Biogeosciences Discuss., 10, C1428–C1439, 2013
www.biogeosciences-discuss.net/10/C1428/2013/

Biogeosciences Discussions

Interactive comment on "Satellite views of global phytoplankton community distributions using an empirical algorithm and a numerical model" *by* C. S. Rousseaux et al.

C. S. Rousseaux et al.

cecile.s.rousseaux@nasa.gov

Received and published: 1 May 2013

The reviewers comments are in bold, our reply are in italics

After a careful reading of the paper entitled "Satellite views of global phytoplankton community distributions using an empirical algorithm and a numerical model", I inform that I recommend major corrections before publication in "Biogeosciences Discussions". You will find below my analysis of this manuscript. Firs, the title of the manuscript is not in agreement with the text and objectives of the work presented. Indeed, the paper is rather a validation, then a comparison, of two different approaches (published before) to estimate the distribution

of major phytoplankton groups: from space and based on a model output. The authors should add a reference to this comparison / errors estimation in the title. The more interesting part is about the seasonality analysis, and this should also be clear in the title.

This paper is in fact both a validation and a comparison. To make this clear we have changed the title to "Comparison and validation of an empirical algorithm and a numerical model in assessing global and seasonal phytoplankton community distributions".

The first part present a quick validation of the satellite and model approaches based on in-situ observations. The second part present a comparison (distribution and seasonality) between the model and the satellite approaches, without a clear link with the first part. The aim of this paper could be potentially interesting if clear information about which method should be preferred in specific regions or for specific groups are added. It's not the case in this first version, sometimes rather confusing. I suggest that the authors change the organisation of their paper, by adding more links with the in-situ validation part (only in the 2.1 part for the moment) when they discuss about the comparison between satellite and model observations/output, ie in each of the 3.2.X part and in the discussion part.

We have now added links between the distribution and seasonality from the two approaches and the validation from in situ data in the section 3.2.X:

- P9, L189-192; These high latitudes regions (Antarctic and North Pacific) also corresponded to the regions where diatoms were overestimated by 60% by the model. Antarctic also corresponded to the only region where diatoms from the satellite-derived approach were underestimated by >10%.

- P9,L200-201; Despite these differences, both approaches were within 13% of in situ data in the North Atlantic and Pacific and Antarctic.

- P9-10,L206-213; The validation with in situ data indicated that coccolithophores in these two regions were within 13% of the data set whereas chlorophytes in these two regions were underestimated by – 23% (North Atlantic) and – 57% (North Pacific).

Although the model agreed with the high abundance of coccolithophores and chlorophytes in the North Atlantic, these two groups were almost absent in the North Pacific. The validation with in situ data indicated that the model underestimated chlorophytes even more than the satellite-derived approach (– 73%) in the North Pacific.
- P10,L217-220; In the North Pacific, both approach were within 16% of in situ data (satellite-derived approach 3% closer to the in situ data set than the model) whereas in the North Indian, the coccolithophores from the model were closer (within - 5%) than those from the satellite-derived approach (17%).

Here are some examples of where we discuss the comparison between the two approaches and the in situ data in the discussion section:
-" P14, L311-312; The group for which both approaches agree most on the seasonal variation and are close to in situ data in most regions is diatoms."
-"P14,L320-322; This suggests that the parameters in the model may need to be adjusted, probably as a function of nutrient, since there is a relatively large overestimate of diatoms when compared to that of the in situ data in the North Pacific and Antarctic"
-"P15, L325-326; Both approaches indicate that the previously, well accepted and reported spring bloom in the North Atlantic and Pacific and austral summer bloom"
-"P15, L331-334 ; For example, Marañon et al. (2000) found that diatoms make up to 80% of the total phytoplankton carbon in the North Atlantic in May and was reduced in September-October which supports both the model and satellite-derived data"
-"P15, L343-346; The vast blooms of coccolithophores observed in the North Atlantic for example (e.g. Okada and McIntyre, 1979;Robertson et al., 1994;Boyd et al., 1997;Balestra et al., 2004) are well represented by both the model and the satellite-derived approach"
-"P16, L350-352. . . however, coccolithophores are abundant using the satellite-derived approach whereas they are absent in the model. The data on the existence of coccolithophores south of 50° S are contradictory. "
-"P16, L354-357;. . . Another region that indicates some deficiencies in the model is

the North Pacific where the model detects very low coccolithophore concentrations. Coccolithophores in the North Pacific can be found in substantial amount (Okada and Honjo, 1973;Lam et al., 2001;Crawford et al., 2003)"
-"P16, L367-369; Particularly striking is the underestimate by the model of chlorophytes in the North Pacific (by -72%) and the overestimate in the Equatorial Indian (by 75%). "
-" P17, L377-378, Another reason for the difference in chlorophytes with in situ data and between approaches could be related to the variety of phytoplankton included as chlorophytes"

