Review of the manuscript submitted to Biogeosciences Discussions: "Global ocean dimethyl sulfide climatology estimated from observations and an artificial neural network", by Wei-Lei Wang, Guisheng Song, François Primeau, Eric S. Saltzman, Thomas G. Bell, and J. Keith Moore.

General comments

The manuscript by Wang and coauthors proposes an interesting methodological development: the use of artificial neural networks (ANN) to produce a global gridded climatology of dimethylsulfide (DMS) concentration at the sea-surface. This is a relevant topic because DMS emission drives aerosol formation in the remote marine atmosphere, with subsequent effects on aerosol and cloud radiative forcing. Measurements of sea-surface DMS concentration are too sparse to be directly usable in studies of atmospheric chemistry and global sulfur biogeochemistry. Therefore, a number of techniques have been used in the past to produce global gridded DMS fields: from objective interpolation (on which the standard climatological product is based; Lana et al., 2011) to empirical remote sensing algorithms or prognostic ocean biogeochemistry models.

The "artificial intelligence" approach proposed by Wang and coauthors is a necessary step to improve existing global DMS products. The article is generally well written and I appreciated the succinct style. However, in its current form the study suffers from a number of important shortcomings:

1. Repetition of results that have been presented more in depth in previous papers. The results of the correlation analysis between DMS(P) and environmental variables are of little interest as they are extremely similar to those reported in previous papers, where they were analyzed more in depth and with a more solid theoretical underpinning. The stepwise multilinear regression (sMLR), which is used mainly to contrast its limited predictive power against the greater predictive power of the ANN, is only partially described.
2. Failure to perform appropriate quality control of the raw DMS, DMSP and chlorophyll (Chl) data, for example following procedures detailed by Galí et al. 2015 for the same global DMS database used by Wang et al. The main flaw is the use of in situ fluorometric Chl measurements and satellite-retrieved Chl as if they were equivalent –they are not.
3. Inaccurate reasoning regarding the utility of data binning for the purpose of calculating monthly climatologies. What is the value of using raw (non-binned) measurements if (1) most of them are matched to climatological fields of the predictor variables? and (2) the final purpose is calculating a monthly climatology at coarser spatial resolution, which by definition aims at smoothing out interannual and small-scale variability? For example, to what extent the introduction of more than 10,000 new measurements, taken at high resolution in a relatively small region that was already quite well documented (temperate NW Atlantic), adds relevant information when it comes to computing monthly climatologies? Wouldn't it be more appropriate to bin all measurements beforehand to the coarsest resolution at which predictor variables are available, and then train the ANN? The authors should treat these issues more accurately and provide evidence for the advantages (if any) of using raw DMS data (including high-resolution transects), e.g. comparing statistics for ANN trained on raw vs. binned data. This being said, I agree that capturing the weight of extremes, ie the non-normal statistical distribution of DMS, is important (Galí et al., 2018).
4. Limited discussion of the advantages of the ANN approach, especially in regions that are challenging for prognostic and empirical models. For example, the ANN method does not outperform the gridded climatology (Lana et al., 2011) in the subarctic Northeast Pacific in August and September, when DMS concentrations are much higher than what one would expect based on global-scale relationships. If the ANN does not outperform the (admittedly limited) objective interpolation approach in a region that contains data, how can we trust ANN predictions in regions with no DMS measurements? An analysis comparing seasonal

DMS patterns across different biogeochemical regimes (eg ANN vs. objective interpolation and remote sensing algorithms) would be very welcome and would strengthen the arguments for adopting the ANN as a standard method to compute climatologies.

5. Misuse or inappropriate citation of some key references (e.g. Simó & Dachs 2002, Toole and Siegel 2004) and, more generally, omission of relevant references from the past 10-20 years. I think the view of marine DMS(P) cycling presented in this article is a bit outdated, especially regarding (1) upper-ocean DMS budgets and DMS turnover times due to biological processes, which ultimately control concentrations, and (2) the role of heterotrophic organisms and processes, which decouple DMS from phytoplankton abundance and taxonomy in much of the global ocean.

In addition to the points above, I have to admit that approaches such as ANN leave me, as a reader, with the feeling I did not learn much about the global distribution of DMS and its controlling factors. This would not necessarily be a criticism as long as the choices the authors made to configure and tune the ANN were sufficiently justified, but I missed some depth of information in this regard. The ANN itself remains a mysterious black box to me and, even if the overall results look reasonable (with exceptions, as highlighted above), I am unable to appreciate whether Wang and coauthors made an optimal implementation of the ANN.

