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

M. Kahru and R. Elmgren

mkahru@ucsd.edu

Received and published: 6 April 2014

We thank S. Kratzer for reading our paper and making comments. We respond point by point below. Before doing that we would like to emphasize that as our goal was to build a multidecadal time series, we had to use methods that are reasonably compatible for satellite sensors over 34 years. We do not have the luxury of using the most advanced algorithm (even if that existed) for the most advanced sensor. Also, we argue in the Discussion that the advanced algorithms that are specifically designed to detect cyanobacteria are extremely sensitive to the existence of the surface scum. As the exact vertical distribution of the accumulations can change dramatically over the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

course of ~15 minutes as winds pick up or slow down, this extreme sensitivity is NOT desirable for building long time series and our less sensitive methods are therefore more desirable. We have stated in the manuscript that the accuracy and sensitivity of the earlier sensors (AVHRR) are not the same as the most recent sensors and therefore we have made a major effort to make the data from the "old" and "new" sensors compatible. As shown in Fig. 7 we have achieved a good correspondence of the FCA between different sensors. Below are our responses to individual comments.

"For the new plots the authors included in their response for showing the difference between MODIS Aqua and MODIS terra, I suggest that you quote the correlation coefficient instead of r2, the coefficient of determination, because we car not talking about a regression here, but about a correlation. How well does the data correlate?"

Correlation coefficient is just the square root of R2. We can report them all. For MODIST/MODISA comparison: R2=0.918, R=0.958, RMSE=4.97E-04; for SeaWiFS/MODISA comparison: R2=0.902, R=0.950, RMSE=6.06E-04; for MERIS/MODISA: R2= 0.783, R=0.885, RMSE=8.37E-04. We don't think that these statistics are crucial in this paper. We have explained in the response to Referee#1 that there are many sources of variability here (e.g. due to the temporal lags). For our paper it is important to just emphasize that there is NO SIGNIFICAN BIAS between the sensors. Moreover, we have validated the compatibility of different sensors using the monthly FCA: SeaWiFS versus MODIST, N = 135 months; MODISA versus MODIST, N = 243 months and VIIRS versus MODIST, N = 54 months and AVHRR versus other sensors with N = 108 months. The error of the monthly FCA was estimated as (1) the mean absolute difference and (2) the median absolute difference between FCA values of different sensors. For MODISA, MODIST and VIIRS the mean absolute differences were from 1.1 to 1.9% FCA and the median absolute differences were between 0.4 and 0.6% FCA. For SeaWiFS the respective errors were slightly higher (2.0% and 1.1% FCA, respectively). As shown in Fig. 7, there is no significant bias between results of the ocean color sensors and the differences between AVHRR and ocean color

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

sensors are corrected by an empirical fit.

"In general, MODIS Terra is known to be unstable and prone to calibration errors, so it is important to assess its response."

We have been working with MODIS-Terra daily since its early versions of data processing in 1999 and are very well aware of the calibration problems involving particularly the 412 and 443 nm bands. It appears that the 667 nm band of MODIST is very similar in performance to the 667 nm band on MODISA. Similarly, the so-called land bands with 250 m and 550 m resolution which are not used in typical ocean color algorithms are also as stable as the corresponding bands on MODISA. We have used those MODIST bands (e.g. Fig. 1 and Fig. 6) and according to our analysis there is no significant difference between MODIST and MODISA when using Rrs667. The correlation of Rrs667 on MODISA is actually better with MODIST compared to either SeaWiFS or MERIS. However, our amphasis is on monthly FCA rather than on pixel by pixel Rrs667 and we have presented extensive statistics validating our monthly FCA values.

"In general, I feel there is a quite loose use of the terminology of reflectance in the paper. It is not clear to me what physical units you use. Sometimes you call it albedo, 'brightness', sometimes water-leaving radiance, and sometimes you use the term reflectance. Please define clearly what you mean in whenever you write about any of these physical quantities."

As clearly stated in sections 2.1, 2.2 and 2.3 we used remote sensing reflectance (Rrs) of the approximately 665 nm band for ocean color sensors and the albedo of the AVHRR channel 1. We developed our algorithm when the standard output of NASA ocean color data was normalized water leaving radiance (nLw) instead of Rrs. Therefore we have some references to nLw. The conversion between Rrs and nLw is straightforward: nLw (lambda) = solar_irradiance (lambda) * Rrs (lambda) and this is included in the text.

