg2283 Title of article: The authors should hyphenate ‘multi-sensor’ as both terms apply to the following noun.

pg 2284 line 2: “Fifteen-year” should be “A fifteen-year” line 8: Authors should expand the sentence “Statistics showed low linearity”. Is this for all the provinces, individually and combined? Is this for all algorithms?

pg 2285 lines 2-6: A very long sentence. Perhaps best to break after Øresund? Change “Great Belt and Øresund, thus leaving the Skagerrak...” to “Great Belt and Øresund. This leaves the Skagerrak...” line 18: It would be more clear if “Thus statistics should be consequently calculated.” were changed to “Therefore the statistical assessment of algorithm performance should be performed on the area as a whole.”

pg2286 line 4: “to cite a few” seems superfluous. line 27: perhaps include chl in the units for “1 mg m-3” otherwise the value could technically refer to concentrations of some other optical component.

pg2287 line 4: “Case 2” to “Case II” for consistency with cited literature.

pg 2288 line 4: change “remains up to date” to “is currently”

pg 2292 lines 3-4: change “that it can be derived” to “that can be derived” line 9: change “BIAS respect to the...” to “BIAS with respect to the...” line 27-28: It might be more clear to state that “In each region, OC4v6 overestimates CHL by > 40% on average.” as “In all cases, OC4v6 overestimates CHL more than 40 %.” could be interpreted to mean that for OC4v6 overestimates CHL by > 40% for all data points.

C1152

pg 2295 line 19: Change “reflect this phenomenon (Fig. 3a). Causes could be” to “reflect this phenomenon (Fig. 3a), possibly due to” line 22: Change “are excessively risen” to “are excessively raised” line 25: Possibly change “eventual coccolithophore” to something like “occasional coccolithophore” or “annual coccolithophore” to reflect the actual frequency of coccolithophore blooms in this region. line 28: Might be clearer to change “However, few spikes in the time series...” to “Additionally, a few spikes in the whole-Baltic time series...”.

pg 2298 line 5: Change “Fifteen years-long merged multi sensor daily CHL data” to “A fifteen-year merged-multi-sensor-daily dataset of CHL” line 14: Change “the dynamics was similar as in the Central Baltic” to “the dynamics were similar to the Central Baltic” lines 16-20: These would be supported by metrics from a cross correlation analysis. line 24: It is not just a higher R2 that would be preferable but a smaller RMS and data collection over a full range of regional conditions to avoid the situation seen at high CHL in the Skagerrak and Kattegat region.

pg 2306 Figure 2: This figure has a colour scale for density of points but there is no legend to show what the colour scale range. Also, it may be of help to the reader to show the linear regression on the plots in addition to the 1:1 line (as is done in figure 3) so that one can get a better understanding of the regression parameters at a glance.

pg2307 Figure 3: The caption states that the line of equal value is dashed and black, but in the plot it is solid and black. Either the text or the plot require correction. This plot also requires the addition of a colour scale legend.

---

Interactive comment on Ocean Sci. Discuss., 12, 2283, 2015.

C1153