Second review round of Wang et al. Manuscript submitted to Biogeosciences Discussions

by Martí Galí

General comments

The manuscript has seen some substantial improvements compared to the first version. I appreciated the more comprehensive comparison between DMS fields generated by the ANN and those generated with simpler empirical algorithms, including remote sensing algorithms. The comparison is informative as to the uncertainties in different global DMS products. However, I was disappointed to find that several inconsistencies and inaccuracies remain in the text. These issues have to be addressed and clarified.

In the previous review I prompted the authors to provide a more up-to-date view of DMS(P) cycling processes. However, this did not require adding long discussions, just improving the quality of the information and the accuracy of literature references. Now, in section 3.3, the authors discuss why some environmental variables got selected as ANN predictors in relation to available knowledge. But they do not show any results about the sign of the effects of each predictor variable used in the ANN, or whether the sign changes from one region to another. Thus, the discussion becomes quite pointless (because our understanding does not increase), and could be more succinct. The distinction between "controlling factors" and "predictors" should also be clearer throughout, as this paper does not shed light on the controlling factors.

As in the previous version of the manuscript, the authors make a strong case against the binning of in situ DMS measurements, and present a new figure (Fig. 2) to support their choice. However, I still find that the "binning issue" is treated in an inconsistent manner. For example, despite advocating the importance of using non-binned data, the authors do not use time-resolved satellite Chl as predictor although such data are available for most of the DMS measurements used to train the ANN. Instead, the authors match non-binned DMS data to a heavily time-binned global climatology of satellite Chl from SeaWiFS, which does not even cover the entire ocean color satellite era. Finally, they produce global 1x1 monthly DMS climatologies by applying the ANN, previously trained on non-binned DMS data, to climatological fields of the predictor variables. To be clear, I am not asking the authors to perform strict matchups between DMS and satellite data, which I think is beyond the scope of their study (note, however, that a DMS-satellite match-up database is available for measurements older than 2012: https://doi.org/10.5281/zenodo.2205131). In my view the paper needs clearer and more consistent reasoning and methodology regarding the utility of binned data to produce climatologies, because (1) the whole point of making climatologies is precisely collapsing temporal and spatial variability, and (2) the trained ANN is not implemented to produce non-climatological fields in this study.

In connection with comment above: no quantitative assessment of overfitting is provided in section 3.4 nor figure 2. Just a qualitative assessment based on "how much the training and validation errors overlap". Although the inclusion of figure 2 is welcome, a quantitative assessment would provide stronger support to the authors' choices (ie training the ANN using non-binned data).

Specific comments

Line numbers refer to the document bg-2020-72-AC4-supplement.pdf

Abstract

The reasoning in the second sentence is backwards, in my opinion. In addition, I think this paper deals with the "predictors", not with the "controlling factors".

Introduction

L45: Some phytoplankton do not produce significant amounts of DMS, just its precursor DMSP. Please clarify.

L44-51: This entire paragraph is misleading. Reading it, one may deduce that the inability of biogeochemical models to capture global DMS distributions stems from the excessive simplification of phytoplankton diversity, lumping together very different taxa in a few phytoplankton functional types. Le Clainche et al. (2010) propose otherwise: it is the excessive dependence of DMS on phytoplankton dynamics in sulfur cycling parameterizations what results in poor model skill. And they explicitly suggest that modulation of community DMS production yields by environmental stressors (light, nutrients) must be included in model parameterizations. For example, by Vallina et al. (2008) showed the importance of adding light stress effects on DMS production in the Sargasso Sea using a model that had a single phytoplankton compartment. I refer the authors to my previous review for papers reporting experimental evidence of nutrient and light stress. Finally, note that citation of Le Clainche's paper does not support the statements made in this paragraph.

L57: "Interpolation from neighboring provinces". I suggest adding "weighted" before interpolation. Actually, what Lana et al. (2011) did in provinces with insufficient monthly data substituting their seasonal cycle by that of the "biogeochemically closest" province, weighted by the "local province" average.

L59-69: The message of this paragraph is unclear and the newly added sentences disrupt the flow. In addition, I think the authors should refer to "predictors" instead of "controlling factors". The ANN approach as implemented here does not reveal the controlling factors.

Methods

L84-85: With the current writing, it is unclear whether measurements were matched to the 1997-2010 SeaWiFS climatology or to the multiyear time series. In the latter case, the period 2011-present would not be covered by matching satellite data. In the former case, one may wonder if the 1997-2010 climatology is representative of the 2011-present period in all ocean areas. Please clarify and provide a solid argumentation for not matching DMS data to satellite data at the best available resolution.
Note: I found the answer later, in L232. As suspected, DMS data were matched to climatological Chl. Note also that global Chl climatologies from ultiple sensors spanning 1997-2020 are available (e.g. GlobColour, ESA CCI).

L84: What reprocessing? Access date? Same applies to L94.

