

Earth Syst. Sci. Data Discuss., author comment AC3
<https://doi.org/10.5194/essd-2022-159-AC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

André Valente et al.

Author comment on "A compilation of global bio-optical in situ data for ocean colour satellite applications – version three" by André Valente et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-159-AC3>, 2022

Thank you very much for the comments. Please find below the responses to each of your comments.

Comment 1

In my opinion the paper is very well crafted and of high very quality, as it is the dataset.

The dataset is a unique and of high interest for the ocean colour community.

Response

Thank you.

Comment 2

However,

I do have a concern,

This paper is regarding a 3rd version of the dataset, meaning that there are already 2 other papers published previously regarding the this dataset. Hence, the dataset is not new, but updated.

As such, this paper shows that the number of observations for the recent years as increased significantly when comparing with the previous version(V2).

This paper, however, is basically a copy from the V2 paper (Valente et al. 2019) with some additional paragraphs to let the reader know how much new data there is. Even the figures, 16 of them, are the same as in the previous paper (except figure 1). Which might be the intentional, as it is indeed an update on the description of a previous version of the same dataset.

I don't know the policy of the journal regarding the publication of papers that describe/discuss the updated of a dataset.

In my view, it would be more interesting, and beneficial to the reader (and public in

general), to present and discuss how the new update changed the past version(s) of the dataset. Whereas there would be no need to maintain the same text/discussions/figures from previous papers. As this knowledge is already published, thus, it already has a doi to be referred to. Unless, of course these results/discussions changed due to the new update.

Response

Regarding this main concern of using a similar structure (text/figures) of previous versions, we followed the policy of ESSD, and specifically we followed the ESSD paper "Global Carbon Budget", which is updated every year (<https://doi.org/10.5194/essd-12-3269-2020>). In each update of that paper the text and figures remain essentially the same. This was considered to be a valid approach for an updated dataset - the updated dataset accumulates all previous co-authors and data, while maintaining its structure (text/figures) and describing the new data, and becomes the new dataset to be cited (i.e., only the last version of the dataset needs to be cited). In our specific case, the updated version 3 does not add new data sources - it updates/adds data from the previous 27 data sources. Thus the methods, and the data sources, need to be essentially the same. It should also be noted that the present version 3, although similar, is not a copy of v2 as there has been numerous changes across the text to improve the paper and to highlight the new data and changes regarding previous version (see track-changes in <https://essd.copernicus.org/preprints/essd-2022-159/essd-2022-159-supplement.pdf>). It is also noted that an additional effort to further reduce the similarity of v3 compared to v2 has already been carried out during the process of submitting to ESSD.

Comment 3

Further general comments below:

- . it would be useful to know up front, i.e. in the introduction, how much data from which project was added in this new version of the dataset.

Response

An upfront description of how much data from which project was added in the present version, will be added to the revised version of the manuscript. This will be added to the third paragraph of section 2.1 (where the updated data is now currently stated) and not in the introduction since the compiled variables and data sources should to be firstly introduced (currently in the first paragraph of section 2.1).

Comment 4

- . the new dataset is improved by
 - . v3 uses AERONET-OC v3
 - . more observations
- . the description of a few datasets do not add anything from the previous dataset paper, for e.g., ARCSSPP does not add anything to a new version as it only provides data from 1954 to 2006
- . the same as above for:
 - . GeP&CO it ran from 1999 to 2002

- . BARENTSSEA 1997 to 2013
- . BIOCHEM 1997 to 2014
- . ESTOC from 1994 to 2011

Response

As noted in the response to the main concern (comment 2), the followed strategy is to present all datasets, including those that were not updated, keeping a similar structure of the paper.

Comment 5:

. Figure 1 doesn't seem to cover the beginning of the data range wavelengths, i.e., 313 nm, it seems to start between 360 and 340 nm

Response

This will be corrected in the revised version of the manuscript.

Comment 6

. ~15% increase of data from Valente et al 2019, from previous existing stations. This information should be in the introduction, so we know up front what the update is.

Response

We will add this information to the third paragraph of section 2.1 (please see also response to comment 3).

Comment 7

. Results section (copy from the v2 paper + some paragraphs stating that there is an increase of obs. in the current version)

As it can be seen in the

4th paragraph - some general results on chl_a (as in V2 paper)

+ a sentence on increasing n obs between V2 and V3 (+5% fluor, +16% hplc)

Response

As noted in the response to the main concern (comment 2), the followed strategy is to present all datasets, including those that were not updated, keeping a similar structure of the paper.

Comment 8

. question: (not that it is significant, but) why are the limits for both types of chl methods different? I.e.,

$0.001 < chl_{a_fluor} < 100$

$0.002 < chl_{a_hplc} < 99.8$

Response

The limits are different because the lowest and highest values of the compiled chl_a_fluor and chl_a_hplc are different (these two variables are measured with different protocols thus they do not need to have the same range).

Comment 9

6th paragraph - relationship between rrs ratio with chl_a

Id it change between different versions? If so, any mechanism that could explain it? (for e.g., predominance of El Nino/La Nina phases during the new coverage?)

. it would be interesting to see if some descriptive relationship statistics evolution between versions, for e.g. coefficient of correlation, between the nasa algorithms and the rrs/chl_a ratio found in the V1, V2 and V3

Response

Compared to the previous version, only a few concurrent Rrs_ratio VS Chl data points were added (N=13), which corresponds to a quite small 0.13 % increase. As already stated in the text: "Compared to the previous version (Valente et al., 2019), the relations between maximum band ratio and chlorophyll are not altered by the additional number of concurrent observations (N=13)."

Comment 8

7th paragraph - general results of aph, adg, bbp

+ only new data for aph (+30%)

8th paragraph - Kd didn't change (as in V2 paper, no need to repeat, already discussed)

9th paragraph - as in V2 paper

F1 is new from V2 paper

F2 to F 16 are as in V2 paper (with some, F6, F12, F13, F16 showing statistic values slightly changed)

Response

As noted in the response to the main concern (comment 2), the followed strategy is to present all datasets, including those that were not updated, keeping a similar structure of the paper.