Please find below some more precise remarks :
-It's not clear why the authors chose to use monthly climatology over 10 years to estimate the errors of their methods compared to in-situ dataset. If there is no clear reason to do that, the authors have to use matchup as the Seawifs documentation (from NASA) defined it : daily pixels observation (from the 9*9km box closer from the in situ measurements, at least) and, ideally within +-3 hours (but at least daily data for satellite part will be better..). Climatology use is not acceptable unless the authors have a good reason for this. The authors should also add the years 2008-2010 for SeaWiFS data ?
In the Seawifs documentation, the match-up data are total chlorophyll. Phytoplankton groups are very sparse. We only have 469 data points for which relative abundances are reported in the literature, from a pretty extensive review. Therefore we use monthly climatology to allow for statistical comparisons. We compare using Level-3 data, which are daily and not resolved at finer time increments. The reason why the period from 2008-2010 was not used is because from 2008 onwards there has been a considerable amount of data missing. More than 40% of the data for 2008 were missing due to sensor issues. Nearly as much (39%) was missing in 2009, and of course the sensor failed in late 2010. Additionally, the satellite drifted in its orbit, significantly changing the viewing and solar angles along with the areas observed. Here are a

few examples of the type of problems that were encountered from 2008 onwards. On January 4, 2008, the satellite was experiencing a telemetry anomaly. Operations continued on April 9, 2008 (http://oceancolor.gsfc.nasa.gov/forum/oceancolor/topic_show.pl?tid=2334). On July 2, 2008, the satellite was not in imaging mode (http://oceancolor.gsfc.nasa.gov/forum/oceancolor/topic_show.pl?tid=2607). It returned to routine operation on August 20, 2008. April 24 - June 16 2009 have no data (http://oceandata.sci.gsfc.nasa.gov/SeaWiFS/Mapped/Daily/4km/NDVI/2009/).

To clarify this we have added the following in the Material and Methods section (p6,L121-123):" The period from 1998 until 2007 was chosen because SeaWiFS started experiencing serious issues with data collection and orbit drift beginning in 2008."

- line p1084-26, the carbon fixation should be presented as an example and some references added.
We have now clarified this by adding: "For example, the intensity of carbon fixation and export is strongly dependent on the phytoplankton community composition (e.g. Falkowski et al. 2000; Laws et al. 2000, Armstrong et al. 2002).

- line p1085-0 to 24, please separate the satellite and the model part, it's rather confusing like that (for example, it's difficult to understand (line 16) if the approaches are those from satellite or from modeling.
We have now separated the satellite and model part. The text is now as follow:
"The approaches to characterize the phytoplankton community composition at a global scale can be roughly classified in two categories: modelling approaches and satellite-derived approaches. In the modelling approaches the phytoplankton composition is generally based on biogeochemical functions. Biogeochemical models coupled to physical circulation can describe the complex interactions between the physics and

biology in the oceans (e.g. Le Quere et al., 2005;Moore et al., 2004;Dunne et al., 2005;Doney and Ducklow, 2006;Gregg et al., 2003;Dutkiewicz et al., 2009). Data assimilation techniques can also be used to constrain the model to track observations time series and to optimize certain variables. Satellite approaches can either be spectral based (using bio-optical algorithm) or based on phytoplankton concentration. The spectral-based approach can be used to estimate particle size distribution (e.g. Ciotti et al., 2002;Mouw and Yoder, 2006), functional groups (e.g. Alvain et al., 2005) or other parameters of interest. The second satellite approach relies on phytoplankton concentration (expressed as chlorophyll concentration or absorption coefficient, e.g. Uitz et al., 2006;Aiken et al., 2007) as an indicator for phytoplankton community composition."

