Biogeosciences Discuss., 11, C566–C569, 2014
www.biogeosciences-discuss.net/11/C566/2014/

Biogeosciences
Discussions
Open Access

Interactive comment on "Satellite detection of multi-decadal time series of cyanobacteria accumulations in the Baltic Sea" *by* M. Kahru and R. Elmgren

Anonymous Referee #1

Received and published: 26 March 2014

The manuscript "Satellite detection of multi-decadal time series of cyanobacteria accumulations in the Baltic Sea" provides a potentially valuable contribution to the literature because it attempts to reconstruct the time series of surface accumulations using a multi sensor approach. There are some potentially interesting findings in the manuscript, particularly around trends in surface accumulation frequency and timing. However, there are some methodological issues that may impact the findings of the study. These should be addressed before any conclusions can be drawn. The primary methodological problem is the cross calibration/cross comparison of products between sensors, and whether the observed trends and patterns are a result of the different sensors or are a reflection of true occurrences. It is quite possible that the patterns

described in the results are nothing more than variations from the different sensors rather than measurements. This is somewhat supported by the EOFs shown in figure 14, where the greater variance in the time series occurs after the inclusion of more sensors. The authors need to demonstrate that this is not the case. There are a number of issues with the manuscript wrapped up in this. I will address them point by point below: 1. The communication of the methodology makes it unclear whether the same approach for cyanobacterial "detection" was taken for the different sensors. In some places the authors state that they used the simple red band approach as a way to cut across sensor differences. But in the methods section, they state that they used standard products for some of the sensor datasets. Which is it? The communication needs to be improved to make this methodological choice clear and consistent throughout the manuscript. 2. After improving the communication, the authors need to do a better job of comparing the different datasets. a. If the same approach (red band) was used, then the authors should demonstrate that the input datasets are comparable (e.g. scatter plots of surface reflectance or radiance between the different sensors). Does the red band look the same in all the different datasets? Please provide quantitative evidence. b. If a different approach was used for each data set, then the authors need to do a better job justifying the different approaches. As it is currently communicated, it sounds like an ad hoc approach to getting the imagery to best "meet." c. Regardless of the approach, I am concerned about the use of the fraction of cyanobacteria accumulations (FCA) in making the cross-sensor comparisons. By using temporally and large spatially binned data, the authors are potentially obscuring the comparability of the datasets by temporally and spatially "smoothing" their results. This may be result in higher R2 values, but that R2 is not informative of the true comparability. While it is important and challenging to summarize remote sensing image data in meaningful ways, those summaries should only be made after the inter comparisons. The FCA seems like a good way to summarize the information for further time series analyses, however, I would like to see a pixel-pixel or window-window comparison between products for the datasets that have the overlaps. While there may be some error in those comparisons due to differences in timing of acquisition and pixel size, I think it provides a far more transparent assessment of the comparability of the different sensors. 3. After the communication around methodological approach is addressed, the communication of the manuscript needs to be improved to explicitly address the assumptions inherent in the cyanobacteria product. The algorithms described in this study are NOT detecting cyanobacteria. They are detecting turbidity. There is strong ancillary evidence to support an interpretation of this turbidity as being caused by cyanobacteria in some places in the Baltic. But clearly this is not true everywhere, as evidenced by the need to mask out areas (e.g. the coast) where turbidity is caused by other mechanisms (e.g. sediment inputs, resuspension). In my opinion it is fine for the authors to make the assumption of cyanobacteria detection from turbidity based on a preponderance of evidence. However, this needs to be clearly and distinctly communicated in the manuscript as there are important implications for the reader. Namely, this algorithm is highly region specific, and is not likely to be widely extensible or generalizable. 4. There is some ambiguity around the coastal masking that needs to be addressed. How were coastal areas defined? What was the threshold in determining whether a bloom is too close to the coast? Sediment plumes from coastal areas can move offshore, and cyanobacteria can accumulate in coastal areas. These are two key processes determining serious confounding factors in your inferences. They should be addressed in the manuscript more explicitly. 5. The manuscript is lacking a conclusions section. In addition to these major methodological concerns, I have a number of smaller, more specific comments and questions about the text. I will list these below. 1. The introduction does not adequately reflect the primary contribution of the paper, which is summarized well in the last portion of the first paragraph in the discussion. Please consider reframing the introduction, specifically taking care to address what the research question and hypotheses of the study were. 2. P. 3323 Line 11: Plant-available nitrogen seems an inappropriate term here. I think “bioavailable” is more widely used. 3. P 3325 Lines 14-16: What does empirically mean? Was it compared against field data? Is a post hoc adjustment? 4. P3325 Line 26: 2°C? Is this the right value? 5. P 3326

Line 6: What tests were skipped? 6. P 3326 Line 20: showing imagery is helpful, but is only qualitative. Please provide a quantitative comparison of the two, on a pixel-pixel or window-window basis. 7. P 3327 Lines 1-2: 670nm is a well known chlorophyll absorption region. I would say that more reflectance at that band is caused not so much by high amounts of scattering as by less absorption. 8. P3328 Lines 5-7: This statement should be supported by quantitative evidence. “Agreement” can and should be quantified here. 9. P 3329 Lines 5-9: However, the quality of MERIS and SeaWIFS data for ocean color applications are much better. Just picking a valid daily observation may not be as good as picking the best valid daily observation. Is there a potential for obscuring and reducing the information content of the time series by electing not to use these sensors? How was this determined? Would you have a better result if you were to use some kind of ensemble of the multiple sensors wherever possible? 10. P3332 Line 9: “accurately detected” This has not been quantified. Please provide the statistics. 11. P3332 Line 10 & Figure 8: It appears to me that the widespread accumulations have much poorer matchups. This is worth delving into and explaining a little bit more, especially in the context of dataset comparisons. 12. P 3333: Is the dependent variable (detections) in the logistic regression from the FCAs or from the matching pixels? 13. P 3336 Line 24-25: I am not convinced this was demonstrated by this study. 14. P 3338 Line 24-26 & Figure 9: It may not be completely legitimate to use logistic regression to compare different independent variables and you may wind up underestimating the effect of the predictor variables. See [Mood, 2010] for a nice commentary on the matter. Mood, C. (2010), Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It, European Sociological Review, 26(1), 67-82.

Interactive comment on Biogeosciences Discuss., 11, 3319, 2014.