Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-72-RC3, 2020

Interactive comment on "Global ocean dimethyl sulfide climatology estimated from observations and an artificial neural network" *by* Wei-Lei Wang et al.

Anonymous Referee #3

Received and published: 23 April 2020

Review of Wang et al., Global ocean dimethyl sulfide climatology estimated from observations and an artificial neural network

This manuscript describes a novel methodology for deriving a global ocean dimethyl sulfide (DMS) climatology, using an artificial neural network (ANN). The authors demonstrate that the ANN is able to explain a greater fraction of variance in the raw available observations of surface ocean DMS concentrations, as compared with a multiple linear regression approach. They also contrast this approach with previous work that used spatial and temporal gap-filling to estimate DMS concentrations, including in data-sparse regions. Instead, the approach presented here derives relationships be-

tween observed environmental parameters and observed oceanic DMS DMS concentrations (using the multiple regression or ANN), and uses these to predict/extrapolate DMS concentrations globally.

The paper is clearly written, the methods are straightforward and appropriate, and it represents a valuable contribution to work on understanding and representing the present-day climatological distribution of DMS concentrations in the surface ocean. Improved climatologies of DMS would be useful for Earth System models, especially if they can offer more insights into how the DMS production would change under past/future climate states. It's unclear (to me, at least) whether a machine learning approach will be able to offer such physical insights. Nevertheless, such approaches can offer a better estimate of the present-day state, and this is useful in itself for Earth System modeling. The uncertainty in ocean DMS climatologies is still quite large, despite advances during the past decade, and new advances in statistical approaches that can reduce errors in these datasets are welcome.

I have only a few minor comments, as follows:

I agree with the comments of the two previous reviewers that the arguments made against data binning are weak. The authors imply that it is an inherently inferior approach, but, this is not necessarily true a prior. There can be good arguments in favor of data binning before analysis, e.g., to harmonize the temporal and spatial scales of multiple datasets before analyzing the relationships between them. When in situ DMS measurements (essentially instantaneous) are being predicted via monthly mean values of chl-a, MLD, etc., it is not at all obvious that it is appropriate to perform the analysis without first binning the data. This point should be treated with more nuance, taking into account the details of the datasets and the processes involved.

p. 5, l. 128-130: I was glad to see that the authors have considered the issue of potential overfishing, but they don't explain how they determined that the setup they used for the ANN is not overfitting (i.e., what methods or criteria were used to determine this). It's common to use multiple rounds of cross-validation (such as k-fold cross-validation or related methods) in order to determine whether a statistical model may be overfitting and to assess the uncertainty in the fit. If I am understanding the description of the method correctly, it seems that while the authors divided the data into training and validation subsets, they did so only once. In this case, the results of the ANN will be sensitive to the specific subset of data that was used for training it. It should be explained how the training/validation subsets were selected, and also whether a multi-round cross-validation method was employed (and if not, why not). Or, if appropriate, the authors could simply carry out a more thorough cross-validation and update the manuscript, since I expect this should not require much effort.

p. 5, l. 133-134: It was not obvious to me what the "random states" refer to – is this a random seed controlling initial parameter values?

p. 8, l. 220: here, it is stated that ANN is able to "capture more of the variance" than "previous extrapolations (Kettle et al., 1999; Lana et al., 2011)". This is a key claim of the paper in terms of the claimed improvement over previous methods, and I can believe this is probably true, but I think the claim ought to be supported by a quantitative value – i.e., the percentage of variance captured by the two previous climatologies – so that readers can compare and see the improvement in this metric. Perhaps these values are in the manuscript somewhere and I overlooked them – in that case I think they should be featured somewhere that is easier to find (e.g., in the abstract or in a table).

p. 11: I tested the links for the code and data availability; the data doi link at zenodo works, but the github link does not seem to be available.

I also noticed a couple of typos: p. 2, l. 40: "result" -> "results" or "result[s]" p. 5, l. 31: "deduction" -> "reduction" p. 5, l. 133: "assemble" -> "ensemble" (?) p. 7, l. 189: "wasters" -> "waters"

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-72, 2020.