Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-72-RC2, 2020

Biogeosciences
Discussions
Open Access

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

Anonymous Referee #2

Received and published: 6 April 2020

The manuscript proposes a new global ocean DMS climatology, or a method to construct it, based on an Artificial Neural Network (ANN). This methodology uses a number of variables and their intelligent combinations as predictors of DMS concentration distribution. It is meant to overcome the limitations of objective analysis based on inter- and extrapolations as well as the limitations of simple linear or logarithmic regressions with few predictors, and to provide better fits of predictions to observations. While developing their ANN application and to claim its better performance, the authors conduct parallel applications of previously published models. Eventually, they indeed obtain a better fit, but very similar seasonal and geographic distributions. The global annual emission to the atmosphere is revised towards the lower end of the hitherto most ac-

cepted estimate.

The topic is timely since, after years of having DMS been dismissed for its role in new particle formation, recent studies are recognizing it again as a central agent in ocean-atmosphere-climate interactions. Atmospheric chemistry and climate models require updated climatologies of DMS emissions.

The text is generally well written and the diplay items are clear and informative, with one exception (see particulars below).

That said, the manuscript reads as though it was written 10 years ago. Even though the ANN methodology is probably state-of-the-art (I am not an expert and can hardly assess every technical aspect), the interpretation arguments are outdated, ignoring many of the discoveries in the last decade. This adds to some bad referencing. But most importantly, when the authors intend to make relevant comparisons with previous similar efforts, they miss the point of the studies they are comparing to, or use them in the wrong way. Finally, besides presenting their new method, they fail to discuss what is new in their findings, they just repeat what is already well known and with much poorer arguments, rather than stressing what is unveiled and why. I will develop these and other concerns hereafter, as they come up in the order of the manuscript.

L28-30: "The weak relationship may be caused by the so-called "summer DMS paradox", which describes a phenomenon where a maximum DMS concentration is commonly detected in low latitude waters when phytoplankton biomass is low (Toole and Siegel, 2004; Vallina et al., 2008)." This is not the summer DMS paradox (a term, by the way, suggested by Simo & Pedros-Alio Nature 1999), which states that the annual maximum of surface DMS commonly occurs in summer, even at the mid and subtropical latitudes where chlorophyll-a (chl-a) is at its annual minimum.

L34-35: "Simó and Dachs (2002) achieved a strong relationship between heavily binned and averaged DMS data and mixed layer depth (MLD)." This is not true. Simo & Dachs (2002) correlated DMS to the MLD and to chl-a/MLD, depending on a chl-a/MLD threshold.

L53-54: “Many provinces lacked adequate data to create a reliable climatology (Fig. A1). In those situations, temporal interpolations were used to fill the blanks, and to create a first-guess map.” This was done where monthly data gaps existed to complete the seasonality. Where data were lacking to even outline a seasonality, this was taken from a neighboring province and adjusted to the existing data.

L61: “Since DMS is produced by marine that algae. . .” This is totally outdated. There are tens of papers showing that this is an oversimplification. DMSP is mainly produced by marine algae, and it is transformed into DMS by marine algae, bacteria and with involvement of zooplankton.

L93-94: “We do not log-transform SST to avoid losing data with temperature below (equal to) zero.” You may have other reasons to not log transform SST, but not this one. A common practice to log transform SST if desired is to convert it to K (Kelvin) first.

If I understand it correctly, you use chl-a data where available, otherwise you take it from SeaWiFS. What efforts have you done to reconcile in situ with satellite chl-a? It is well known that algorithms for satellite estimates of chl-a are developed and calibrated against HPLC chl-a, and there is an important shift between this and Turner fluorometric chl-a. Therefore, putting together in situ (Turner, perhaps HPLC too?) and satellite chl-a data will mess up your statistics.

Calculation of air-sea fluxes: I agree that Nightingale 2000 is quite a standard. But, why not using a more updated linear relationship of Kw to u10? Marandino proposed one with one of the coauthors. Also, you use monthly means of wind speed. Since you are using a nonlinear dependence of Kw on the u10, how do you deal with the fact that a mean u10 will not give the same result as a mean Kw?

L170-176: It reads as though you did not know of the existence of Gali et al. BGS 2016.

L182: “On the other hand, negative correlations between DMS and Chl a have also been detected in coastal waters of the Mediterranean and in the Sargasso Sea (Toole and Siegel, 2004).” Toole & Siegel did not do anything with Med Sea data. The original data from the Sargasso Sea were from Dacey et al DSR 1996, and data from the coastal Med Sea were reported by Vila-Costa et al. LO 2008.

