Biogeosciences Discuss., 11, C653–C665, 2014
www.biogeosciences-discuss.net/11/C653/2014/

Biogeosciences Discussions

BGD

11, C653–C665, 2014

Interactive Comment

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

M. Kahru and R. Elmgren

mkahru@ucsd.edu

Received and published: 1 April 2014

We appreciate the comments made by Anonymous Referee #1 and we will respond point by point below. Before doing that we would like to emphasize that to our knowledge this is the longest satellite-derived time series in biological oceanography anywhere in the world and that it responds to the problem that has been challenging Baltic oceanographers at least since 1970s (Horstmann, 1975): Are the cyanobacteria accumulations that cover huge swaths of the Baltic Sea in July-August increasing in frequency and extent and are they being favored due to eutrophication? The annual cyanobacteria bloom in the Baltic is a most conspicuous oceanographic phenomenon that - when it occurs – attracts the attention of the general public and is often reported in front page articles of local newspapers. Due to the huge spatio-temporal variability of the cyanobacteria distributions and the limited sampling capability of regular ship-based monitoring during the approximately 2-month period when these accumulations occur, oceanographers have not been able to answer the questions put forward at least 40 years ago (Finni et al. 2001). In this work we tried to compile a time series that is as long as possible. Well calibrated satellite ocean color sensors covering the July-August period have been in operation since 1998 and using only these data would be insufficient to analyze multidecadal variability. We therefore included data from earlier sensors and made a major effort to make the data from the "old" and "new" sensors compatible. Below are our responses to individual comments by Anonymous Referee #1. "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." We strongly disagree. In fact, sections 2.4, 2.5 and 3.1 as well as figures 7, 8 and 9 show that the estimates from different satellite sensors are in good agreement with each other and also with in situ measurements. Figure 14 shows changes in main spatial modes (EOF1 and EOF2) which are dominating since 1999 but were weak in earlier time period. This north-east versus south-west pattern has been dominant since 1999 but the magnitude of these patterns was obviously close to zero during the period (1985-1997) when the accumulations were very rare or missing. Also, the distribution patterns in earlier years were different from those in recent years. This has nothing to do with "the inclusion of more sensors" as the products from different sensors agree well with each other (Fig. 7). "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?" In Introduction we claim that there are no "standard satellite products" available for detecting cyanobacteria in the Baltic Sea. Therefore we created our own product based on high remote sensing reflectance (Rrs) of the approximately 665 nm band for ocean color sensors and on the high albedo of the AVHRR channel 1. Sections 2.1, 2.2 and 2.3 give detailed description of the methods. "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." Space agencies (NASA and ESA) devote significant resources to vicarious calibration of the water leaving radiances (or Rrs) produced by their sensors. We did not consider it necessary to validate the Rrs products of individual sensors as our derived estimates of cyanobacteria fractions agreed very well (section 2.4, Fig. 7). Meaningful comparison of the Rrs667/Rrs670 of individual pixels between different satellite sensors is complicated as in each location, a good overpass of a typical ocean color sensor occurs every 2nd day; the local overpass times are different (e.g. MODIST at 10:30 and MODISA at 13:30). Also, while the differences due to the different viewing geometry and solar geometry are supposed to be corrected for, some residual errors are inevitably present, e.g. when comparing a pixel in the middle of the swath with the same ground location but observed at the edge of a satellite swath. The accumulations are expected to drift over many pixels during the time lag between satellite overpasses, even if we have corresponding satellite passes in the same day. Therefore, errors that are difficult to quantify are present in these comparisons. Figs. A1 and A2 show scatter diagrams of the Rrs667 measured by MODIST versus that measured by MODISA and of SeaWiFS versus that of MODISA. All these data are from July 10, 2005 that corresponds roughly to the images in Figures 1, 4, 5, 6. Fig. A1 shows that most of the Rrs667/Rrs670 points of different sensors fall close to the unity line but there is also considerable scatter. More scatter is evident in the SeaWiFS data compared to MODIST data. That is probably explained by the lower SNR of SeaWiFS. Additionally, it may be significant