"I would like to go back to using a broad AVHRR channel to derive information on the

reflectance of cyanobacteria. You stated yourself that you did not do any atmospheric correction to the AVHRR data, and that you used the information from channel 1 as an indicator of 'brightness' or albedo."

It has been stated clearly that we used albedo of the AVHRR channel 1, i.e. not Rrs. Calculating Rrs of AVHRR data is error prone due to uncertainty in calibration and was not attempted in this work.

"Clouds etc. were masked out using the information from other channels. I would like to point out that some of the information that is included in your so-called 'albedo' also includes information about atmospheric aerosols, as you are using TOA radiance (is it radiance you are using?). So, some of the variability you get in the AVHRR image may be due to the variability in atmospheric aerosols or SPM in the water. This may also affect the final conclusions you draw, i.e. an earlier on-set of the bloom. Can you be sure that this is not caused by fluctuations of atmospheric aerosols? I also feel that you have not sufficiently addressed the different sensitivity of the different sensors. Sensitivity not only refers to signal-to-noise ratio, but also to the actual radiometric sensitivity and the dynamic range of the instrument, i.e. how well it is adapted to sense dark or highly reflective targets."

When the launch of SeaWiFS was delayed in 1997, we were involved in a NASA contingency plan to use AVHRR to derive water leaving radiances while no proper ocean color sensor were not available. After the successful launch of SeaWiFS these plans were abandoned as calculating Rrs from AVHR data was deemed too error-prone. As stated in section 2.2.1, the low sensitivity and poor calibration accuracy of AVHRR's two broadband spectral channels make accurate atmospheric correction difficult and highly error-prone. We are not aware of any significant effort for vicarious calibration of AVHRR that can be useful for ocean waters. Even with the best available calibration coefficients, atmospheric correction of AVHRR bands often resulted in physically impossible negative values of the water-leaving radiance (Stumpf and Fryer, 1997). Fortunately for this study, the cyanobacteria accumulations in the Baltic have some of

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

the highest reflectance of any blooms in the world which makes it possible to estimate the extent of the accumulations even without accurate atmospheric correction. Figures posted in our response to Referee#1 show that satellite-derived Rrs667 covers more than 2 orders of magnitude and we are only interested in the top part of it. Figures 1, 4, 5, 6 in the manuscript have all been created WITHOUT removing atmospheric aerosols and have been corrected only for for the gaseous absorption and Rayleigh scattering. We assume that when detecting dense surface accumulations we can safely ignore the smaller differences caused by aerosols. We do visually examine each and every single image and remove features that seem erroneous. It is well known that correction for atmospheric aerosols is prone to large errors in turbid and high-biomass waters. In fact, atmospheric correction of aerosols has been deemed unnecessary and error-prone even when using the high-quality MERIS data (Matthews et al. 2012) for detecting high-concentration blooms.

"MERIS, e.g. has a very high dynamic range and is adapted both to highly reflective land and dark ocean targets; SeaWiFS e.g. needs a few pixels to adapt from highly reflective (land or clouds or cyanobacteria blooms) to dark water pixels; this is called bight-pixel recovery effect."

Yes, this is well known. We exclude the coastal areas and also expand the cloud mask to eliminate suspicious pixels along cloud edges.

"The AVHRR has very broad channels which may increase its sensitivity, but at the same time it makes it impossible to detect the difference of different optical constituents, e.g. SPM versus cyanobacteria (which are both highly reflective). Turbidity is not the same as cyanobacteria blooms!"

Yes, this has been discussed extensively in this paper. We are able to detect cyanobacteria accumulations in the open Baltic Sea as they are the dominant phenomenon that produces high reflectance in the open Baltic Sea in July-August.

"Turbidity is a physical measure (scatter at a certain angle) which may also be caused

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

by particles. In the southern and eastern Baltic there is a high load of suspended particulate matter and the sediment plumes may reach 100-120 km off-coast, whereas research in the Northern Baltic Proper showed that the extent of sediment influence is in the range of 10-15 km (Kratzer and Tett, 2008). That paper shows that the effect of inorganics tends towards zero at 10-15 km off coast."

