

Earth Syst. Sci. Data Discuss., referee comment RC3
<https://doi.org/10.5194/essd-2021-236-RC3>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on **essd-2021-236**

Murat Aydin (Referee)

Referee comment on "Third revision of the global surface seawater dimethyl sulfide climatology (DMS-Rev3)" by Shrivardhan Hulswar et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-236-RC3>, 2021

This is the third revision of global surface seawater DMS climatology. DMS is an important natural source of atmospheric sulfur. The new climatology includes a substantial amount of new data and represents a significant upgrade over the previous version; it is likely to be widely used in various modeling applications. Overall, it is a reasonably well-written manuscript with sound methodology, although some aspects of the data processing were difficult to understand. My concerns, questions, and suggestions are explained in detail below, followed by more specific comments, most of which are fairly trivial.

One overarching concern I have is with regards to the decisions made during the initial stages of the data processing. I understand that subjective decisions are unavoidable when it comes to data filtering, but proper justification is important and the impact of these decisions should be quantified if possible.

First data "clean-up" decision is the omission of the near zero and negative measurements. I find it difficult to believe that the negative values are incorrectly reported numbers (Ln 182). Negative measurements, unless they are extreme negative numbers, commonly occur when working at or below detection limits due to uncertainties in the calibration and/or a system blank that is being subtracted from the signal. I would expect there to be roughly an equal number of small positive measurements that balance out the negatives so that the average is equal to zero within the uncertainties. I do not see a reason to throw out these data as doing so will introduce a positive bias to the data set, especially if data filling schemes are implemented after data removal. In the current case, I realize that a very small fraction of the data is thrown out and the impact of this data clean-up is likely insignificantly small. Still, I do not find the justification convincing. Do the negative numbers cause problems with the data processing that follows? A more elegant solution might be to replace all data below detection limit with zeros, or a very small positive number if zeros cause problems too.

The second item is filtering out high numbers. I was glad to see they did not go with the

L11 method of using a 99.9% filter on the entire data set. I think the biggest issue with that approach is it will filter out much more than 0.1% of the data from provinces with high DMS concentrations. The question I have is, is there really a need to filter out any data at all? If you think that the highest 0.1% from each province do not represent surface ocean concentrations, it is necessary to filter out these measurements, I agree. However, if you think these high end measurements might represent blooms (Ln 186), I don't think it is justified to throw these data out, even if there are repeat samplings of the same bloom. My guess is blooms might be severely underrepresented in these data sets because I suspect they occupy a small fraction of the surface ocean, or is there evidence to the contrary?

The next critical data handling decision is how to combine the newer high resolution mass spectrometer (HRMS) data sets with the older low temporal resolution GC data. Rather, the 1 hr averaging applied to the HRMS data before data unification. This decision should be primarily data driven and the related discussions should include two important criteria: 1) The fidelity of the HRMS data to real variability in the surface ocean. How much temporal averaging to HRMS data is needed to filter out the noise from issues related to these instruments and sampling methods to capture the variability in the surface ocean? 2) The relative representativeness of HRMS vs. GC data. Are the GC data representative of a larger volume of seawater averaged over a longer period than the HRMS? The paper does not discuss these points. Instead, there is an emphasis on the need to balance the influence of historical vs. modern data on the climatology. What is the reason to worry about this balance? This is not adequately explained. It is possible that the quantitative impact of this 1 hr averaging is not all that large, but it would be good to know because HRMS data will become even more dominant in the future.

Is it possible to demonstrate how these decision impacts the final climatology by repeating what was done for testing the sensitivity of results to using dynamic vs. static boundaries for provinces and VLS?

I didn't understand parts of section 2.5. It is not clear how the 1degx1deg product was created and how it was merged with the province averages-based product? Are you doing a 1x1 bin averaging of the underlying raw data? It is stated that the province-based averages and the binned data are superimposed to merge them. I suspect some kind of weighted averaging is done; this needs to be explained. In the Barnes filter step, the ROI is not explained (in one sentence the average distance between a grid point and the data points that influence it can be mentioned), but the sensitivity of results to this parameter is discussed using patchiness as a measure of accuracy. This is problematic because the climatology itself should be the best representation of reality, yet this makes it sound like you already know what reality should look like. VLS, which is actually based on data although it is not clear from which reference(s) the underlying VLS data are taken, is used as ROI in the Barnes filter. I'm not an expert in this kind of thing, but this made sense to me. However, I struggled to reconcile short VLS values with the starting point of this climatology, which is the notion that DMS concentration should more or less be the same within the same biogeochemical provinces. Perhaps a sentence or two are needed to bring everything together here.

A final suggestion: It would be useful if the new DMS flux estimates are used to calculate

DMS emissions. The emissions could be presented in tabular form for each province and for 30deg latitude bands, and also as a graph vs. latitude, which could include the L11-based emissions for comparison. This would be helpful for atmospheric scientists studying atmospheric oxidation products of DMS.

The specific comments below are listed in the order they appear in the paper, not in the order of importance.

Specific comments:

Ln 31: Missing full stop.

Ln 91-93: I don't understand the sentence. Delete "incorporating" perhaps?

Ln 94: No need for comma.

Ln 109: Which database?

LN 139-141: It is stated here that data from highly productive regions lead to patchiness in L11. These sentences, and some others later, seem to imply DMS emissions from blooms should be ignored because the climatology should not look patchy. I agree that patchiness is not good when introduced during data analysis, for example due to introduction of provinces and related assumptions to come up with first guess estimates of DMS concentrations. However, if there is patchiness inherent to the underlying data, don't we need to accept that as reality?

Ln 178: It is not clear what is meant by data should be within a range of 25%.

Ln 195: I think the 0.001% is a typo. Should be 0.1% for 99.9% threshold.

Ln 209-210: Should be less than or equal to 24 hours?

Ln 211: Fig. 2a does not show any data collected less frequently than 24 hours. If lower frequencies are bundled into the 24 hour bin, this should be noted in the caption.

Ln 214: Fig. 2b shows high-res campaigns are 15% of the total.

Ln 215: No need for "which is based on observations."

Ln 218-219: I don't understand what is meant by "matching?" Is 1 hour chosen to have roughly equal contribution from CIMS/MIMS vs. GC data?

Ln 235: "Computing a mean value across the provinces" is confusing to me, I think you mean calculating a mean value that is representative of the entire province?

Ln 280: I think this is the first instance a province acronym being used. It would be easier to read the manuscript if the number associated with the provinces accompanied the acronyms every time an acronym is used because the maps only show the numbers.

Ln 300-304: I suggest seasonal cycle or month-to-month variability instead of annual trend.

Ln 350: Is there really data from under the sea-ice, or is this a result of calculating average values for provinces?

Ln 373-375: Would the use of wind data from more recent periods change the fluxes appreciably? My guess is no. A one-to-one comparison to L11 argument is not that convincing because many aspects of the methods have been changed.

Ln 418-419: Do you mean southwest coast of Africa and the west coast of South America?

Ln 441: You switched what number is reported in the parenthesis, I think.

Ln 583: Fig. S11b color bar label is not in percentage.

Ln 640: Should be Fig. S13.

Ln 646-648: Fig. S14 shows L11 annual average is slightly less than Rev3. Obviously, this depends on what is subtracted from what to make those figures. Caption should include this information.