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

that SeaWiFS data here are based on GAC (4 km ground resolution) pixels instead of the 1-km ground resolution pixels of MODISA and MODIST. As mentioned above, some of the scatter is due to the spatial and temporal aliasing as accumulations may drift during the time lag between overpasses of the different satellite sensors. For the detection of cyanobacteria accumulations we are concerned with the high end values and can ignore the increased scatter in the low end. We don't consider it necessary to include those plots of standard NASA variables in this paper as we have validated the derived products shown in Fig. 7. "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."" As described in detail in sections 2.1, 2.2 and 2.3 we used remote sensing reflectance (Rrs) of the approximately 670 nm band for ocean color sensors and the albedo of the AVHRR channel 1. We have also explained that due to the relatively poor spectral calibration and low SNR of the AVHRR sensors, accurate calculations of Rrs are very difficult (practically impossible) and therefore we used primarily channel 1 albedo. We have spent considerable amount of text justifying the use of the specific algorithms that are different for AVHRR, namely the spectrally wide-band (∼100 nm) and low signal-to-noise ratio (SNR) channels compared to modern ocean color sensors. "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.” As already mentioned above, direct comparison of pixel by pixel values between different satellite sensors is not possible without major errors due to differences in timing and viewing/solar geometries. In each location, a good overpass of a typical ocean color sensor occurs every 2nd day. The local overpass times are different (e.g. MODIST at 10:30 and MODISA at 13:30), therefore it is expected that the accumulations may drift over many pixels during the time lag between satellite overpasses. The huge small-scale variability of the accumulations is well known and differences of up to 2 orders of magnitude are common over distances of a few meters. While the within-pixel variability is averaged in the satellite pixel of approximately 1 km2 , a typical speed of drift by the accumulations of $\sim$10 km per day adds an error that makes it difficult to interpret the differences observed in such pixel by pixel comparisons. However, as shown in Fig. A1, the different ocean color sensors have no significant bias. “The algorithms described in this study are NOT detecting cyanobacteria. They are detecting turbidity.” This has been emphasized repeatedly, starting from line 7 in the Abstract. “Namely, this algorithm is highly region specific, and is not likely to be widely extensible or generalizable.”

We don’t see a problem here. We solve a practical problem that has been troubling Baltic oceanographers for the last 40 years. Our method is applied to accumulations consisting primarily of Nodularia sp. in the Baltic Sea. Major Nodularia accumulations are known to occur ONLY in the Baltic Sea (minor accumulations have also been observed in the past in a small estuary in Australia and some inland saline lakes).

“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.”

Coastal areas have been masked and it is clearly stated and clearly shown in the map of Fig. 2. The masked areas were determined by depth being less than 30 m. We added that clarification to the text. Because Nodularia spumigena blooms typically occur in the open sea (e.g.,Wasmund, 1997), the exclusion of the coastal zone did not greatly reduce the total detected bloom area.

"5. The manuscript is lacking a conclusions section."

We have added a Conclusions section.

"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."

We have rewritten the first paragraph of the Introduction to accommodate this remark, and instead shortened the discussion.

"2. P. 3323 Line 11: Plant-available nitrogen seems an inappropriate term here. I think "bioavailable" is more widely used."

Since cyanobacteria are not plants, strictly speaking, we replaced "plant-available" with "bio-available".

"3. P3325 Lines 14-16: What does empirically mean? Was it compared against field data? Is a post hoc adjustment? "

Cyanobacteria accumulations have characteristic spatial structures that are well visible and demonstrated in Figs. 1, 4, 5, 6. The albedo limits of 2.3-4% mentioned here have been used in the past analysis (Kahru et al., 1993, 1994, 2000, 2007; Kahru, 1997) and were determined visually as the lower limit of detection of the characteristic spatial structures and as the approximate lower limit of clouds.

"4. P3325 Line 26: 2C? Is this the right value?"

BGD

11, C653–C665, 2014

Yes, that’s what we have used.

“5. P 3326 Line 6: What tests were skipped? “

We clarified the sentence and specified that the tests involving AVHRR infrared bands 4 and 5 were skipped. These tests did not have a significant impact.