L185-190: This is a very poor interpretation of the DMS vs MLD coupling, and a misuse of the original relationship suggested by Simo & Dachs GBC 2002. As a matter of fact, you cite Simo & Pedros-Alio GBC 1999 because they brought it up for the first time, but the occurrence of a negative relationship between DMS and MLD over large regions of the global ocean was reported by Simo & Dachs. However, the relationship was logarithmic, DMS = aÂůLn(MLD) + b, and there are reasons for this to occur, related to exposure to solar radiation. Trying to correlate DMS directly to MLD (or in a log-log manner) is not expected to provide good prediction.

L189-199: There are a number of papers that should be cited here – besides Toole et al. and Sunda et al, sevarl papers by Marti Gali deal exactly with the effects of solar radiation, and particularly UV, on enhancing DMS production and concentration.

“Climatological PAR is the second strongest predictor (R2 = 0.12, n = 54,683) of raw DMS data with a positive correlation. (. . .) Strong correlation between monthly binned and averaged solar radiation dose (SRD) and DMS concentration has been reported (R2 = 0.94) at the Blanes Bay Microbial Observatory located in the coast of northwest Mediterranean (Vallina and Simó, 2007).” Again, you compare your statistics with that of a previous study, but applying a different calculation. According to the methods description, you used monthly PAR, i.e., monthly surface irradiance. Vallina & Simo 2007, conversely, computed what they called the solar radiation dose, which is the daily averaged solar radiation integral in the mixed layer. This is very different from surface irradiance, because it takes into account the mixed layer depth (and a median light attenuation coefficient). Later on, in L200-211, you infer that, contrasting to Vallina & Simo, you did not get a good correlation to light, and attribute it to the number of original data and to data binning. But you did not use the same light metrics as the other authors, and ignored the arguments given by V&S to use the SRD instead of the surface irradiance, and ignoring the Gali & Simo GBC 2015 meta-analysis too.

L201: “Simó and Dachs (2002) obtained a high R2 value between DMS concentration and the ratio of Chl a and MLD (Chl/MLD).” This is not true. As already mentioned above, the Simo & Dachs (2002) model correlated DMS to the MLD (logarithmic) and to chl-a/MLD (linear), depending on a chl-a/MLD threshold.

All in all, if you are to compare your statistics to those of S&D 2002 and V&S 2007, everything here has to be recomputed and rewritten.

The arguments against binning the data are poor. It is true that binning reduces the variance, but you are using monthly climatologies (heavily averaged and also binned) to relate raw DMS data to potential predictors. Also, binning must be used if you want to avoid giving too much predictive weight to the regions thoroughly sampled over the undersampled. This is becoming more important as we are bringing in underway data at unprecedented spatial resolution, like the NAAMES data incorporated here.

L231-233: “The summertime high DMS concentration at high latitudes is consistent with the hypothesis that phytoplankton use DMSP as a cryoprotectant (Karsten et al., 1992). It is found that the same phytoplankton (Antarctic macroalga) contains higher DMSP concentration in the polar regions than in the temperate regions (Karsten et al., 1990).” Poor again, if not wrong. See recent papers on DMS in polar regions (e.g. Webb et al Sci Rep 2018, Gali et al. PNAS 2019). And macroalgae are not phytoplankton.

Subsequent discussion: The seasonality and geographic distribution of DMS have been profusely (and much better) discussed by Lana et al. GBC 2011 and others, including regional studies. You should rather focus on new features unveiled with respect to others, particularly Lana 2011.

L305-306: “By contrast, objective interpolation methods are spatial/temporal averages of sparse data with no underlying basis in environmental variability.“ Again, this is not totally true. In Lana et al. 2011, to create a first guess field, biogeographic provinces were used, which is an informed approach to extrapolation. These provinces are defined from environmental descriptors. And a distance weighted interpolation from original data was used for interpolation.

Figure 2: An annual average is not very informative. I would even argue it is misleading in the case of highly seasonal variables like DMS, because summer maxs and winter mins cancel out each other. I would recommend splitting the map into two or four seasons to show hemispherical patterns.

Figure 4: Some differences are outstanding but you do not discuss them. For instance, Lana 2011 captures the September max of DMS concentration in the subarctic NE Pacific, because it is well covered with data. Conversely, your ANN does not capture it. This warrants some discussion, as it will reveal some of the caveats of the ANN approach.

In summary, I think that the ANN is an interesting approach that will help improve DMS (and other) climatologies, especially where data are lacking, as it will do better than inter- and extrapolations. However, the present manuscript does not go much beyond the application of the ANN; when it intends to do so, too often it uses the wrong arguments and is not fair with previous studies. It fails to mobilize what we have learned about DMS in the last one or two decades.

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