L114: This contradicts L91, where the authors say that "For consistency, we use only Chl-adata retrieved from SeaWiFS in the following multilinear and network models".
L117: The statement that "there is no available climatological [DMSP] dataset to fill the missing values" is not entirely true. There is no climatological DMSP dataset based on objective interpolation of in situ data. But there is a global sea-surface DMSPt climatology based on the remote sensing algorithm of Galí et al. 2015, available here: https://doi.org/10.5281/zenodo.2558511.

L118 and elsewhere: Please replace SiO by SiO_4.

L154-162: It is unclear what the authors did. I could not understand whether data had finally been shuffled or not, and why.

Section 2.4.2: I don't see the point of training the ANN using non-binned data to, afterwards, compute the DMS climatology from 1x1 gridded climatologies of the predictor data. If training the model on non-binned data was so advantageous, I would expect the authors to first compute global gridded DMS time series at the highest possible resolution, and only afterwards collapse the multiyear fields into a climatology. The advantage of training the ANN with non-binned DMS data would be demonstrated if a climatology produced using the minimal possible amount of climatological predictor fields outperformed another one computed directly from climatological predictor fields.

Equation 7: this equation produces negative DMS flux at WS < 1.33. Although such low wind speeds are infrequent in a climatology, did the authors add a correction to avoid potential negative values?

Results and Discussion

L224: I suggest replacing "it is relatively easy to parameterize in a biogeochemical model" by something like "it is a priori more amenable to simple parameterizations".

L248: "the model roughly predicts the level of DMS concentrations" sounds a bit vague. Can you please report skill metrics like RMSE or bias in addition to R2?

L257: [the ANN] "also incorporates diurnal and seasonal signals present in the data". Please add something like "but see below", or place this sentence after discussing why time of day does not improve ANN predictions. In fact this sentence is a bit contradictory with the finding described two paragraphs later.

L264: "As shown in Fig. 2… sampling location and date alone can explain 44% of the validation data variance". Figure 2 does not show R2. Please rephrase.

L265: The discussion that starts here is interesting but needs a little revision. The analysis of Vallina & Simó 2007 did not address diel cycles, so I do not agree with this sentence: "Given the strong correlation between solar radiation and265DMS concentration reported by Vallina and Simó (2007), one would expect that adding sampling time would improve the model performance".
The occurrence of predictable diel cycles was assessed by Royer et al. 2015 (Small−scale variability patterns of DMS and phytoplankton in surface waters of the tropical and subtropical Atlantic, Indian, and Pacific Oceans) using continuous underway DMS data collected across the global oligotrophic oceans. They concluded there was no such a universal diel cycle, for the reasons pointed out in the manuscript (different possible outcomes of the balance between DMS sources and sinks over the die cycle).

L281: I do not think the cryoprotectant role has anything to do with global DMS patterns, considering that water temperature in most oceanic regions never decreases to freezing temperatures. Please remove or elaborate a better explanation. In addition, what is the sign of the SST influence on DMS?
On the contrary, I agree with the sentence that concludes the paragraph. SST and MLD have known for decades to define and capture a large deal of the biogeographic patterns of the ocean (e.g., Fay & McKinley, 2014, ESSD).

L290: "rate" should be "rate constant" (= specific rate). Not the same as rate!

L295: "loss" should be "transport" because sometimes turbulent vertical transport can result in net inputs to the upper mixed layer.

L351: Please remove “runoff”. I would use expressions like “higher freshwater content” or “salinity stratification”. Freshwater may come from ice melt, continental runoff, etc… And in addition much of the stratification in high northern latitudes is due to encounter between fresher Pacific-derived waters and more salty Atlantic waters.

L361: It is risky to use these maxima to illustrate your point. If anything, I would rather use quantiles. Very high Chl from SeaWiFS may result from algorithm artifacts in CDOM-ladden waters, whereas Chl data from PMEL were not quality-controlled and some were collected in estuarine areas (Galí et al., 2015, RSE).

L371-374: In the marginal ice zone, DMS can reach high concentrations during-after ice breakup without any need for sea-ice-algae release. Simply because the biomass of high-DMS(P) phytoplankton (*Phaeocystis*) in the water column can be very high. See the compilation by Levasseur (2013; Nat Geo). Out of the seasonal ice zone, the argument on thhe cryoprotection function does not make sense, since subpolar waters where coccolithophores bloom in booth hemispheres do not reach freezing temperatures, and for example Emiliania huxleyi blooms typically happen late in summer at tempertaures >10C. Please rephrase.

L396: In the text you report a global DMS emission of 15.9 Tg S yr^{-1} using the GM12 parameterization. However, when I look at Fig. 8, I see a mean monthly DMS emission of about 1.5 Tg S yr^{-1} (and, for sure, greater than 1.4). Multiplying 1.5x12, we get an annual emission of 18 Tg S yr^{-1}. Therefore, I wonder how do you arrive at the value of 15.9 Tg S yr^{-1}.

L430-434: Agree with this point.

<u>Technical corrections</u>

L84: Check verb tense in this sentence and concordance with the previous one.
L94: in the paragraph above the resolution was 9.2 km, now it is 93 km. Please check.
L173: coordination or coordinate?
L348: please add “concentrations” after “higher”.
L426: principal, not “principle”.