“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.“

Fig. A1 above provides quantitative comparisons of the Rrs values in these images. As already mentioned above, direct comparison of pixel by pixel values between different satellite sensors is not possible without major errors due to differences in timing and viewing/solar geometries. In each location, a good overpass of a typical ocean color sensor occurs every 2nd day. The local overpass times are different (e.g. MODIST at 10:30 and MODISA at 13:30), therefore it is expected that the accumulations may drift over many pixels during the time lag between satellite overpasses and the accumulations are expected to drift ∼10 km per day. This makes it difficult to interpret the differences observed in such pixel by pixel comparisons. As shown in the added Fig. A1 (above in this document) Rrs values correspond well between different satellite sensors in a statistical sense (R2 > 0.9 and slope ∼ 1) but the scatter is difficult to interpret. “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. “

That is not correct. The decrease in reflectance due to increased chlorophyll absorption is a minor effect compared to the huge increase in reflectance due to increased scattering.

“8. P3328 Lines 5-7: This statement should be supported by quantitative evidence. “Agreement” can and should be quantified here.”

As already mentioned above, direct comparison of pixel by pixel values between dif-

BGD

11, C653–C665, 2014

Interactive Comment

ferent satellite sensors is not possible without considerable errors due to differences in timing. We use this example of images from July 10-11, 2005 to show a visual comparison of images. Our quantitative time series are using monthly or 2-month (July-August) FCA and TA estimates (Fig. 12, 13) and therefore we are also using monthly FCA to compare the performance of different sensors (Fig. 7).

"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?"

We are not sure on what basis the reviewer claims that "the quality of MERIS and SeaWIFS data for ocean color applications are much better". As shown by our analysis, all modern ocean color sensors used in this work (SeaWiFS, MODIST, MODISA, VIIRS) have very similar performance and the accumulations estimated with these sensors agree very well with each other (e.g. Fig. 7). We therefore used data from ALL available sensors (excluding MERIS due to practical consideration as during the time o f this work access to Level-2 MERIS data was cumbersome). As shown in Table 4, during the time period of available MERIS data (2002-2011) we have over 300 datasets per July-August period from other sensors; we therefore concluded that adding MERIS to that mix would not add much to the FCA and TA variables that were the variables that we were targeting in this work. SeaWiFS data were used during the 1998-2004 period (when full-resolution MLAC data were available) but were not used in 2005-2010 when the high-resolution (MLAC) data were not available. The available GAC data had lower spatial resolution and were deemed unnecessary due to the large amount of high-resolution data from MODIST and MODISA during that period. As mentioned above, we did not use MERIS data as we had enough coverage during the years of MERIS operation but just for curiosity we completed a comparison of the MERIS Rrs670 data versus similar MODISA data as we did in Fig. A1 for other sensors. As shown in Fig. A3, MERIS/MODISA comparison shows much higher scatter (at least on that particulate day) compared to the comparison with other sensors in Fig. A1 and A2. The increased scatter is difficult to explain but may be caused by the different processing methods used by ESA for MERIS compared to the methods used by NASA for all the other sensors in Fig. A1 and A2.

“10. P3332 Line 9: “accurately detected” This has not been quantified. Please provide the statistics.“

What we meant was that the peaks corresponding to the accumulations in the Bay of Gdansk area (Fig. 8, A-B) lined up very well considering the quite different spatial sampling of the ship and satellite products. To be more accurate, we rephrased the sentence to: “The accumulations started to be detectable in the beginning of July in the Bay of Gdansk (near 19 °E longitude) and the narrow patches of accumulations were lined up very well on both the ship transect and on the satellite map (Fig. 8A-B).”

“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. “

The accumulations are very patchy spatially and probably also vertically. While the satellite pixel averages over approximately 1 km2, the exact location of the ship transect within the pixel introduces significant variability. Also, the ship measurements are taken at the depth of 5-6 m while the surface accumulations seen by the satellite are more affected by the accumulations at the very surface or just under it. The implications of depth variability are discussed with a reference to a recent paper by Groetsch et al. (2012). We see no reason why the differences between satellite sensors should significantly influence the differences between the ship transect and the satellite measurements.

“12. P 3333: Is the dependent variable (detections) in the logistic regression from the

FCAs or from the matching pixels? “

The dependent variable is the set of detected/undetected values of the corresponding satellite image. FCA is calculated as a statistical measure over a period of time, e.g. 1 month or 2 months (July-August). We clarified that in the rephrased sentence: “As our output variable, i.e. the presence or absence of accumulation in the satellite pixel nearest to the ship measurement, was binary. . .”

“13. P 3336 Line 24-25: I am not convinced this was demonstrated by this study. “

We agree. We did not show that in this study. We added a reference to Gregg et al. (2009) demonstrating this problem.