We are well aware of that. Uncertainty in shallow and coastal regions was the reason for excluding those. We have been careful to eliminate sediment plumes originating from the coast.

"The cyanobacteria are often driven into bays and coastal areas by wind, so one should really aim to establish methods that are also applicable in the coastal regions where most of the production happens, and where it mostly affects tourism and maybe even the well-being of animals. Maybe one should look into different pattern recognition methods that differ between the typical patterns of cyanobacteria blooms and effects from coastal run-off (inorganics), and if not included in the paper, this should at least be discussed."

This is a wonderful idea for a future project but is outside the scope of this paper. As emphasized previously, the goal was to create the longest time series possible using available satellite data. Studying bays and coastal areas requires the use of completely different imagery, e.g. of the type of Landsat TM with ∼20-30 m resolution. This is very different from the wide-swath imagery with the ∼1 km resolution used in this work. Even the 250 m bands of MODIS and ∼300 m resolution FRS data of MERIS do not have the enough spatial resolution to properly resolve the variability in bays and coastal areas. Sensors with ∼20-30 m resolution or those with even higher resolution must be used. Unfortunately, the revisit time of these sensors is ∼14-16 days and the chances are that many of these days happen to be cloudy and provide no data. Therefore, obtaining a representative time series with a few high-resolution images per season is highly unlikely. While the coastal areas are certainly important, for the multidecadal time series of the Nodularia blooms happening mostly in the open Baltic, they are not

crucial.

"It may well be that some of the features that you see on the highly sensitive ocean-colour images really derive from suspended matter (which in this case also indicate the typical eddies and other meso-scale features), and I wonder if the shift of the starting point of the blooms that you detect are maybe partially related to the increased sensitivity of ocean colour sensors and also the improved dynamic range of MERIS."

As explained in the Introduction and Discussion of the paper, MERIS data were not used in this work. Therefore the improved dynamic range of MERIS is irrelevant. We are applying thresholding and classifying only the highest reflectance areas as cyanobacteria accumulations. Therefore, the increased sensitivity at the lower end of reflectance is not important. Various errors are caused by different orbits, viewing and solar geometries, etc. The statistics that we present show no significant bias between the modern sensors and no bias with AVHRR after conversion and when detecting large-scale blooms that matter most in the compilation of the FCA statistics.

"How sure can you be that these changes in instrument specification do not affect your results?"

The statistics given in section 2.4 and Fig. 7 demonstrate clearly that we have no detectable bias between different ocean color sensors and that we have converted the output from AVHRR to correspond to the results obtained by the more sensitive ocean color sensors.

"Other comments to improve the quality of writing I also find the paper rather difficult to read. The methods part seems rather technical, but I feel it does not sufficiently describe the methods applied to each sensor -as already commented by reviewer 1#, and as I already mentioned you need to define the physical units you use."

We have revised the text to make it easier to read. The methods are the same for the ocean color sensors (high Rrs at approximately 667 nm) and the high albedo of

the channel 1 for AVHRR. Sections 2.1, 2.2 and 2.3 give detailed description of the methods. We believe that the simple conversion between nLw and Rrs that is given in section 2.2.2 is sufficient to explain the relationship with the standard data format was used earlier (i.e. nLw).

"Also, to compare the results from satellite-derived information to ship transects with a water intake at 5 m depth is rather misleading, as the cyanobacteria tend to accumulate in the top meter during high isolation and strong stratification."

We agree that there are significant differences between the surface and 5-m depth but it is quite obvious that at the horizontal scales of hundreds of kilometers there is high correlation between satellite measurements and the ship transects at 5 m depth. In fact, the correspondence between some of the features is STRIKING (e.g. Fig. 8A) and we certainly do not consider that "misleading". The ship transects are not used in any of the satellite time series, but are only used to demonstrate that patterns observed by our satellite methods are similar to the patterns observed by ship transects. Obviously, the ship-measured signal at 5-m depth is correlated with the satellite-measured surface but the correlation may be locally higher or lower depending on the vertical stratification (e.g. Groetsch et al. 2012). It has been known for a long time (e.g. Kahru et al. 1993) that the cyanobacteria accumulations detected by satellite correlate well with chlorophyll fluorescence at the depth of ∼5 m (see Figs. 2, 3, 5 in Kahru et al. 1993; http://spg.ucsd.edu/People/Mati/1993_Kahru_et_al_Cyano_cause_heating_MEPS.pdf). As a result of the analysis done in this paper, we now know that phycocyanin fluorescence is even better correlated with the satellite-detected accumulations and that is in good agreement with the pigment composition and fluorescence properties of the Baltic cyanobacteria (Seppälä et al. 2007).