I prompt the authors to address the formal, conceptual and technical criticisms made above. I honestly hope these constructive criticisms will improve the study and, more broadly, pave the way towards sensible implementation of AI techniques to compute DMS climatologies (which will likely become routine in the near future).

Finally, note that in the specific comments below I will frequently refer to my own papers, simply because some of them are very relevant for the present study and, in some cases, the only ones available. Of course, the authors are free to decide what citations they incorporate. For all these reasons, and for the sake of transparency, I decided to sign the review.

Martí Galí

Specific comments

Abstract

Please reshape taking into account the main criticisms, especially concerning the amount of variance captured when using raw or binned data (general point 3). Raw DMS data variance could be biased towards high-resolution data representative of small-scale variability if no homogenization of the spatial-temporal scales covered by the measurements is applied (the PMEL database consists mostly of coarse resolution data).

Introduction

L19: OK, but the approach proposed here does not reveal the factors controlling DMS variability. Rephrasing suggested.

L24: If the authors insist on the mechanistic point of view (not sure is the right line of thought in this paper), I suggest adding “process rate measurements” here. They hamper predictive models even more strongly than limited observations of DMS concentration.

L30: To the best of my knowledge, the term “summer paradox” was coined by Simó & Pedrós-Alió 1999 (Nature), so I suggest crediting them for it.

L49: Relevant citations here are Le Clainche et al. 2010 (GBC) (S cycling model inter-comparison seeking to understand the processes responsible for the summer paradox) and Tesdal et al. 2016

(Env Chem) (the most extensive comparison among gridded climatologies, empirical and prognostic models published so far, to my knowledge).

L61: DMS is produced by some marine algae, some bacteria, and mostly as a result of food web interactions (Kiene et al. 2000; Simó et al. 2001, Stefels et al. 2007, Curson et al. 2011, Moran et al. 2012, etc.). Please nuance and refine.

Methods

L73: Over 10,000 measurements came from NAAMES alone. Not binning the data might give too much weight to particular conditions sampled during NAAMES, at least in the multilinear regression.

L75: The DMSPt, Chla, SST and SSS data in the PMEL database require some quality control. This was documented by Galí et al. 2015. Quality controlled datasets with stringent satellite match-up criteria (ie minimizing the use of climatological coarse resolution data), as well as a piece of code used to clean data, are publicly available on github: https://github.com/mgali/DMS-SAT_DATA_DEV_VAL.

L75: Fluorometric Chl is on average around 40% higher than HPLC Chl (Sathyendranath et al., 2009), and the proportion sometimes varies quite a bit. Satellite Chl is validated against HPLC measurements (e.g. Morel et al., 2007).

L77: Less than 3.5 years do not make a good Chl climatology in many ocean areas in my experience. Data products covering much longer periods are available on NASA's ocean colour website. Please update datasets, and specify also what reprocessing was used.

L79: The more recent climatology of Holte et al. 2017 seems to outperform that of Schmidtko et al. 2013 in areas of deep winter convection (subpolar North Atlantic) or where deep mixing prevails (Southern Ocean circumpolar current). In some cases the differences are important. Please consider using the Holte et al. 2017 MLD climatology.

L81: Please specify the nutrient datasets used.

L105-107: I see some contradiction here. Data extremes typically arising from nonlinear dynamics are often smoothed out when averaging data. Your predicted variable (DMS) retains full variability but predictor fields do not, because apart from SST they largely originate from monthly climatologies. How can meaningful nonlinear relationships be identified?

L128-130: Are these parameters default ones, or tuned manually to achieve reasonable fits in this particular study?

L137: Inclusion of time of day is interesting, although diel variability was not been mentioned earlier in the manuscript. Are hourly predictions useful for computing climatologies? Although DMS can oscillate on diel time scales (Galí et al., 2013; Royer et al., 2016) diel cycles do not seem to follow a fixed pattern, at least in low-latitude high-resolution datasets (e.g. Royer et al., 2015).

Results

L170 and 177: Note Galí et al. (2018) reported an R2 of 0.42 (r of 0.65) between DMS and DMSPt using the same datasets with stricter quality control. Similarly for DMS vs. Chl: R2 of 0.20 (r of 0.45).