“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.”

We agree that logistic regression is tricky to use but in our case it is completely legitimate to compare the estimates as our independent variables are part of the same dataset and used against the same sample of the dependent variable. Moreover, our conclusion that phycocyanin fluorescence is a better indicator of cyanobacteria concentrations than chlorophyll-a fluorescence was also supported by visual inspection of the transects as well as by previously reported comparison between ship-based data with microscopic cyanobacteria counts (Seppala et al., 2007).

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

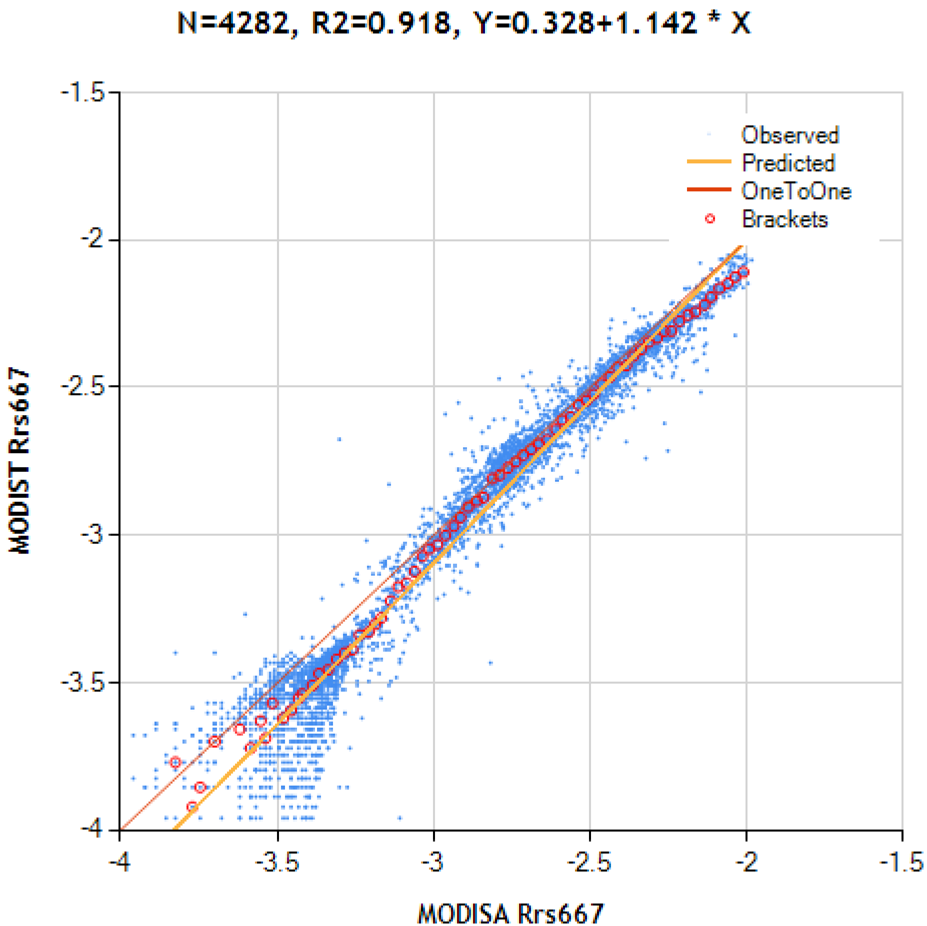

Fig. 1. Pixel-by-pixel c comparison of the Rrs667/Rrs670 data over the Baltic Sea (see Fig. 2A) on July 10, 2005. A, MODIST versus MODISA. B, SeaWiFS versus MODISA. All axes have log10 scaling.