- line p1086- 0 to 25 : clear lack of references in this part...! ie for page 1087. Please check your references for each previous results or an hypothesis.
We have added references as following: "Seasonal cycles have different forcing depending on the region. For example, the spring-summer blooms in subpolar waters are related to winter mixing that replenishes surface waters with nutrients (Sverdrup 1953, Pingree et al. 1977). Winter mixing is followed by spring-summer increases of incident solar irradiance and water column stratification leading to a well lit, initially nutrient-rich mixed layer conducive to phytoplankton growth and biomass increase. Winter blooms in subtropical waters are also responses to winter mixing (Longhurst 1995), although mixing in the subtropics is generally weaker than that which occurs farther poleward. In the subtropics, comparatively high winter solar irradiance at the lower latitudes and shallow mixed layers leads to an immediate phytoplankton response (winter blooms) to nutrients."

- The authors should address and discuss the question of satellite observation bias due to underestimation or overestimation of chlorophyll-a concentration in

some regions (see Arrigo et al. 1998 or Dierssen et al 2000, for the Antarctic for example). This could also be an explanation of some errors they noticed.
The reviewer is right that persistent bias in the total chlorophyll could introduce bias in the phytoplankton composition. We have therefore added the following in the discussion (p.13, L298-300):
"Furthermore, the total chlorophyll a concentration in some regions such as the Antarctic is known to be biased (e.g. Arrigo et al. 1998, Dierssen et al. 2000) and may therefore bias the phytoplankton community composition derived from the satellite approach."

-page 1088, line 5 : is it the NOMAD dataset ? If not, the author should use this dataset also.
No this is not the NOMAD dataset. The dataset we use can be downloaded at http://gmao.gsfc.nasa.gov/research/oceanbiology/data.php (as in Material and Methods, pp7, L142).
The NOMAD dataset does not contain phytoplankton composition data. It does contain pigment concentration. Since there are a variety of methods and a lack of consensus on the conversion of pigment concentration to phytoplankton composition data, we only use published data of phytoplankton composition.

This part is really too short :
-The authors have to explain the method they used to define which phytoplankton is present in the water based on their in-situ observations.
We have now clarified this in the text in Material and Methods:
" The phytoplankton groups from the model and satellite-derived approach are validated against in situ data. This data set includes 469 surface-layer observations of phytoplankton group concentrations (Figure 1). The full list of the referenced paper used to build this database can be found in Gregg and Casey (2007) and downloaded

directly from http://gmao.gsfc.nasa.gov. Data are converted when necessary into percent abundance of the entire population to compare with the model. The co-located, coincident match ups are assembled over ocean basins and over the months for a year. "

- The impact of the use of a climatology for the seasonal study should be at least precisely discussed.
Although the use of a climatology here is not ideal, the comparison of the phytoplankton composition from both approaches with this independent in situ data set allows us to quantitatively assess both approaches. This has now been added to the discussion. (p13, L285-287):
"Although the use of a climatology here is not ideal, the comparison of the phytoplankton composition from both approaches with this independent in situ data set allows us to quantitatively assess both approaches."

- A table for validation results could really help the reader.
We are confused with this comment. Table 1 in the manuscript presents the percent difference in relative abundance between the model or the satellite-derived approach and in situ data. The intention of making a table was to help the reader interpret the results.

- Parts 3.2.1, 3.2.2... please add some reference to the "in-situ based" errors estimation part to make these parts useful. Which method is the best... ? Presented like this, this paper can only be used by people in charge of the development of the satellite and model approaches. With more information it could be useful for the community also.
As indicated earlier, we have now added several references to how the distribution from the two approaches compared with the validation:

- P9, L189-192; These high latitudes regions (Antarctic and North Pacific) also corresponded to the regions where diatoms were overestimated by 60% by the model. Antarctic also corresponded to the only region where diatoms from the satellite-derived approach were underestimated by >10%.
- P9,L200-201; Despite these differences, both approaches were within 13% of in situ data in the North Atlantic and Pacific and Antarctic.
- P9-10,L206-213; The validation with in situ data indicated that coccolithophores in these two regions were within 13% of the data set whereas chlorophytes in these two regions were underestimated by – 23% (North Atlantic) and – 57% (North Pacific). Although the model agreed with the high abundance of coccolithophores and chlorophytes in the North Atlantic, these two groups were almost absent in the North Pacific. The validation with in situ data indicated that the model underestimated chlorophytes even more than the satellite-derived approach (– 73%) in the North Pacific.
- P10,L217-220; In the North Pacific, both approach were within 16% of in situ data (satellite-derived approach 3% closer to the in situ data set than the model) whereas in the North Indian, the coccolithophores from the model were closer (within - 5%) than those from the satellite-derived approach (17%).