L172-176: These sentences look a bit contradictory and may need further elaboration (or else, can be removed). Is it straightforward or not to predict DMS from DMSPt? Are measurements sufficient or not? Regarding DMSP prediction, another relevant study is that by McParland & Levine (2018). Regarding the relationship between DMSPt and DMS in the global PMEL database, Galí et al. (2018) is a relevant reference.

L179: The weaker relationship between DMS and Chl in the entire dataset probably results from the higher proportion of oligotrophic low latitude data (where DMS is anticorrelated to Chl over the seasonal cycle) compared to in situ Chl-DMS data pairs. The difference between in situ fluorometric Chl and satellite Chl may also play a role. Finally, note that the global PMEL DMS database is biased

towards productive conditions (Galí et al. 2018; figure 7) which influences global DMS-Chl correlations. In summary, the correlation between DMS and Chl in global datasets is not really meaningful as it depends strongly on how evenly represented are the different ocean biomes.

L187-190: This is incorrect. Dilution is not the main explanation for the negative relationship between MLD and DMS (as originally proposed by Aranami and Tsunogai, 2004). The main explanation is the different balance between biological DMS sources and sinks, as explained by Galí & Simó (2015). In the handful of studies that have made DMS budgets including the vertical mixing term, vertical DMS transport has never been found to dominate DMS budgets in the MLD over relevant (~daily) time scales. Check for example Bailey et al. 2008 (DSR), Herrmann et al. 2012 (CSR), Galí et al. 2013 (GBC), Royer et al. 2016 (Sci Rep), etc. DMS turnover in the surface layer due to vertical transport is generally an order of magnitude slower than biological turnover or biological + photochemical turnover, at least. Please correct.

L191-199: The strongest evidence for light-driven DMS production in natural plankton assemblages comes from recent work by myself and colleagues (Galí et al. 2013a, b, c; Royer et al. 2016). Evidence for light-driven DMS production in Toole et al. (2006) (otherwise, a great piece of work!) is indirect as that study focused on DMS removal processes.

Section 3.2: see general comment 3 on data binning.

L201-204: It is unclear here if the authors made the appropriate comparisons with Simó & Dachs 2002 and Vallina & Simó 2007 empirical models. Was DMS compared with surface PAR or with the solar radiation dose in the mixed layer as done by Vallina and Simó 2007? Similarly, did the authors correctly apply the double algorithm used by Simó and Dachs 2002, where different equations are used depending on the value of Chl/MLD? Or just computed a single regression of DMS against Chl/MLD?

Section 3.3: Methods section mentions 8 initial variables (PAR, MLD, Chl a, SSS, SST, DIN, DIP, and SiO), but, what predictor variables were included in the final multilinear regression model (MLR)? Does the R2 of the MLR refer to linear of log space?

Section 3.4: Since this subsection describes the main technical innovation of this paper, a deeper explanation of why the ANN gives these results would be very welcome. See general comment 4.

Section 3.5: For the authors' information, global DMS fields produced with the remote sensing algorithm of Galí et al. 2018, as well as the algorithms of Simó & Dachs 2002 and Vallina & Simó 2007, are available in this repository: https://doi.org/10.5281/zenodo.2558511. Corresponding Matlab and R codes are available on a linked github repository:

https://github.com/mgali/DMS-SAT_ALGORITHM.

L235: These references are not appropriate here. Please cite studies that actually documented DMS(P) dynamics in subpolar or polar blooms of coccolithophores or *Phaeocystis*.

L247-248: I do not find reasonable that DMS decreases below 0.1 nM in a subtropical gyre in winter. By examining the maps in Fig. 3 I would say ANN DMS is mostly between 0.1 and 0.5 nM, which still looks a bit low but more realistic according to my experience. DMS concentrations lower than 0.1 nM are extremely rare in both the PMEL database and in global estimates made with empirical algorithms (see Galí et al. 2018 figure 7).

L250-254: Please check Galí and Simó 2015 (GBC) for a mechanistic explanation of the summer paradox.

In general, I suggest making a figure showing the climatological seasonal cycles in different ocean biomes or regions, to better support the description made in the text.

L273-293: Here I strongly suggest citing Tesdal et al. 2016.

Figures

Figure 3 and 5: I suggest using a color scale with different colors to help readers appreciate concentration patterns.

Figure 6: I strongly recommend splitting results into northern and southern hemisphere given the strong seasonality of DMS (also wind speed and SST), which results in opposed seasonal patterns.