"One should really use the same spatial resolution for the comparison- i.e. use the 1 km resolution for all sensors. This is because with the high horizontal heterogeneity of the cyanobacteria blooms you get different accuracies for the different spatial resolution."

This is exactly why we used 1 km satellite data for ALL the sensors in the quantitative analysis! To make it clearer, we have added 2 more sentences to section 2.1: "While the MODIS sensors on Terra and Aqua have bands with 250 m and 500 m resolution (e.g. Fig. 1), these bands have typically lower signal to noise ratio. For compatibility between different sensors, all satellite data used in the quantitative analysis in this work had approximately 1 km spatial resolution."

"The accuracy for MERIS e.g. may improve by 100% when going from 1 km resolution down to 300 m resolution. "

It depends how you define "accuracy". Anyway, we did NOT use MERIS data at all, therefore this is irrelevant.

"If you go from 1km to 4 km resolution you therefore may have a large decrease in accuracy. In order to make the methodology more coherent 1 km resolution should be used for the whole data set!"

That's exactly why we DID NOT use SeaWiFS data in 2005-2010 when only the GAC data were available.

"I also see a great disadvantage of using 5x5 pixel averages to compare to the ship measurements. Usually, in such heterogeneous waters, one extracts 3x3 pixels (5x5) is really only appropriate in open ocean where you can assume that the water does does not significantly change over a 5 km distance. But in this scenario it really seems not appropriate!"

The time lag between the ship sampling along the transect between Helsinki and Travemünde is variable and as accumulations may drift around ∼10 km during that time lag, we used the 5x5 pixel neighborhood. We also made the same analysis with 3 x 3 pixel window and found no significant differences.

"General comments: In general I would also try to avoid writing in the first person in a scientific article. It is ok on the odd occasion, but I find the frequent use of 'we' very

disturbing, and I would like you to consider to change to a more neutral and factual way of describing your methodology and results."

This is a personal choice and we do not find this "disturbing". In fact, many scientific journals actually encourage the more personal style instead of the anonymous "was done" style.

"For example: 'While we have a better data coverage' can be easily rephrased to 'whilst there is better data coverage' – it is not necessary to include the word 'we' here at all! Please check and modify the whole document for this!"

In this case we completely agree and we have changed the sentence to "there is better data coverage".

"Page 3322 line 11: use biologically available rather than plant-available or bio-available"

We have already changed that to "bio-available". We will change it to "biologically available".

"Page 3322 line 25: Aphanizomenon sp. often dominates deeper layers in the water column (it is adapted to lower light conditions) ".

The sentence is "While the surface accumulations consist primarily of Nodularia spumigena, other species of cyanobacteria, primarily Aphanizomenon sp., often dominate in the water column (Hajdu et al., 2007; Rolff et al., 2007)." This is exactly what we want to say. We see no reason to change this. We just state the facts here.

"Page 3322 line 2: what do you mean with 'brightness'?"

Brightness is an attribute of visual perception in which a source appears to be radiating or reflecting light. For example, we easily recognize the accumulations by brightness when looking down at the Baltic Sea from an airplane. This is the basic idea of detecting the accumulations. We could use reflectance here but we consciously tried to avoid

the use of more specific quantities, e.g. Rrs here.

"Page 3322 line 26: MERIS operated from 2002-2012; a time series only based on MERIS would be too short"

Yes, that is correct. We don't see anything wrong with that.

"Page 3323 end of introduction: include here the aims of your study, try to avoid 'we'

Thanks for the suggestion. We added the following sentence to Introduction: "The aim of this study was to create a quantitative multidecadal time series of cyanobacteria accumulation characteristics in the Baltic Sea using those compatible algorithms."