- p1090 : could you precise the method used to compute the p value ?
We have added the reference for the method we used to calculate the p-value in the material and methods section: "The p-values are computed as in Gibbons (1985)."

- The seasonality part analysis is better even if it looks like a succession of results.
Thank you

- The discussion should address separately the potential bias due to model or satellite method, then specficity of regions. Once again, a discussion based

on a careful validation exercise is needed, unless the authors focused their article only on a comparison between a specific satellite approach and a specific model, but this will be really less interesting for the community.
Here are some examples of where we discuss the comparison between the two approaches and the in situ data in the discussion section:
-" P14, L311-312; The group for which both approaches agree most on the seasonal variation and are close to in situ data in most regions is diatoms."
-"P14,L320-322; This suggests that the parameters in the model may need to be adjusted, probably as a function of nutrient, since there is a relatively large overestimate of diatoms when compared to that of the in situ data in the North Pacific and Antarctic"
-"P15, L325-326; Both approaches indicate that the previously, well accepted and reported spring bloom in the North Atlantic and Pacific and austral summer bloom"
-"P15, L331-334 ; For example, Marañon et al. (2000) found that diatoms make up to 80% of the total phytoplankton carbon in the North Atlantic in May and was reduced in September-October which supports both the model and satellite-derived data"
-"P15, L343-346; The vast blooms of coccolithophores observed in the North Atlantic for example (e.g. Okada and McIntyre, 1979;Robertson et al., 1994;Boyd et al., 1997;Balestra et al., 2004) are well represented by both the model and the satellite-derived approach"
-"P16, L350-352. . . however, coccolithophores are abundant using the satellite-derived approach whereas they are absent in the model. The data on the existence of coccolithophores south of 50° S are contradictory. "
-"P16, L354-357;. . . Another region that indicates some deficiencies in the model is the North Pacific where the model detects very low coccolithophore concentrations. Coccolithophores in the North Pacific can be found in substantial amount (Okada and Honjo, 1973;Lam et al., 2001;Crawford et al., 2003)"
-"P16, L367-369; Particularly striking is the underestimate by the model of chlorophytes in the North Pacific (by -72%) and the overestimate in the Equatorial Indian (by 75%). "

-" P17, L377-378, Another reason for the difference in chlorophytes with in situ data and between approaches could be related to the variety of phytoplankton included as chlorophytes"

- P1095 : this part is not clear and rather vague. It's difficult to extract a clear information from the discussion part as it is written in this first version. - The 'phaeocystis part' (in the conclusion part) is a hypothesis and should be presented like that, at this stage, it's not clear that it will improve this inter-comparison exercise. - The two last sentences are off topic if we consider previous results addressed in this paper.
We agree that we are unsure that the addition of a Phaeocystis group in the model would lead to improvement. To clarify this we have changed the wording to reflect this in the text: "The departure from in situ data and the disagreement between both approaches for chlorophytes may be due to the inclusion of Phaeocystis spp. as chlorophytes in the model and observations and the lack of understanding on the specifics of how they relate to other transitional groups. We hypothesize that adding a group in the model that would represent Phaeocystis sp. may improve these estimates."

The last two sentences are:" Future satellite ocean color missions may enhance our capacity to distinguish phytoplankton composition at a global scale by increasing the number of wavelengths to allow the differentiation of an increasing number of phytoplankton groups. This is especially true for the intermediate, transitional groups between the functional extremes (diatoms and cyanobacteria), such as the chlorophytes and non-coccolithophore prymnesiophytes, which are the most problematic for both approaches."Because upcoming missions like PACE are in the process of determining the type of direction the mission should take, we added these sentences to show the importance of this study in the broader context of satellite research. The results of this study provide new information on the strengths and weaknesses of

satellite-derived data and may help in making decision about the characteristics that future sensors/missions may have. The last sentence reminds the community that in situ data are crucial and that our understanding of some of these phytoplankton groups is still rather limited.

- Table and figures are clear.
With many thanks in advance to take into consideration this review,
Sincerely yours

Interactive comment on Biogeosciences Discuss., 10, 1083, 2013.