Minor corrections

L131: What does “epochs” mean in this context? Please use synonym for readers that are not expert in ANN or similar techniques.

L218: The “tracer-tracer” term is not needed here (quite specific to bgc modelling).

Reviewer references (only if not cited by the authors)

Aranami, K., & Tsunogai, S. (2004). Seasonal and regional comparison of oceanic and atmospheric dimethylsulfide in the northern North Pacific: Dilution effects on its concentration during winter. *Journal of Geophysical Research: Atmospheres*, *109*(D12).

Bailey, K. E., Toole, D. A., Blomquist, B., Najjar, R. G., Huebert, B., Kieber, D. J., ... & Del Valle, D. A. (2008). Dimethylsulfide production in Sargasso Sea eddies. *Deep Sea Research Part II: Topical Studies in Oceanography*, *55*(10-13), 1491-1504.

Curson, A. R., Todd, J. D., Sullivan, M. J., & Johnston, A. W. (2011). Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nature Reviews Microbiology*, *9*(12), 849-859.

Galí, M., Ruiz-González, C., Lefort, T., Gasol, J. M., Cardelús, C., Romera-Castillo, C., & Simó, R. (2013). Spectral irradiance dependence of sunlight effects on plankton dimethylsulfide production. *Limnology and oceanography*, *58*(2), 489-504.

Galí, M., Simó, R., Vila-Costa, M., Ruiz-González, C., Gasol, J. M., & Matrai, P. (2013). Diel patterns of oceanic dimethylsulfide (DMS) cycling: Microbial and physical drivers. *Global Biogeochemical Cycles*, *27*(3), 620-636.

Galí, M., Simó, R., Pérez, G., Ruiz Gonzalez, C., Sarmento, H., Royer, S. J., ... & Gasol, J. M. (2013). Differential response of planktonic primary, bacterial, and dimethylsulfide production rates to static vs. dynamic light exposure in upper mixed-layer summer sea waters.

Galí, M., & Simó, R. (2015). A meta-analysis of oceanic DMS and DMSP cycling processes: Disentangling the summer paradox. *Global Biogeochemical Cycles*, *29*(4), 496-515.

Herrmann, M., Najjar, R. G., Neeley, A. R., Vila-Costa, M., Dacey, J. W., DiTullio, G. R., ... & Vernet, M. (2012). Diagnostic modeling of dimethylsulfide production in coastal water west of the Antarctic Peninsula. *Continental Shelf Research*, *32*, 96-109.

Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An Argo mixed layer climatology and database. *Geophysical Research Letters*, *44*(11), 5618-5626.

Kiene, R. P., Linn, L. J., & Bruton, J. A. (2000). New and important roles for DMSP in marine microbial communities. *Journal of Sea Research*, *43*(3-4), 209-224.

Le Clainche, Y., Vézina, A., Levasseur, M., Cropp, R. A., Gunson, J. R., Vallina, S. M., ... & Bopp, L. (2010). A first appraisal of prognostic ocean DMS models and prospects for their use in climate models. *Global biogeochemical cycles*, *24*(3).

Moran, M. A., Reisch, C. R., Kiene, R. P., & Whitman, W. B. (2012). Genomic insights into bacterial DMSP transformations. *Annual review of marine science*, *4*, 523-542.

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., & Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sensing of Environment*, *111*(1), 69-88.

Royer, S. J., Mahajan, A. S., Galí, M., Saltzman, E., & Simó, R. (2015). Small-scale variability patterns of DMS and phytoplankton in surface waters of the tropical and subtropical Atlantic, Indian, and Pacific Oceans. *Geophysical Research Letters*, *42*(2), 475-483.

Royer, S. J., Galí, M., Mahajan, A. S., Ross, O. N., Pérez, G. L., Saltzman, E. S., & Simó, R. (2016). A high-resolution time-depth view of dimethylsulphide cycling in the surface sea. *Scientific reports*, *6*, 32325.

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., ... & Platt, T. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Marine Ecology Progress Series*, *383*, 73-84.

Simó, R., & Pedrós-Alió, C. (1999). Role of vertical mixing in controlling the oceanic production of dimethyl sulphide. *Nature*, *402*(6760), 396-399.

Tesdal, J. E., Christian, J. R., Monahan, A. H., & von Salzen, K. (2016). Evaluation of diverse approaches for estimating sea-surface DMS concentration and air–sea exchange at global scale. *Environmental Chemistry*, *13*(2), 390-412.