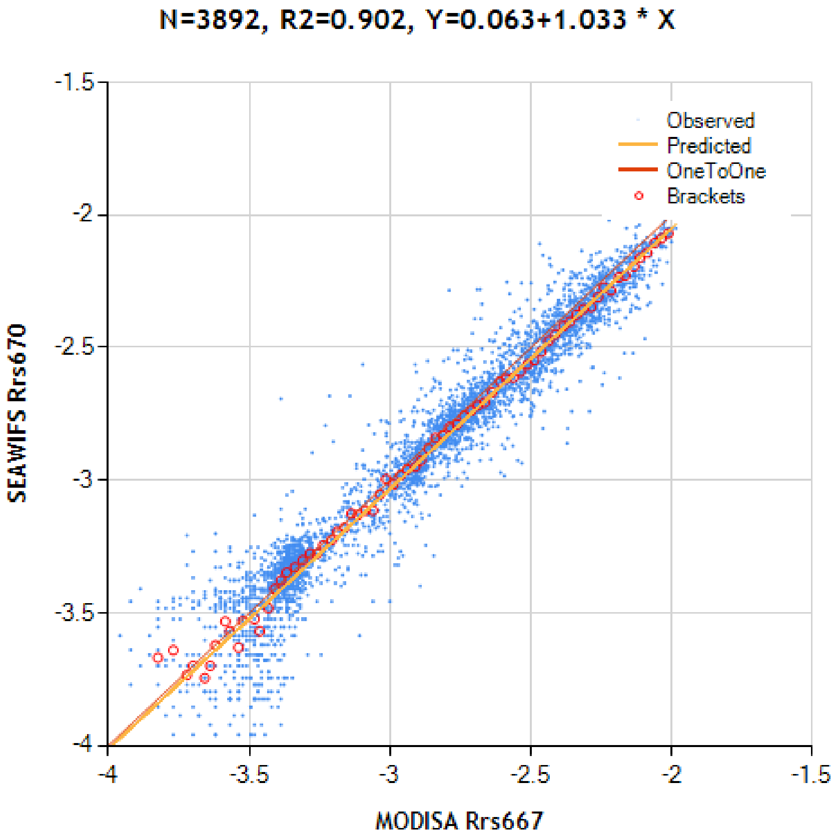

Fig. 2. . B, SeaWiFS versus MODISA. All axes have log10 scaling. The values plotted here are from binned and mapped 9-km pixels

BGD

11, C653–C665, 2014

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

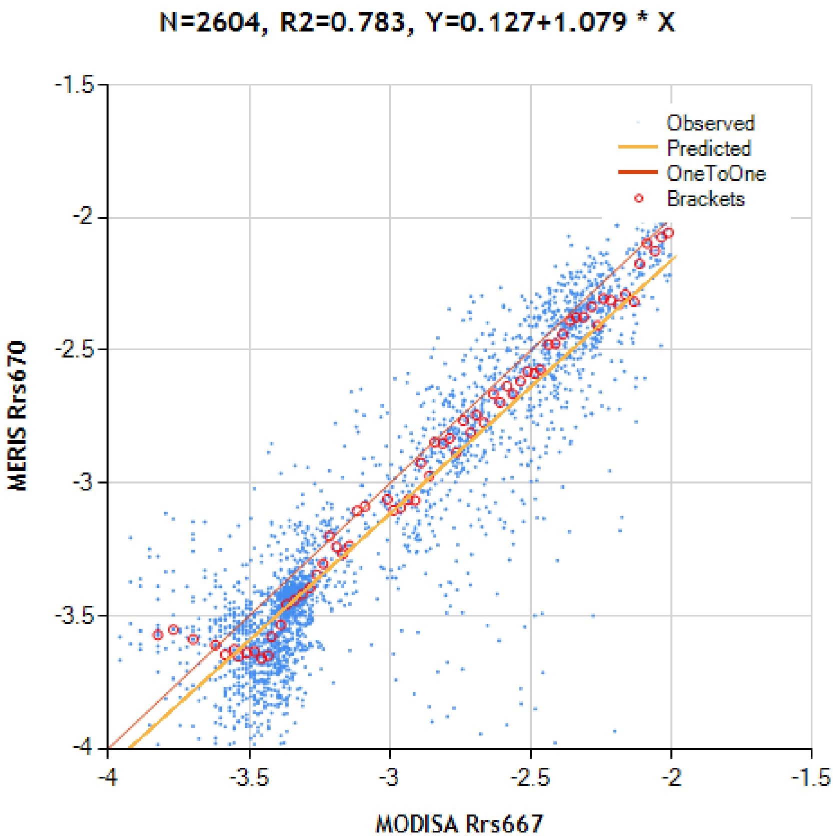

Fig. 3. As Fig. A1 but showing MERIS Rrs670 versus MODISA Rrs667 on July 10, 2005 over the Baltic Sea

BGD
11, C653–C665, 2014