"Page 3323 line 24 use 'highly reflective' instead of bright blooms; brightness is not what we detect, it is the change in reflectance; brightness is not really a physical measure or quantity. What are you quantities what are your units?"

We agree and we have changed "bright" to "highly reflective".

"Page 3323 in the last paragraph you talk about the poor sensitivity and the low spectral resolution of AVHRR; I feel this belongs into the discussion, as this may affect your results, but you do not really have a clear way to correct for these effects."

We already have several sentences on AVHRR in the Introduction and in the Discussion. We have extensively discussed the problems with AVHRR and the method that we used to converting FCA estimated with AVHRR to FCA compatible with ocean color sensors. We have presented evidence (e.g. Fig. 7D) that the detection of large-scale accumulations by AVHRR is compatible with the detection by ocean color sensors.

"Page 3324: several times 'we' – is not necessary-please use passive form instead"

We already changed one case according to your request and we don't see any other cases that need changes.

"Page 3325 line 18: I thought that SeaWiFS does not have a 4 km 'mode' – it always

measures in 1 km resolution- the data is binned to 4 km after reception (level 3 product)"

First, you probably refer to page 3324 and not page 3325. Second, our statement that we did not use SeaWiFS data when only the GAC mode (4 km resolution) was available is correct. We see no reason to discuss the specifics of the GAC mode when we did not use it. The GAC mode picks every 4th pixel (in line) and every 4th line. It has nothing to do with "binning" as we did not use Level-3 products in this work.

"Page 3327: this full page should be rewritten so it is easier to follow. Some of the sentences from below need to be moved up for logical consistency. How do you define your thresholds? You do not correct for the band shifts – what effect may this have?What are your underlying assumptions – that the effects are insignificant? Can you be sure?"

We have revised the text to make it read better. The threshold value for defining turbid waters is that used by NASA ($Rrs670 > 0.012$ sr-1). If by "band shifts" you mean the slight differences between the bands of different sensors specified by the nominal (center) wavelength, e.g. 667 nm or 670 nm, then these small differences are irrelevant in this study. The FCA values obtained during temporally overlapping periods (Fig. 7) show the compatibility of the FCA estimates.

"A flag is not 'set' but 'raised'; if you use the term 'set' it sounds like it is actively done by the operator- if it is raised- it sounds like it is done by an automatic process, e.g. by reaching a certain threshold."

"Set" is used extensively in the appropriate technical literature, e.g. http://oceancolor.gsfc.nasa.gov/VALIDATION/flags.html

"What are your thresholds?"

The threshold is given in Page 3327, line 13: $Rrs670 > 0.012$ sr-1

"Use term 'particle backscatter'"

We changed it to "strong backscatter of particles that are either in the water column near the surface or directly at the surface".

"Line 27: instead of 'were manually filled' use term 'manually denoted' or designated? Filled sounds strange."

"Fill" is well-known operation in digital image analysis, e.g. filling inside a contour, and we don't see anything strange about it. The other terms ("denote" and "designate") seem to be less appropriate here.

"Page 3328 The Gulf of Riga and Gdansk are known to have high SPM loads.Quote some of the relevant papers!"

We added a reference to Liblik and Lips, 2011.

"Line 5. Which algorithms???"

We changed to "algorithms described in this section".

"Line 10: how did you choose your threshold value? Why just this value? Any statistics to quote?"

We used this threshold value as it produced patterns that were consistent with the patterns of Nodularia accumulations.

"Page 3328: a map is usually projected, not registered".

"Registered" is a term used when mapping data from an irregular grid of pixels (e.g. satellite Level-2 data) to a defined map.

"Page 3330 from line 21: does this not belong in the discussion?"

This single sentence is used to introduce the presentation of numerical values comparing FCA estimated with AVHRR. Discussion has more on this.

"Page 3333 what do you mean with the relationship is less tight?"

It refers to what is illustrated in Fig. 9, i.e. the probability of detecting accumulations with Turbidity and particularly with Chl fluorescence is lower than that with phycocyanin fluorescence over most of the variability range of the independent variable.

───────────────────────────────

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